

# GEBF-Abstractband

## »Die Perspektiven verbinden«

2. Tagung der GEBF in Frankfurt am Main

03.-05. März 2014



## **Kompetenzveränderung im Längsschnitt**

Chair(s): Juliane Grünkorn (DIPF), Jens Fleischer (Universität Duisburg-Essen), Eckhard Klieme (DIPF)

Diskutant/in: Detlev Leutner (Universität Duisburg-Essen)

Die Forschung zur theoriegeleiteten Modellierung und Messung von Kompetenzen im Bildungsbereich hat in den letzten Jahren große Fortschritte erzielt. So wurden in zahlreichen Wissensdomänen kognitionspsychologisch und fachlich fundierte Kompetenzstruktur- und Kompetenzniveaumodelle entwickelt und darauf aufbauend geeignete Messinstrumente zur empirischen Überprüfung dieser Modelle konstruiert. Um darüber hinaus Auskünfte über zeitliche Entwicklungsverläufe zu geben und damit wichtige Informationen z. B. für unterrichtliche Förderung bereitzustellen, werden Kompetenzentwicklungsmodelle benötigt. Diese lassen sich aus Kompetenzstruktur- und Kompetenzniveaumodellen, die in der Regel in querschnittlich angelegten Studien untersucht werden, nicht direkt ableiten. Die Bemühungen zur Entwicklung theoretisch und empirisch hinreichend fundierter sog. *Learning Progression Models* bilden auch international derzeit einen Fokus des *Educational Measurement*. Von zentraler Bedeutung sind Fragen, wie z. B.: Inwiefern verändern sich Kompetenzstrukturen über die Zeit? Wie können diese Veränderungen adäquat gemessen werden? Welche Implikationen ergeben sich daraus für die gezielte Förderung von Kompetenzen in Schule und Ausbildung?

Das Symposium umfasst vier Beiträge aus Forschungsprojekten des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“. Sie bearbeiten Fragen der zeitlichen Entwicklung von Kompetenzen in verschiedenen Domänen und aus der Perspektive unterschiedlicher Fachdisziplinen wie der Psychologie, der Erziehungswissenschaft, der Naturwissenschaftsdidaktik, der Literaturdidaktik sowie der Berufspädagogik. Der erste Beitrag von Neumann befasst sich mit der Validierung eines Kompetenzmodells zur Entwicklung des Energieverständnisses im Physikunterricht der Sekundarstufe I. Er verbindet quantitative und qualitative Forschungsansätze und liefert Ergebnisse, die zur Optimierung des Physikunterrichts genutzt werden können. Eggert, Bögeholz, Watermann und Hasselhorn untersuchen im zweiten Beitrag Möglichkeiten der gezielten Förderung von Bewertungskompetenz im Regelunterricht der Jahrgangsstufe 7. Sie kombinieren dazu eine kooperative Lernumgebung mit einem zusätzlichen metakognitiven Training und entwickeln ein Testinstrument, welches zur Veränderungsmessung der Bewertungskompetenz geeignet ist. Der dritte Beitrag von Roick, Meier, Henschel und Frederking befasst sich mit der Struktur und Veränderung von literarischer Textverstehenskompetenz zum Ende der Sekundarstufe I. Sie liefern Anhaltspunkte, wie Entwicklungspotenziale hinsichtlich literarischen Textverstehens im Deutschunterricht besser zu nutzen sind. Schließlich stellen Nickolaus, Schmidt und Gschwendtner im vierten Beitrag eine Studie zur Veränderungen der Struktur berufsfachlicher Kompetenzen im Verlauf der Ausbildung zum Kfz-Mechatroniker bzw. zur Kfz-Mechatronikerin vor. Sie liefern Hinweise darauf, dass sich die Kompetenzstruktur im Verlauf der Ausbildungszeit sowohl ausdifferenziert als auch verschmilzt und begründen dies mit der curricularen Schwerpunktsetzung sowie mit unterschiedlichen Anforderungskontexten.

Montag,  
03.03.2014,  
11:15 - 13:00,  
Casino-Gebäude  
Raum 823

**Die Entwicklung physikalischer Kompetenz in der Sekundarstufe I am Beispiel des Energiekonzepts**

Die Vermittlung einer naturwissenschaftlichen Grundbildung hat sich national und international als zentrales Ziel des Unterrichts in den Naturwissenschaften durchgesetzt (vgl. Waddington, Nentwig & Schanze, 2007). Statt eines isolierten Faktenwissens sollen die Schülerinnen und Schüler nun ein Verständnis der zentralen Konzepte der Naturwissenschaften erwerben (z. B. KMK, 2005; NRC, 2012). Bisher fehlt es jedoch an Modellen, die die systematische Entwicklung eines solchen Verständnisses im Unterricht in den jeweiligen Fächern erlauben würden. Ziel des vorliegenden Projekts war die Entwicklung und empirische Prüfung eines solchen Modells für das Energiekonzept und darauf aufbauend die Untersuchung der Entwicklung des Energieverständnisses im Physikunterricht der Sekundarstufe I.

Theoretischer Hintergrund: Bisherige Untersuchungen weisen darauf hin, dass sich das Energieverständnis von Schülerinnen und Schülern der Sekundarstufe I entlang von vier Konzeptualisierungen entwickelt: 1) Energie kann in verschiedenen Formen auftreten, 2) Energie kann von einer Form in eine andere umgewandelt werden, 3) Energie besitzt einen Nutzwert und 4) Energie ist eine Erhaltungsgröße (Dawson-Tunik, 2006; Lee & Liu, 2010; Liu & McKeough, 2005). Unklar ist bisher jedoch, wie Schülerinnen und Schüler ein Verständnis der einzelnen Konzeptualisierungen erwerben. In der Vergangenheit wurde allerdings verschiedentlich vermutet, dass die Entwicklung konzeptuellen Verständnisses mit der Entwicklung einer zunehmend komplexeren Wissensbasis einhergeht (z. B. Alao & Guthrie, 1999; für eine Übersicht siehe Bransford, Brown & Cocking, 2000). Kauertz (2008) gelang es zu zeigen, dass sich die Komplexität des zur Bearbeitung physikalischer Aufgaben zu zentralen physikalischen Konzepten erforderlichen Wissens zur Vorhersage der Aufgabenschwierigkeit eignet. Auf dieser Basis wurde im Rahmen des vorliegenden Projekts ein Modell der Entwicklung des Energieverständnisses entwickelt. Es nimmt an, dass sich das Energieverständnis entlang der vier Konzeptualisierungen entwickelt und dass die Schülerinnen und Schüler im Verständnis der einzelnen Konzeptualisierungen voranschreiten. Letzteres indem sie ein zunehmend komplexeres Wissen bzgl. der Konzeptualisierungen entlang von vier Komplexitätsniveaus (Fakten, Wissen über Zusammenhänge, Zusammenhangswissen, Konzeptuelles Verständnis) entwickeln.

Fragestellung: Die beiden zentralen Fragestellungen des Projekts lauteten: 1) Inwieweit eignet sich das entwickelte Modell zur Beschreibung der Entwicklung des Energieverständnisses in der Sekundarstufe I und 2) Wie hängt die Entwicklung des Energieverständnisses von der Strukturierung der Inhalte in der Sekundarstufe I ab?

Methode: Das Projekt war als kombinierter Quer- und Längsschnitt über die Sekundarstufe I angelegt. In einem ersten Schritt wurde das Modell der Entwicklung des Energieverständnisses durch Aufgaben operationalisiert. D. h. zu jeder Kombination von vier Konzeptualisierungen und vier Komplexitätsstufen wurden Aufgaben, insgesamt 120, konstruiert. Diese wurden N = 1856 Schülerinnen und Schülern der Jahrgänge 6, 8 und 10 vorgelegt, um zu untersuchen, inwieweit sich das entwickelte Modell zur Beschreibung der Entwicklung des Energieverständnisses eignet. In einem anschließenden Längsschnitt wurde die Entwicklung des Energieverständnisses von N = 504 Schülerinnen und Schülern ausgewählter Schulen vom 6. bis 9. Jahrgang erhoben und als Funktion der schulspezifischen Curricula untersucht.

Ergebnisse: Die Ergebnisse der Querschnittsuntersuchungen bestätigen im Wesentlichen die Modellannahmen. So konnte eine statistisch bedeutsame Abhängigkeit der Aufgabenschwierigkeit von der Konzeptualisierung festgestellt werden,  $F(3,98) = 12.58$ ,  $p < .001$ ,  $\eta^2 = 0.28$ . Zudem zeigte sich eine Abhängigkeit der Personenfähigkeit vom Jahrgang (als Proxy für Lerngelegenheiten),  $F(2,1853) = 161.29$ ,  $p < .001$ ,  $\eta^2 = .15$ . Ein Effekt der Komplexität auf die Aufgabenschwierigkeit konnte zunächst nicht nachgewiesen werden. Eine Re-Analyse der Daten, bei der unterschiedliche Modelle des Wissenserwerbs verglichen wurden, zeigte, dass das Modell eines kumulativen Wissenserwerbs dem Modell der Entwicklung eines integrierten Wissens unterlegen ist ( $AIC_{\text{kum}} = 42619 < AIC_{\text{int}} = 42924$ ). Die (qualitative) Auswertung der längsschnittlichen Entwicklung des Energieverständnisses im Vergleich unterschiedlicher Schulen liefert darüber hinaus Hinweise darauf, dass die Strukturierung der Inhalte einen wesentlichen Einfluss auf die Entwicklung des Energieverständnisses hat. In der Summe bietet das Projekt damit wichtige Erkenntnisse über die Entwicklung des Energieverständnisses und damit für die Optimierung des Physikunterrichts.

Alao, S. & Guthrie, J. T. (1999). Predicting conceptual understanding with cognitive and motivational variables. *The Journal of Educational Research*, 92, 243–254.

Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*. Washington, D. C.: National Academy Press.

Dawson-Tunik, T. L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 111–136). Maple Grove, USA: JAM Press.

Kauertz, A. (2008). *Schwierigkeitserzeugende Merkmale von Physikaufgaben*. Berlin: Logos.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] (2005). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.

Lee H.-S. & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, 94(4), 665–688.

Liu, X. & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching*, 42(5), 493–517.

National Research Council [NRC] (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D. C.: The National Academies Press.

Waddington, D., Nentwig, P. & Schanze, S. (2007). *Making it comparable. Standards in science education*. Münster: Waxmann.

Montag,

03.03.2014,

11:15 - 13:00,

Casino-Gebäude

Raum 823



## **Förderung von Bewertungskompetenz bei Schülerinnen und Schülern im Biologieunterricht durch zusätzliche metakognitive Strukturierungshilfen – Ein Beispiel für Veränderungsmessung**

Theoretischer Hintergrund: Seit Verabschiedung der nationalen Bildungsstandards ist Bewertungskompetenz einer von vier vorgeschriebenen Kompetenzbereichen in den naturwissenschaftlichen Fächern (KMK, 2004). Lernende benötigen Bewertungskompetenz insbesondere für die Bearbeitung von Themen moderner Naturwissenschaften, z. B. einer nachhaltigen Entwicklung unserer Umwelt. Derartige Umweltproblemsituationen („socioscientific issues“) sind naturwissenschaftlich und ethisch komplex, haben keine optimale Lösung, müssen multiperspektivisch bearbeitet werden und basieren z. T. auf unsicherem oder sich widersprechendem Wissen (z. B. Eggert & Bögeholz, 2010; Kolsto, 2001; Ratcliffe & Grace, 2003; Sadler, Barab & Scott, 2007). Für die Bearbeitung derartiger Themen im naturwissenschaftlichen Unterricht bieten sich kooperative Lehr-Lernarrangements an. Für das Gelingen kooperativer Lernumgebungen ist jedoch der Grad der (Vor-)strukturierung von Lernprozessen zentral (Huber, 1995; Sumfleth, Nicolai & Rumann, 2004). Eine Möglichkeit kooperative Lernumgebungen zusätzlich zu strukturieren, ist der Einsatz metakognitiver Trainings beispielsweise durch den Einsatz von Fragetechniken (z. B. Mevarech & Kramarski, 2003). Ziel der Studie war, einerseits den Effekt eines kooperativen Lehr-Lernarrangements mit integriertem metakognitiven Training auf die Bewertungskompetenz der Lernenden zu analysieren. Andererseits wurde der methodischen Frage nachgegangen, inwiefern das eingesetzte Testinstrument Kompetenzzuwächse im Sinne von Veränderungsmessung erfassen kann.

Die Forschungsfragen sind: Inwiefern hat ein Treatment Kooperativen Lernens mit integriertem metakognitiven Training einen stärkeren positiven Einfluss auf die Bewertungskompetenz von Lernenden als ein kooperatives Treatment allein? Inwiefern kann die Qualität der Bewertungskompetenz der Lernenden zum zweiten Messzeitpunkt als ein Lernzuwachs im Vergleich zum ersten Messzeitpunkt interpretiert werden?

Methode: 258 Schülerinnen und Schüler der Jahrgangsstufe 7 an Gymnasien (Durchschnittsalter: 12.84; 128 Mädchen) nahmen an der Interventionsstudie teil. 118 Lernende arbeiteten in einer kooperativen Lernumgebung (KOOP), 140 Lernende in derselben kooperativen Lernumgebung mit zusätzlichem metakognitiven Training (KOOP META). Die Unterrichtseinheit zum Thema Fließgewässerbewertung und -gestaltung wurde zuvor mit Lehrkräften entwickelt, die zufällig einer der beiden Varianten zugeordnet wurden und diese unterrichteten. Der Lernzuwachs wurde in einem Prä-Posttest Kontrollgruppendesign mittels eines paper-and-pencil Tests zu Bewertungskompetenz erfasst (WLE-Reliabilitäten Prätest = .73; Posttest = .85). Die Auswertung erfolgte mit dem Rasch Partial Credit Model. Für den Vergleich der Personenfähigkeiten zu beiden Messzeitpunkten wurden die Itemschwierigkeiten auf Messinvarianz geprüft. Dies erfolgte mittels differential item functioning (dif) Analysen. Für den Treatmentvergleich (KOOP vs. KOOP META) wurden die Personenparameter varianzanalytisch ausgewertet.

Ergebnisse: Im Rahmen der dif Analysen wurden die Itemschwierigkeiten beider Messzeitpunkte mittels eines grafischen Modelltests verglichen. Dabei wurden einige Items bzw. Itemsteps als problematisch hinsichtlich dif ermittelt. Eine Analyse der Logitdifferenzen bestätigte dies (vgl. Wilson, 2005). Einige Items wurden zum zweiten Messzeitpunkt unverhältnismäßig einfach, wie z. B. das Benennen von Pro- und Kontraargumenten bei der Bewertung von Handlungsoptionen. Andere Items wurden zum zweiten Messzeitpunkt unverhältnismäßig schwierig, wie z. B. das kritische Überprüfen von Bewertungs- und Entscheidungsprozessen. Diese Items bzw. Itemsteps wurden bei der Schätzung der Personenfähigkeiten frei geschätzt und trugen nicht zur Analyse des Kompetenzzuwachses bei. Für den anschließenden Treatmentvergleich zeigte eine zweifaktorielle Varianzanalyse mit Messwiederholung signifikante Haupteffekte der Zeit ( $F(1,256) = 17.240, p < .001$ ) sowie der Treatmentgruppe ( $F(1,256) = 6.136, p < .05$ ) aber keine signifikante Interaktion. Bei einer zusätzlichen Kovarianzanalyse über die Posttestwerte mit den Prätestwerten als Kovariate wurde das 5 % Signifikanzniveau knapp verfehlt ( $F(1,256) = 3.598, p = .06$ ).

Diskussion: als eingesetzte Messinstrument zu Bewertungskompetenz kann – mit Einschränkungen – für Veränderungsmessungen eingesetzt werden, da zumindest partielle Messinvarianz gegeben ist. Der Ausschluss mehrerer dif Items war jedoch mit einem Informationsverlust bezüglich des Kompetenzzuwachses verbunden. Dennoch bilden die Ergebnisse dieser probabilistischen Auswertung den Kompetenzzuwachs der Lernenden realistischer ab.

Es zeigte sich, dass Lernende in beiden Treatmentvarianten einen Kompetenzzuwachs erfahren haben, ein differentieller Effekt des zusätzlichen metakognitiven Trainings konnte jedoch nicht nachgewiesen werden. Dennoch liefern die Ergebnisse wichtige Erkenntnisse bezüglich der Umsetzung des Trainings im Unterricht sowie einer Optimierung des Trainings von Lehrkräften hinsichtlich der Förderung von Bewertungskompetenz.

Eggert, S. & Bögeholz, S. (2010). Students' Use of Decision-Making Strategies With Regard to Socioscientific Issues – An Application of the Rasch Partial Credit Model. *Science Education*, 94, 230–258.

Huber, G. L. (1995). Lernprozesse in Kleingruppen: Wie kooperieren die Lerner? *Unterrichtswissenschaft*, 23, 316–331.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] (2004). Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss vom 16.12.2004. München: Wolters Kluwer.

Kolsto, S. D. (2001). Scientific Literacy for Citizenship: Tools for Dealing with the Science Dimension of Controversial Socioscientific Issues. *Science Education*, 85, 291–310.

Mevarech, Z. R. & Kramarski, B. (2003). The effects of metacognitive training versus worked-out examples on students' mathematical reasoning. *British Journal of Educational Psychology*, 73, 449–471.

Ratcliffe, M. & Grace, M. (2003). *Science Education for Citizenship. Teaching Socio-Scientific Issues*. Maidenhead: OUP.

Sadler, T., Barab, S. & Scott, B. (2007). What Do Students Gain by Engaging in Socio-scientific Inquiry? *Research in Science Education*, 37, 371–391.

Sumfleth, E., Nicolai, N. & Rumann, S. (2004). Schulische und häusliche Kooperation im Chemieanfängerunterricht. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 284–302). Münster: Waxmann.

Wilson, M. (2005). *Construction Measures: An Item Response Modeling Approach*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.





Thorsten Roick (Humboldt-Universität zu Berlin), Christel Meier (Friedrich-Alexander-Universität Erlangen-Nürnberg), Sofie Henschel (Humboldt-Universität zu Berlin), Volker Frederking (Friedrich-Alexander-Universität Erlangen-Nürnberg)

## **Entwicklung der literarischen Textverstehenskompetenz zum Ende der Sekundarstufe I: Erste Befunde und Schlussfolgerungen**

Theoretischer Hintergrund: Die Vermittlung von Kompetenzen im Umgang mit literarischen Texten nimmt in der Literaturdidaktik (z. B. Schilcher & Pissarek, 2013) wie auch in den Bildungsstandards für das Fach Deutsch in der Sekundarstufe I (KMK, 2004) eine prominente Stellung ein. Lange Zeit fehlten allerdings belastbare empirische Erkenntnisse zu der Frage, ob das Verstehen literarischer und nicht-literarischer Texte voneinander abgrenzbare Kompetenzbereiche darstellen. Im DFG-Projekt „Literarästhetische Urteilskompetenz“ (vgl. Roick, Stanat, Dickhäuser, Frederking, Meier & Steinhauer, 2010) ist der Versuch unternommen worden, einen Beitrag zur Beseitigung dieses Forschungsdesiderats zu leisten. Im Zentrum stand dabei die empirische Überprüfung eines auf der Grundlage von Eco's ästhetischer Semiotik (Eco, 1990) entwickelten Modells literarischer Textverstehenskompetenz, welches zwischen inhaltsbezogenen (semantischen) und formbezogenen (idiolektalen) Verstehensleistungen unterscheidet (z. B. Frederking, Henschel, Meier, Roick, Stanat & Dickhäuser, 2012). Im Rahmen verschiedener Querschnittsanalysen ist es gelungen, die Abgrenzbarkeit des Verstehens literarischer Texte (LTV) vom Verstehen expositorischer Texte (ETV) anhand differenzieller Zusammenhängen zu motivationalen und affektiven Determinanten des Textverstehens zu erhärten (z. B. Henschel & Roick, 2013; Henschel, Roick, Martin & Stanat, 2013). Erkenntnisse über die längsschnittliche Entwicklung des literarischen Textverstehens liegen indes bislang nicht vor.

Fragestellung: Untersucht werden soll, wie sich literarisches und expositorisches Textverstehen über einen längeren Zeitraum entwickeln und wechselseitig determinieren. Eine besondere Betrachtung soll dabei den Subfacetten des literarischen Textverstehens zukommen. Auf Basis der Ergebnisse soll diskutiert werden, in welchen Bereichen Ansatzpunkte für eine evidenzbasierte Entwicklung literarischer Textverstehenskompetenz im Literaturunterricht bestehen.

Methode: An der Untersuchung nahmen ca. 1400 bayerische Jugendliche aus 62 Klassen der neunten Jahrgangsstufe (51 % Mädchen) von Realschulen und Gymnasien teil. Die Jugendlichen bearbeiteten Testverfahren zum literarischen und expositorischen Textverstehen zu Beginn und zum Ende der 9.

Klassenstufe sowie zum Ende der 10. Klassenstufe. Zur Analyse der längsschnittlichen Leistungsentwicklungen wurden autoregressive Modelle sowie latente Wachstumskurvenmodelle unter Berücksichtigung von Geschlecht und Schulformzugehörigkeit berechnet.

Ergebnisse: Sowohl das expositorische (ETV) als auch das literarische Textverstehen (LTV) weisen über den Zeitraum von 1,5 Schuljahren, d.h. von Anfang 9. Klasse (K9) bis Ende 10. Klasse (K10), eine hohe Konstruktstabilität auf (ETVK9 --> ETVK10:  $\beta = .70$ ; LTVK9 --> LTVK10:  $\beta = .76$ ). Betrachtet man die wechselseitigen kreuzverzögerten Effekte, so wird deutlich, dass die literarische Textverstehenskompetenz auch die Varianz des expositorischen Textverstehens fast vollständig erklären kann (LTVK9 --> ETVK10:  $\beta = .52$  bei ETVK9 --> ETVK10:  $\beta = .13$ ). Umgekehrt gilt dies jedoch nicht (ETVK9 --> LTVK10:  $\beta = .02$  bei LTVK9 --> LTVK10:  $\beta = .67$ ). Weitere Analysen über die Subfacetten des literarischen Textverstehens deuten darauf hin, dass Unterschiede in den Entwicklungsverläufen des formbezogenen literarischen Textverstehens vor allem bei Jungen sowie Schülerinnen und Schülern aus Realschulen auftreten (Regression der Slope-Werte auf Geschlecht/ Jungen:  $\beta = .14$  und auf Schulform/ Gymnasium:  $\beta = .20$ ).

Diskussion: Die Befunde liefern erste Hinweise darauf, in welchen Bereichen Ansatzpunkte für eine evidenzbasierte Entwicklung der literarischen Verstehenskompetenz im Literaturunterricht bestehen. Hierzu zählen Überlegungen, wie die Entwicklungspotenziale beim inhalts- und formbezogenen literarischen Textverstehen besser genutzt werden können. Weiter soll skizziert werden, um welche Subfacetten das Modell der literarischen Textverstehenskompetenz zu erweitern ist, um die Anwendbarkeit im Literaturunterricht zu optimieren.

Eco, U. (1990). Die Grenzen der Interpretation. München: dtv.

Frederking, V., Henschel, S., Meier, C., Roick, T., Stanat, P. & Dickhäuser, O. (2012). Beyond functional aspects of reading literacy: Theoretical structure and empirical validity of literary literacy. L1-Educational Studies in Language and Literature, 12, 35-58.

Henschel, S. & Roick, T. (2013). Zusammenhang zwischen Empathie und dem Verstehen literarischer Texte. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 45, 103-113.

Henschel, S., Roick, T., Martin, B. & Stanat, P. (2013). Leseselbstkonzept und Textart: Lassen sich literarisches und faktuales Leseselbstkonzept trennen? Zeitschrift für Pädagogische Psychologie, 27, 181-191.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] (2004). Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. München: Luchterhand.

Roick, T., Stanat, P., Dickhäuser, O., Frederking, V., Meier, C. & Steinhauer, L. (2010). Strukturelle und kriteriale Validität der literarästhetischen Urteilskompetenz. Projekt literarästhetische Urteilskompetenz. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. 56. Beiheft der Zeitschrift für Pädagogik (S. 165-174). Weinheim: Beltz.

Schilcher, A. & Pissarek, M. (2013). Auf dem Weg zur literarischen Kompetenz. Ein Modell literarischen Lernens auf semiotischer Grundlage. Baltmannsweiler: Schneider.

## **Strukturelle und graduelle Entwicklungen berufsfachlicher Kompetenzen im Ausbildungsverlauf**

Theoretischer Hintergrund: Im Anschluss an die Drei-Phasen-Theorie, in der davon ausgegangen wird, dass sich der Fertigkeitserwerb in drei Phasen mit zunehmenden Automatisierungsprozessen vollzieht (Ackerman, 1992), und Befunde zur beruflichen Ausbildung (z. B. Abele, 2013; Gschwendtner, 2011; Nickolaus, Geißel, Abele & Nitzschke, 2011; Nickolaus u.a. 2012; Nickolaus & Seeber, im Druck; Winther, 2010) kann begründet vermutet werden, dass sich im Ausbildungsverlauf einerseits eine Ausdifferenzierung der Kompetenzstruktur vollzieht. Andererseits finden in stark vertrauten Anforderungskontexten auch Verschmelzungsprozesse statt, die zugleich mit graduellen Entwicklungen verbunden sind.

Fragestellung: In diesem Beitrag gehen wir diesen Entwicklungen am Beispiel der Ausbildung zum Kfz-Mechatroniker bzw. zur Kfz-Mechatronikerin gestützt auf eine längsschnittliche Studie, die durch Querschnitts-untersuchungen ergänzt wurde, nach und prüfen, inwieweit sich bezogen auf unterschiedliche Formen des Fachwissens (nicht direkt handlungsbezogenes und direkt handlungsbezogenes Wissen) strukturelle und graduelle Entwicklungen vollziehen. Ausdifferenzierungsprozesse scheinen primär durch curriculare Einflüsse begünstigt. Verschmelzungsprozesse können vor allem dann erwartet werden, wenn durch die Anforderungskontexte integrative Verarbeitungsprozesse angeregt und intensiviert werden.

Methode: Im Beitrag gehen wir theoriegeleitet (vgl. Ackerman, 1992) primär zwei Hypothesen nach. Erstens prüfen wir, ob sich handlungsbezogenes Wissen, dessen Handlungskern/ Anforderungsstruktur als weitgehend betrieblich normiert angenommen wird, über die Ausbildungszeit dimensional reduziert (Verschmelzungsprozesse). Zweitens prüfen wir, ob sich nicht direkt handlungsbezogenes Wissen, dessen Genese stärker durch zwischenschulische Varianzen beeinflusst wird, über die Ausbildungszeit dimensional ausdifferenziert. Die Hypothesen prüfen wir auf der Basis zweier Stichproben (Längsschnitt 2. Aj., n=235; Querschnitt 3. Aj., n=241), in welchen das nicht direkt handlungsbezogene Fachwissen und das handlungsbezogene Wissen in zwei im Vertrautheitsgrad variierenden Anforderungsbereichen erhoben wurden.

Ergebnisse: Zu den strukturellen Entwicklungen zeigen konfirmatorische Faktorenanalysen erwartungskonforme Befunde, d. h., es kommt über die Ausbildungszeit einerseits zu Ausdifferenzierungen des nicht direkt handlungsbezogenen Wissens. Erste Analysen zur Strukturentwicklung des handlungsbezogenen Wissens deuten darauf hin, dass innerhalb dieser Subdimension im Ausbildungsverlauf eine schwächere Relationierung des darin anforderungsbezogen erhobenen Sachwissens und des Handlungswissens beobachtet werden kann (Reduktion der latenten Korrelationen von 0.8 auf 0.68). Zur graduellen Entwicklung sind die Analysen noch nicht vollständig abgeschlossen. Die Sensitivität der Instrumente scheint allerdings auf der Basis einer ersten klassischen Datenanalyse gewährleistet, so dass auch dazu Ergebnisse im Rahmen des Tagungsbeitrags präsentiert werden können.

Diskussion: Die Ausdifferenzierungsprozesse des Fachwissens scheinen vor allem durch die curricularen Schwerpunktsetzungen verursacht und vollziehen sich primär entlang von fachlichen Strukturen. Die zunehmend losere Relationierung des Sach- und Handlungswissens steht in Einklang mit der Drei-Phasen-Theorie von Ackerman (1992).

Abele, S. (2013). Modellierung, Entwicklung und Determinanten berufsfachlicher Kompetenz in gewerblich-technischen Ausbildungsberufen. Analysen auf Basis von testbasierten, berufsschulischen und betrieblichen Leistungsdaten sowie Prüfungsergebnissen. Aachen: Shaker.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 598-614.

Gschwendtner, T. (2011). Die Ausbildung zum Kraftfahrzeugmechatroniker im Längsschnitt. Analysen zur Struktur von Fachkompetenz am Ende der Ausbildung und Erklärung von Fachkompetenzentwicklungen über die Ausbildungszeit. *Zeitschrift für Berufs- und Wirtschaftspädagogik – Beihefte (ZBW - B)*, 24, 55-76.

Nickolaus, R., Geißel, B., Abele, S. & Nitzschke, A. (2011). Fachkompetenzmodellierung und Fachkompetenzentwicklung bei Elektronikern für Energie- und Gebäudetechnik im Verlauf der Ausbildung – Ausgewählte Ergebnisse einer Längsschnittstudie. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 25, 77-94.

Nickolaus, R., Abele, S., Gschwendtner, T., Nitzschke, A. & Greiff, S. (2012). Fachspezifische Problemlösefähigkeit in gewerblich-technischen Ausbildungsberufen – Modellierung, erreichte Niveaus und relevante Einflussfaktoren. *Zeitschrift für Berufs- und Wirtschaftspädagogik (ZBW)*, 108(2), 243-272

Nickolaus, R. & Seeber, S. (im Druck). Berufliche Kompetenzen: Modellierungen und diagnostische Verfahren In: Frey, A., Lissmann, U. & Schwarz, B. (Hrsg.), *Handbuch berufspädagogischer Diagnostik*.

Winther, E. (2010). *Kompetenzmessung in der beruflichen Bildung*. Bielefeld: W. Bertelsmann.

Dominique Rauch (DIPF)

***Determinanten der Kompetenzentwicklung von Personen mit Zuwanderungshintergrund:  
Perspektiven der Erziehungswissenschaften, Psychologie und Soziologie***

Chair(s): Dominique Rauch (DIPF)

Diskutant/in: Petra Stanat (IQB, Humboldt-Universität zu Berlin)

Ein selbstbestimmtes und erfolgreiches Leben zu führen erfordert in modernen Gesellschaften eine Grundausstattung an Kompetenzen, die in der Regel institutionell vermittelt werden. Aus der Perspektive der Erziehungswissenschaft, Psychologie und Soziologie untersucht das Symposium individuelle, institutionelle und durch verschiedene Bildungssysteme geschaffenen Bedingungen, unter denen Menschen mit Zuwanderungshintergrund über die Lebensspanne Kompetenzen erlernen. Der Beitrag von Hachfeld, Anders und Tuftensammer analysiert anhand von Daten der Evaluation des Programms „Offensive Frühe Chancen“ die kulturellen Überzeugungen und Selbstwirksamkeitserwartungen der am Programm teilnehmenden fröhpädagogischen Zusatzfachkräften und setzt diese mit der Prozessqualität und der Sprachentwicklung von Kindern mit Zuwanderungshintergrund in Verbindung. Der Beitrag von Rauch, Bergann und Stanat untersucht anhand der deutschen PISA 2009 Daten den Zusammenhang zwischen Geschlechterrollenorientierungen von Jugendlichen und Eltern mit Zuwanderungshintergrund und Lesekompetenz. Der Beitrag von Teltemann zieht internationale Daten der PISA Erhebungen 2000 bis 2012 heran, um die Frage zu beantworten, inwieweit sich Zusammenhänge zwischen Merkmalen des Bildungssystems (z.B. der Grad der Schulautonomie) und Bildungserfolg für Zugewanderte und Einheimische unterscheiden. Der Beitrag von Maehler schließlich untersucht auf Basis der deutschen PIAAC Daten des Jahres 2012, welche Indikatoren (soziodemografische, ökonomische und psychologische) für den Migrationshintergrund die Grundkompetenzen von erwachsenen Personen mit Migrationshintergrund in der deutschen Gesellschaft am besten vorhersagen.

Montag,

03.03.2014,

11:15 - 13:00,

IG-Farben-Haus

Raum 311



## **Frühförderung von Kindern mit Zuwanderungshintergrund im Blick: Orientierungsqualität von Fachkräften im Modellprojekt „Offensive Frühe Chancen: Schwerpunkt-Kitas Sprache & Integration“**

Mehr als ein Drittel der Kinder unter fünf Jahren in Deutschland hat einen Zuwanderungshintergrund (Cinar, Otremba, Stürzer, & Bruhns, 2013). Statistiken zeigen, dass diese sich nicht gleichmäßig auf das vorhandene institutionelle Betreuungsangebot verteilen, sondern dass Kinder mit Zuwanderungshintergrund (1) seltener eine Kindertagesstätte besuchen als Kinder ohne (Sachverständigenrat deutscher Stiftungen für Integration und Migration, 2013) und (2) häufiger eine Kindertagesstätte besuchen, in der die Mehrheit der Kinder ebenfalls einen Zuwanderungshintergrund hat (Cinar et al., 2013; Kiziak, Kreuter, & Klingholz, 2012). Sprechen die Kinder zusätzlich zuhause eine andere Sprache als Deutsch, kann von ungünstigen Bedingungen für einen frühzeitigen Erwerb der deutschen Sprache gesprochen werden. Der Spracherwerb gilt als zentrales Schlüsselthema kindlicher Entwicklung und als Grundvoraussetzung für den späteren Bildungserfolg sowie die gesellschaftliche Teilhabe. Die Offensive Frühe Chancen: Schwerpunkt-Kitas Sprache & Integration setzt hier an und fördert rund 4000 Kitas, die auch Kinder unter drei Jahren betreuen. Sie konzentriert sich dabei auf Kitas, die überdurchschnittlich häufig von Kindern mit Zuwanderungshintergrund und/oder aus sozial benachteiligten Familien besucht werden. Die teilnehmenden Kitas unterscheiden sich dabei hinsichtlich der Unterstützung, die sie im Rahmen des Programms erhalten (im Folgenden Intervention genannt; z.B. Konsultationskitas, Video-Coaching). Sprachförderansätze für den vorschulischen Bildungsbereich werden schon seit mehreren Jahren diskutiert und erprobt, jedoch mit ernüchternden Resultaten. Mittlerweile wird davon ausgegangen, dass langfristig erfolgreiche Ansätze der sprachlichen Bildung – im Gegensatz zu (Sonder-) Förderansätzen – in den pädagogischen Alltag der Einrichtungen integriert werden und eine stabil hohe Prozessqualität aufweisen sollten. Ebenfalls eine wichtige Dimension der Bildungsqualität und mit der Prozessqualität in Zusammenhang stehend ist die Orientierungsqualität. Es wird davon ausgegangen, dass eine auf interkulturelle und mehrsprachliche Bildung bezogene hohe Orientierungsqualität eine Voraussetzung für eine qualitativ hochwertige in den Kita-Alltag integrierte sprachliche Bildung ist. Auf der individuellen Ebene der Fachkraft bezieht sich die Orientierungsqualität auf Überzeugungen, pädagogische Werte und Einstellungen. Diese werden als Teil der professionellen Haltung und damit als wichtige professionelle Kompetenzen diskutiert (Anders, 2012). Hinsichtlich der (mehr-)sprachlichen Bildung sind diejenigen Überzeugungen hervorzuheben, die eine interkulturelle Dimension bei der Gestaltung von Bildungs- und Erziehungsprozessen anerkennen, und somit die kulturelle und sprachliche Vielfalt für die Interaktion und die Lernprozesse in Kindergruppen als Ressource nutzen. Multikulturelle Überzeugungen (MKÜ) beschreiben in diesem Sinne, wie sehr pädagogische Fachkräfte die Beibehaltung kultureller und sprachlicher Integrität unterstützen und somit in ihrer Arbeit stärker auf die verschiedenen Herkunftskulturen sowie Familiensprachen der Kinder einzugehen versuchen (Hachfeld, Schroeder, Anders, Hahn & Kunter, 2012). Im Rahmen des Beitrages wird die Orientierungsqualität der an der Offensive Frühe Chancen teilnehmenden Kita-Leitungen und Sprachförderkräfte (N = 257) untersucht, und zwischen den verschiedenen Interventionsgruppen verglichen. Es wird der Frage nachgegangen, ob die Teilnahme an einem speziellen Förderprogramm zu einer höheren Zustimmung zu MKÜ und höheren Selbstwirksamkeitserwartungen (SWE) in Bezug auf die Arbeit mit Kindern mit Zuwanderungshintergrund führt, und wie diese mit der Prozessqualität und der Sprachentwicklung der Kinder zusammenhängen. Erste Ergebnisse weisen auf eine überdurchschnittliche Zustimmung zu MKÜ (M = 5.43, SD = .62) unabhängig von der Interventionsgruppe. Höhere MKÜ hängen ebenfalls mit höheren SWE zusammen ( $r = .39, p < .001$ ). Je mehr die Sprachförderkraft sich zutraut, Kinder mit Zuwanderungshintergrund genügend fordern und fördern zu können (SWE), desto stärker wird auch die individuelle Förderung von Kindern mit Zuwanderungshintergrund forciert, was wiederum mit besseren Werten von Prozessqualität (KES-R; Tietze, Schuster, Grenner, & Roßbach, 2007) einhergeht. Erste Hinweise für Programm-Effekte zeigten sich in den Zusammenhängen von SWE mit der Dauer der Tätigkeit im Programm. Weitere Analysen sollen die sprachliche Entwicklung der Kinder einbeziehen. Die Ergebnisse werden in Hinblick auf ihre Bedeutung für Frühförderung im Kontext wachsender kultureller Heterogenität diskutiert.

Anders, Y. (2012). Modelle professioneller Kompetenzen für frühpädagogische Fachkräfte. Aktueller Stand und ihr Bezug zur Professionalisierung. Expertise zum Gutachten „Professionalisierung in der Frühpädagogik“ im Auftrag des Aktionsrats Bildung. München: vbm.

Cinar, M., Otremba, K. Stürzer, M. & Bruhns, K. (2013). Kinder-Migrationsreport. Ein Daten- und Forschungsüberblick zu Lebenslagen und Lebenswelten von Kindern mit Zuwanderungshintergrund. München: Deutsches Jugendinstitut e.V.

Hachfeld, A., Schroeder, S., Anders, Y., Hahn, A. & Kunter, M. (2012). Herkunft oder Überzeugung? Welche Rolle spielen der Zuwanderungshintergrund und multikulturelle Überzeugungen für das Unterrichten von Kindern mit Zuwanderungshintergrund.

Zeitschrift für Pädagogische Psychologie, 26, 101-120. doi: 10.1024/1010-0652/a000064

Kiziak, T., Kreuter, V. & Klingholz, R. (2012). Dem Nachwuchs eine Stimme geben. Was frühkindliche Sprachförderung leisten kann. Discussion Paper Nr. 6. Berlin-Institut für Bevölkerung und Entwicklung: Berlin.

Sachverständigenrat deutscher Stiftungen für Integration und Migration (2013).Hürdenlauf zur Kita: Warum Eltern mit Zuwanderungshintergrund ihr Kind seltener in die frühkindliche Tagesbetreuung schicken. Berlin: SVR GmbH.

Tietze, W., Schuster, K.-M., Grenner, K. & Rossbach, H.-G. (2007). Kindergarten- Skala (KES-R). Feststellung und Unterstützung pädagogischer Qualität in Kindergärten (3. Auflage). Weinheim: Beltz.





## ***Effekte traditioneller Geschlechterrollenorientierungen von Jugendlichen und Eltern mit und ohne Zuwanderungshintergrund auf Lesekompetenz: Analysen von Daten aus der PISA-2009-Erhebung.***

In der vorliegenden Untersuchung soll der Frage nachgegangen werden, ob und inwieweit Geschlechterrollenorientierungen (GRO) mit dem Bildungserfolg von Mädchen und Jungen mit und ohne Zuwanderungshintergrund in Zusammenhang stehen. Geschlechterrollenorientierung wird dabei als Einstellung verstanden, die bei Individuen eher traditionell oder egalitär ausgeprägt sein kann. Zum Zusammenhang zwischen Geschlechterrollenorientierung und Schulleistung können analytisch zwei Wirkmechanismen unterschieden werden: Der erste bezieht sich auf die Einstellungen der Jugendlichen selbst. Es wird angenommen, dass Einstellungen und Vorstellungen von Kindern und Jugendlichen zu Männlichkeit und Weiblichkeit ihr Lernverhalten beeinflussen und sich daher auf ihren Schulerfolg auswirken (Hannover, 2008; Stanat & Bergann, 2009). Ausgehend von dieser Grundannahme werden in der Literatur zwei entgegengesetzte Hypothesen diskutiert. So könnten traditionelle GRO bei Mädchen zu einer Verminderung von schulischer Anstrengung führen, während sich Jungen möglicherweise besonders in der Schule anstrengen, um gute Leistungen zu erzielen und später beruflich erfolgreich zu sein. Andere Autoren vermuten hingegen, dass eine traditionelle GRO bei Jungen mit der Gleichsetzung von Leistung und Schulerfolg als „weiblich“ einhergeht und dies die Motivation zu schulerefolgsbezogenem Verhalten bei Jungen mindert (Hadjar & Lupatsch, 2010). Der zweite Wirkmechanismus bezieht sich auf die Einstellungen der Eltern, die den außerschulischen Lernkontext der Jugendlichen prägen. Es wird angenommen, dass traditionelle GRO der Eltern, die insbesondere in Familien mit Migrationserfahrung beobachtet werden können, dazu führen, dass von Mädchen ein stärker konformes Verhalten, auch in der schulischen Umgebung, verlangt wird. Jungen würden hingegen mehr Freiheiten hinsichtlich ihres Verhaltens zugestanden. Die geringere elterliche Kontrolle der Jungen könnte dann zu einem im Vergleich mit Mädchen geringeren Schulerfolg führen (Feliciano & Rumbaut, 2005; Suarez-Orozco & Baolian Qin, 2006). Repräsentative Aussagen über die GRO von Jugendlichen und Eltern mit und ohne Zuwanderungshintergrund in Deutschland liegen bislang nicht vor. Zudem wurden beide skizzierten Wirkmechanismen des Zusammenhangs von GRO und Bildungserfolg bislang nicht systematisch gegeneinander abgewogen. Die vorliegende Studie überprüft deshalb, ob es einen Zusammenhang zwischen der GRO der Eltern und der Jugendlichen mit dem Bildungserfolg gibt und ob die Richtung dieser Zusammenhänge den aus dem ersten oder zweiten Wirkmechanismus abgeleiteten Hypothesen entspricht. Datengrundlage bietet die PISA-2009-Erhebung der 9. Klassen (N=8178, 49,5 % Mädchen, 51,5 % Jungen). 25,3 Prozent der Jugendlichen in der Stichprobe haben einen Zuwanderungshintergrund. Die Skala Geschlechterrollenorientierung (9 Items,  $\alpha=.89$ ) erfasst die GRO in den Bereichen Politik, Arbeit, Familie und Schule. Die Eltern der Schülerinnen und Schüler wurden anhand derselben Skala zu ihrer GRO befragt. Als Maß für den Bildungserfolg wurde als schulische Schlüsselkompetenz das Leseverstehen herangezogen. Die Ergebnisse zeigen zunächst, dass Mädchen signifikant egalitärer eingestellt sind als Jungen (Differenz  $z=.79$ ;  $p<.005$ ). Weiterhin bestätigt sich, dass Jugendliche mit Zuwanderungshintergrund signifikant traditioneller eingestellt sind als Jugendliche ohne Zuwanderungshintergrund (Differenz  $z=.30$ ;  $p<.005$ ). In multivariaten, nach Herkunftsland und Geschlecht getrennt modellierten, Regressionsanalysen zeigt sich unter Kontrolle des Bildungsgangs, des sozioökonomischen und kulturellen Kapitals der Familie sowie der Bildung der Eltern und der allgemeinen kognitiven Fähigkeiten der Jugendlichen, dass die GRO der Jugendlichen in einem substantiellen Zusammenhang mit Lesekompetenz steht. Dieser Zusammenhang ist sowohl für Mädchen und Jungen als auch für Jugendliche ohne und mit Migrationshintergrund negativ, d.h. je traditioneller die GRO, desto schwächer die Leistungen im PISA-Lesekompetenztest. Eine egalitäre Einstellung zum Geschlechterverhältnis ist demnach sowohl für Mädchen als auch für Jungen förderlich für Bildungserfolg, allerdings scheint dieser Effekt bei Mädchen zumeist deutlicher zu sein. Die GRO der Eltern weist hingegen unter Kontrolle der GRO der Jugendlichen keinen Zusammenhang mit der Lesekompetenz der Jugendlichen auf. Die Ergebnisse werden vor dem Hintergrund der dargestellten theoretischen Annahmen zu differenziellen Zusammenhängen von GRO und Bildungserfolg bei Mädchen und Jungen mit und ohne Zuwanderungshintergrund diskutiert.

Feliciano, C. & Rumbaut, R. (2005). Gendered Paths: Educational and Occupational Expectations and Outcomes among Adult Children of Immigrants. *Ethnic and Racial Studies*, 28, 1087-1118.

Hadjar, A. & Lupatsch, J. (2010). Der Schul(miss)erfolg der Jungen. *Kölner Zeitschrift für Soziologie*, 62, 599 – 622.

Hannover, B. (2008). Vom biologischen zum psychologischen Geschlecht: Die Entwicklung von Geschlechtsunterschieden. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 339-388). Bern: Huber.

Stanat, P. & Bergann, S. (2009). Geschlechtsbezogene Disparitäten in der Bildung. In: Tippelt, R. & Schmidt, B. (Hrsg.). *Handbuch Bildungsforschung* (2. Auflage) (S. 513-528). Opladen: Leske und Budrich.

Suárez-Orozco, C. & Baolian Qin, D. (2006). Gendered Perspectives in Psychology: Immigrant Origin Youth. *International Migration Review*, 40, 165–198.

## **Effekte von Bildungssystemen auf Leistungsunterschiede zwischen Migranten und Einheimischen: Ein internationaler Vergleich auf Basis von PISA**

Die Analyse von Bildungsungleichheiten zwischen Schülern mit und ohne Migrationshintergrund gehört inzwischen zum klassischen Programm (nicht nur) der empirischen Bildungsforschung. Dabei hat sich der überwiegende Teil der Forschung zur Entstehung ungleicher Bildungserfolge von Migranten und Einheimischen bislang auf Erklärungsfaktoren auf der Individual-, Familien- oder Gruppenebene konzentriert (u.a. Betts & Lofstrom, 2000; Kao & Thompson, 2003; Kristen & Granato, 2007; van Ours & Veenman, 2003). Insbesondere die Erklärung solcher Ungleichheiten, die auch unter Kontrolle von Merkmalen des sozioökonomischen familiären Hintergrunds bestehen, ist Gegenstand der migrantenbezogenen (soziologischen) Bildungsforschung. Ansätze zur Erklärung solcher „ethnic penalties“ verweisen in der Regel auf diskriminierende Praktiken im Bildungssystem, kulturelle Prägungen, die den Bedingungen für Bildungserfolg entgegenstehen, oder mangelndes strategisches Wissen über solche Erfolgsbedingungen (Gomolla & Radtke, 2002, Kao, 1995, Karakasoglu & Nieke, 2002, Pfeffer, 2008). Mit der Verfügbarkeit großer international vergleichender Schulleistungstests wurde deutlich, dass die Ungleichheit zwischen Migranten und Einheimischen in verschiedenen Staaten unterschiedlich stark ausgeprägt ist (u.a. Stanat & Christensen, 2006). Die unterschiedlich großen Leistungslücken lassen sich zu einem Teil durch selektive Migrationsprozesse erklären, die zu (in Bezug auf für den Bildungserfolg relevante Merkmale) unterschiedlichen Zusammensetzung der Migrantengruppen in verschiedenen Ländern führen. Doch auch hier haben einschlägige Auswertungen der vergleichenden Schulleistungstudien gezeigt, dass nach Konstanthaltung von soziokulturellen Hintergrundmerkmalen der Schüler die Ungleichheit zwischen Migranten und Einheimischen substantiell zwischen verschiedenen Staaten variiert. Nicht zuletzt durch diesen Befund ist der Einfluss von Kontextfaktoren auf ungleiche Schulleistungen, etwa auf der Ebene von Bildungssystemen, verstärkt in den Blick genommen worden (Entorf & Minoiu, 2005; Buchmann & Parrado, 2006; Park & Sandefur, 2010). Der Beitrag zeigt zunächst, wie Effekte von Institutionen auf Schulleistungen von Migranten theoretisch erfasst werden können, um Hypothesen zum Einfluss einzelner Merkmale von Bildungssystemen auf Schulleistungsunterschiede zwischen Migranten und Einheimischen entwickeln zu können. Das „Modell der soziologischen Erklärung“ (Esser, 1999) in Verbindung mit der Werterwartungstheorie und der Theorie sozialer Produktionsfunktionen bietet hierfür die Basis. Dabei geht es darum, durch die Rekonstruktion der Lebenssituation von Migranten Opportunitäten und Restriktionen zu identifizieren, die durch bestimmte institutionelle Merkmale moderiert werden können. In einem weiteren, empirischen Schritt zeigt der Beitrag anhand von Daten aus den PISA-Wellen 2000 bis 2012, welche Eigenschaften von Bildungssystemen mit Schulleistungsunterschieden zwischen Migranten und Einheimischen in Zusammenhang stehen und inwieweit sich Bildungssystemeffekte auf Leistungen zwischen Migranten und Einheimische unterscheiden. Die PISA-Studien haben die Möglichkeiten, Zusammenhänge zwischen Schul- bzw. Bildungssystemmerkmalen und Kompetenzen auch im internationalen Vergleich zu untersuchen, deutlich erweitert. Dabei stellt sich jedoch das Problem, dass aufgrund des querschnittlichen Designs Variationen in den Leistungen nicht kausal einzelnen Merkmalen von Bildungssystemen zugeschrieben werden können. Durch die Verknüpfung von Daten mehrerer PISA-Wellen ist es jedoch möglich, Zusammenhänge zwischen sich verändernden Merkmalen von Bildungssystemen zu untersuchen. Der Beitrag greift kurz die methodischen Probleme, die mit einer solchen Analysestrategie verbunden sind auf, und zeigt, wie Veränderungen auf der Ebene von Bildungssysteminstitutionen mit den Leistungsunterschieden zwischen Migranten und Einheimischen in Zusammenhang stehen. Die Ergebnisse für PISA 2000 bis 2009 zeigen, dass sich Bildungssystemeffekte für Migranten und Einheimische unterscheiden können: so geht etwa eine Ausweitung der fächerbezogenen Leistungsgruppierung mit besseren Leistungen von Einheimischen einher, während Schüler mit Migrationshintergrund in Ländern mit einem wachsenden Schüleranteil an Schulen mit vollständiger Leistungsgruppierung besser abschneiden. In Bezug auf Veränderungen im Grad der Schulautonomie zeigen einheimische Schüler in Ländern mit einer Ausweitung der schulischen Handlungsbefugnis in Bezug auf Prüfungen bessere Leistungen, Migranten können hier jedoch nicht profitieren.

Betts, J. R., Lofstrom, M. 2000: The Educational Attainment of Immigrants. Trends and Implications. In: Borjas, G. J. (Hg.), Issues in the economics of immigration. Chicago: University of Chicago Press, 51–117.

Buchmann, C., Parrado, E. A. 2006: Educational achievement of immigrant-origin and native students: a comparative analysis informed by institutional theory. In: Baker, D. P. (Hg.), The Impact of Comparative Education Research on Institutional Theory. Amsterdam: Elsevier JAI Press, 345–377.

Entorf, H., Minoiu, N. 2005: What a Difference Immigration Policy Makes: A Comparison of PISA Scores in Europe and Traditional Countries of Immigration. In: German Economic Review, 6. Jg., Heft 3, 355–376.

Esser, H. 1999: Soziologie. Allgemeine Grundlagen. Frankfurt/Main ; New York: Campus.

Gomolla, M., Radtke, F.-O. 2002: Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule. Opladen: Leske + Budrich.

Kao, G., Thompson, J. S. 2003: Racial and Ethnic Stratification in Educational Achievement and Attainment. In: Annual Review of Sociology, 29. Jg., 417–442.

Kao, G. 1995: Asian Americans as Model Minorities? A Look at Their Academic Performance. In: American Journal of Education, 103. Jg., Heft 2, 121–159.

Karakasoglu, Y., Nieke, W. 2002: Benachteiligung durch kulturelle Zugehörigkeit? In: Weegen, M. (Hg.), Bildungsforschung und Politikberatung. Schule, Hochschule und Berufsbildung an der Schnittstelle von Erziehungswissenschaft und Politik. Weinheim: Juventa, 199–218.

Kristen, C., Granato, N. 2007: The educational attainment of the second generation in Germany: Social origins and ethnic inequality. In: Ethnicities, 7. Jg., Heft 3, 343–366.

Ours, J. C., Veenman, J. 2003: The educational attainment of second-generation immigrants in the Netherlands. In: Journal of population economics, 16. Jg., Heft 4, 739–753.

Park, H., Sandefur, G. D. 2010: Educational Gaps Between Immigrant and Native Students in Europe: The Role of Grade. In: Dronkers, J. (Hg.), Quality and Inequality of Education: Cross-National Perspectives Assessing the Quality of Education and its Relationships With The Inequality in European and Other Modern Societies. Berlin: Springer, 113–136.

Pfeffer, F. T. 2008: Persistent Inequality in Educational Attainment and its Institutional Context. In: European Sociological Review, 24. Jg., Heft 5, 543–565.

Stanat, P., Christensen, G. 2006: Where Immigrant Students Succeed. A Comparative Review of Performance and Engagement in PISA 2003. Paris: OECD.



## **Grundkompetenzen erwachsener Migranten in Deutschland: eine Frage der Operationalisierung des Migrationsstatus**

Die Definition des Migrationsstatus variiert sowohl in verschiedenen Veröffentlichungen als auch über verschiedene Disziplinen hinweg. Häufig ist diese von den im jeweiligen Datensatz zur Verfügung stehenden Variablen abhängig. Bisherige Forschungsarbeiten anhand von Daten aus dem Mikrozensus zeigen, dass zum Beispiel die Bildungsbeteiligung und somit der Bildungserfolg von Erwachsenen je nach Operationalisierung des Migrationshintergrundes (z.B. Geburtsland, Staatsangehörigkeit, Muttersprache) stark variiert (z.B. Gresch & Kristen, 2011; Kogan, 2011; Stanat & Segeritz, 2009). Dies führt wiederum zu inkohärenten Aussagen über das Maß der strukturellen Integration in Deutschland. Basierend auf bisheriger Forschung soll im Rahmen dieses Beitrages untersucht werden, welche Effekte das Kriterium zur Operationalisierung des Migrationsstatus auf die in der deutschen Gesellschaft vorhandenen Grundkompetenzen hat. Die Analyse der PIAAC Daten (Programme for the International Assessment of Adult Competencies) zeigt, dass ein Migrant – definiert als eine Person mit mindestens einem ausländischen Elternteil – fast doppelt so häufig (25 % von N = 1600) in der untersten Lesekompetenzstufe (unter Stufe I) vorzufinden ist, wie ein Autochthone (14 % von N = 3730). Gleiches gilt für das Kriterium Staatsangehörigkeit (z. B. Nicht-EU Staaten: 39 % von N = 301; Deutsch: 15 % von N = 4860). Wird jedoch das Kriterium Muttersprache zur Operationalisierung herangezogen, ist der Anteil der Migranten (41 % von N = 589) in der untersten Kompetenzstufe (unter Stufe I) sogar ca. dreimal so groß wie der Anteil Autochthone (15 % von N = 4786). Dabei können Personen auf der untersten Kompetenzstufe lediglich sehr einfache, elementare Leseaufgaben lösen. Die gestellten Aufgaben in PIAAC erfordern meistens das Lesen relativ kurzer Texte, wobei eine einzelne Information im Text aufgefunden werden muss. Kenntnis eines Basiswortschatzes, das Verstehen von Sätzen und das Lesen von Abschnitten werden dabei erwartet (vgl. Zabal et al., 2013). Der anteilmäßige Unterschied in den untersten Kompetenzstufen zwischen Personen mit Migrationshintergrund beim Klassifikationskriterium „Geburtsland der Eltern“ (mindestens ein Elternteil im Ausland geboren) und Autochthone beträgt 11 Prozentpunkte. Demgegenüber sind 23 Prozentpunkte mehr Personen mit Migrationshintergrund beim Klassifikationskriterium Staatsangehörigkeit (z. B. Drittstaaten/ Nicht EU-Staaten) und 26 Prozentpunkte beim Kriterium Muttersprache (Deutsch vs. nicht Deutsch) der untersten Lesekompetenzstufe zugeordnet als Autochthone. Darüber hinaus wurde mittels multivariater Analysen untersucht, welche Indikatoren (soziodemografische, ökonomische und psychologische) für den Migrationshintergrund die Grundkompetenzen von erwachsenen Personen mit Migrationshintergrund in der deutschen Gesellschaft am besten vorhersagen. Unter den migrationsbezogenen Indikatoren (z.B. Generation, Staatsangehörigkeit, Sprache oder Wertvorstellungen) kommt der Sprache die größte Bedeutung zu. Es zeigt sich jedoch, dass den migrationsbezogenen Indikatoren insgesamt eine sehr geringe Bedeutung zukommt. Faktoren wie Bildungsabschluss und Alter induzieren deutlich stärkere Effekte. Die Auswertungen legen nahe, dass in der Berichterstattung klar gestellt werden muss, welche migrationsbezogenen Indikatoren verwendet werden, um eine angemessene Ableitung politischer Maßnahmen zu ermöglichen.

Gresch, C. & Kristen, C. (2011). Staatsbürgerschaft oder Migrationshintergrund? Ein Vergleich unterschiedlicher Operationalisierungsweisen am Beispiel der Bildungsbeteiligung. Zeitschrift für Soziologie, 40 (3), S. 208-227.

Kogan, I. (2011, November). Probleme mit Klassifizierungen von Migrationshintergrund bzw. dessen Erfassung. Vortrag auf Herbstreffen des BMBF-Promotionsbegleitprogramms, Förderschwerpunkt „Sekundäranalysen“, Frankfurt.

Stanat, P., & Segeritz, M. (2009). Migrationsbezogene Indikatoren für eine Bildungsberichterstattung. In Tippelt, R. (Hrsg.): Steuerung durch Indikatoren (S. 141-156). Opladen u.a.: Budrich.

Zabal, A., Martin, S., Klaukin, A., Rammstedt, B., Baumert, J. & Klime, E. (2013): Grundlegende Kompetenzen der erwachsenen Bevölkerung in Deutschland im internationalen Vergleich. In

Rammstedt, R. (Hrsg.): Grundlegende Kompetenzen Erwachsener im internationalen Vergleich – Ergebnisse von PIAAC 2012 (S.31-76). Münster: Waxmann.



## **Personenorientierte empirische Bildungsforschung?**

Chair(s): Manuel Voelkle (Max-Planck-Institut für Bildungsforschung)

Diskutant/in: Ulman Lindenberger (Max-Planck-Institut für Bildungsforschung)

Natürgemäß ist die Suche nach den zugrundeliegenden Mechanismen individueller Entwicklungs- und Lernprozesse in der empirischen Bildungsforschung schwierig. In aller Regel ist weder die Unabhängigkeit, noch die Zufälligkeit, bei der Auswahl der zu untersuchenden Individuen gewährleistet. Darüber hinaus können randomisierte Kontrollgruppenexperimente höchstens flankierend, selten aber als primäres oder gar einziges Instrument des Erkenntnisgewinns zum Einsatz kommen, denn dies hieße in letzter Konsequenz, Bildung aktiv vorzuenthalten. Entsprechend groß ist die Bedeutung von Längsschnittstudien, nicht nur zur Dokumentation von Veränderungen und Zusammenhängen, sondern insbesondere zur Untersuchung von Mechanismen \*individueller\* Entwicklungs- und Lernprozesse mit dem Ziel, diese zu beschreiben, zu erklären, vorherzusagen und zu verbessern.

Ziel des vorgeschlagenen Symposiums ist es daher einen Überblick über aktuelle Probleme und Möglichkeiten bei der Identifikation von Mechanismen individueller Effekte in nicht-experimentellen Forschungsdesigns zu geben und zu diskutieren auf welcher „Ebene“ (Bildungs-) Forschung am besten angesiedelt sein soll. Im ersten Teil des Symposiums werden Oliver Lüdtke und Alexander Robitzsch der Frage nachgehen ob für Daten mit einer Mehrebenenstruktur immer eine Mehrebenenanalyse durchgeführt werden muss und wann die Verwendung von population-average versus cluster-spezifischen Effekten indiziert ist. Daran anknüpfend vergleichen Ulrich Pötter und Götz Rohwer verteilungsabhängige und verteilungsunabhängige Skalierungsansätze für die statistische Auswertung von (personenorientierten) Kompetenzverläufen. Während die methodische und inhaltliche Relevanz der Unterscheidung von within-person und between-person Effekten unbestritten ist, geht man—mit wenigen Ausnahmen—oft implizit von der Äquivalenz von within-person und between-person Messmodellen aus, obwohl die Problematik einer solchen Annahme bereits seit langem bekannt ist (cf. Molenaar, 2004). Im zweiten Teil des Symposiums, werden Manuel Voelkle et al. diese Annahme daher etwas genauer beleuchten und Möglichkeiten einer direkten empirischen Überprüfung diskutieren. Darauf aufbauend werden Florian Schmiedek et al. der Frage nachgehen was uns die Struktur interindividueller Unterschiede in kognitiven Fähigkeiten über die Struktur intraindividuelle Veränderungen in kognitiven Leistungen sagt und wo die Möglichkeiten und Grenzen eines solchen Vergleichs liegen. Bedarf es tatsächlich dem von Molenaar und Campbell (2009; p. 112) für die Psychologie eingeforderten Kuhnschen Paradigmenwechsel auch in der Bildungsforschung? Wir hoffen, dass das Symposium durch die interdisziplinäre Perspektive auf prominente methodische Verfahren und aktuelle Entwicklungen in der Längsschnittdatenanalyse sowie das Aufwerfen und die Diskussion grundsätzlicher Fragen in der aktuellen Bildungsforschung auf breites Interesse bei einem heterogenen Publikum stoßen wird.

Montag,

03.03.2014,

11:15 - 13:00,

PEG-Gebäude

Raum 1.G168

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2(4), 201-218.

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, 18, 112–117. doi: 10.1111/j.1467-8721.2009.01619.x



## **Mehrebenenanalyse in der empirischen Bildungsforschung:**

### **Muss für Daten mit einer Mehrebenenstruktur immer eine Mehrebenenanalyse durchgeführt werden?**

Datensätze in der empirischen Bildungsforschung besitzen häufig eine Mehrebenenstruktur, die dadurch gegeben ist, dass z.B. Schüler innerhalb von Klassen und Klassen innerhalb von Schulen geschachtelt sind. Generell werden zwei Ansätze bei der statistischen Analyse von Datensätzen mit einer Mehrebenenstruktur bzw. Clusterdaten unterschieden (Raudenbush & Bryk, 2002; Subramanian & O'Malley, 2010). Der erste Ansatz betrachtet die Mehrebenenstruktur als einen „Störfaktor“. So wird z.B. häufig eine Stichprobe von Klassen aus einer Population von Schulklassen ausgewählt, aus denen alle Schülerinnen und Schüler bzw. wiederum eine Stichprobe berücksichtigt werden. Aufgrund dieses Multistage-Samplings sind die Beobachtungen nicht – wie in einfachen Zufallsstichproben – unabhängig voneinander. Diese korrelierten Beobachtungen führen dazu, dass die effektive Stichprobengröße bei der Verwendung von statistischen Verfahren überschätzt wird, was mit einer Unterschätzung der Standardfehler verbunden ist. Es wurde eine Reihe von Verfahren vorgeschlagen, mit denen die Standardfehler in statistischen Analysen mit Clusterdaten korrigiert werden können (Scott & Holt, 1982; Wooldridge, 2003). Der zweite Ansatz betrachtet die Mehrebenenstruktur nicht nur als „Störfaktor“, sondern erklärt die verschiedenen Ebenen zum Gegenstand seiner Untersuchung (Raudenbush & Bryk, 2002). So setzen sich z.B. die interindividuellen Unterschiede zwischen den Schülerinnen und Schülern – unter der Annahme einer Zwei-Ebenen-Struktur – aus zwei Variationsquellen zusammen, nämlich der Variation auf der individuellen Ebene und der Klassenebene. Der individuelle Wert  $Y_T$  eines Schülers lässt sich danach zerlegen in den Klassenmittelwert  $Y_B$  seiner Klasse und die individuelle Abweichung  $Y_W$  des Schülers von diesem Klassenmittelwert (Cronbach, 1976):  $Y_T = Y_W + Y_B$ . Im Rahmen dieses Ansatzes ist häufig von Interesse, inwiefern sich Effekte innerhalb von Klassen (Within-Class) von Effekten zwischen Klassen (Between-Class) unterscheiden. Anhand der Frage des Zusammenhanges zwischen Leseleistung und sozioökonomischem Status (SES) wird auf Grundlage der Daten einer großen Schulleistungsstudie demonstriert, dass beide Ansätze zu einer sehr unterschiedlichen Schätzung der Stärke des Effektes von SES auf die Leseleistung kommen können. Im ersten Ansatz zeigte eine Regressionsanalyse mit korrigierten Standardfehlern (Liang & Zeger, 1986) einen relativ starken Effekt ( $\beta = 0.326$ ,  $s.e. = .018$ ) des SES auf die Leseleistung. Für das im zweiten Ansatz spezifizierte random-intercept Modell fiel dagegen der Effekt nur vergleichsweise schwach aus ( $\beta = 0.097$ ,  $s.e. = .012$ ). In der statistischen Literatur wird der Schätzer des ersten Ansatzes auch als der population-averaged Effekt bezeichnet, da er unabhängig von der Zuordnung der Personen zu den Clustern ist. Dagegen stellt der zweite Schätzer einen cluster-spezifischer Effekt dar, da er auf die Zugehörigkeit zu den Clustern bedingt (Gardiner et al., 2009). Es wird argumentiert (siehe Graubard & Korn, 1994), dass bei Interesse am Effekt des individuellen SES auf die individuelle Leseleistung (marginaler Effekt) der population-averaged Effekt zu bevorzugen ist, da die Clusterung nur als ein Störfaktor angesehen wird. Die beiden Ansätze werden sich vor allem dann unterscheiden, wenn die Zugehörigkeit zu den Clustern nicht exogen ist. Dies wird im vorliegenden Beispiel der Fall sein, da die Zuweisung zu den Schulen sowohl mit der Kovariaten SES als auch mit der Leseleistung korreliert sein dürfte. Der Unterschied zwischen dem population-averaged und dem cluster-specific Effekt soll zusätzlich mit Hilfe einer Datensimulation illustriert werden (Cronbach, 1976). Des Weiteren soll für typische Anwendungen der Mehrebenenanalyse in der empirischen Bildungsforschung (z.B. Analyse von Kontexteffekten oder Faktorenanalysen) diskutiert werden, ob eher der population-averaged oder der cluster-specific Effekt von Interesse ist. Es wird argumentiert, dass das zentrale Kriterium bei der Wahl des Analysemodells die theoretische Fragestellung sein muss. Abschließend sollen diese Überlegungen auf längsschnittliche Designs übertragen werden. Analog zur Betrachtung von Within-Class und Between-Class Effekten bei querschnittlichen Mehrebenenstrukturen werden auch hier häufig Within-Person von Between-Person Effekten separiert (Curran & Bauer, 2011), um die Effekte von Faktoren innerhalb von Personen von den Effekten von Faktoren zwischen Personen zu trennen.

Cronbach, L. J. (1976). Research on classrooms and schools: Formulations of questions, design, and analysis. Stanford, CA: Stanford Evaluation Consortium.

Curran, P.J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. Annual Review of Psychology, 62, 583-619.

Gardiner, J. C., Luo, Z., & Roman, L. A. (2009). Fixed effects, random effects and GEE: What are the differences? Statistics and Medicine, 28, 221-239.

Graubard, B. I., & Korn, E. (1994). Regression with clustered data. Statistics in Medicine, 13, 509-522.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. Biometrika, 73, 13-22.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models (2nd ed.). Thousand Oaks: Sage.

Scott, A. J., & Holt, D. (1982). The effect of two-stage sampling on ordinary least squares methods. Journal of the American Statistical Association, 77, 848-854.

Subramanian, S. V., & O'Malley, A. J. (2010). Modeling neighborhood effects: The futility of comparing mixed and marginal approaches. Epidemiology, 21, 475-478.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. American Economic Review, 93, 133-138.



## **Statistische Auswertungen von Kompetenzverläufen**

Kompetenzverläufe werden aus zeitlich aufeinander folgenden Kompetenztests konstruiert. Dabei können Skalierungen der einzelnen Kompetenztests verteilungsabhängig bzw. verteilungsunabhängig konstruiert werden. Bei verteilungsunabhängigen Konstruktionen beruhen die Kompetenzwerte einer Person ausschließlich auf ihren Testergebnissen zu einem Zeitpunkt, bei verteilungsabhängigen Konstruktionen aber auch auf den Ergebnissen anderer Personen und Zeiten. Das ist insbesondere dann der Fall, wenn Skalierungen mit IRT-Modellen gewonnen werden.

Wir möchten zeigen, dass diese Unterscheidung bei statistischen Auswertungen berücksichtigt werden muss. Das betrifft zunächst Möglichkeiten für personenorientierte Analysen, die mit verteilungsabhängigen Kompetenzwerten nicht sinnvoll durchgeführt werden können. Weitere Konsequenzen betreffen die Verwendung der Kompetenzverläufe als abhängige Variablen in Regressionsmodellen. Wir diskutieren Möglichkeiten, den Konstruktionsprozess der Kompetenzverläufe konzeptionell in die Spezifikation von Regressionsmodellen einzubeziehen.

Manuel Voelkle (Max-Planck-Institut für Bildungsforschung), Annette Brose (KU Leuven),  
Florian Schmiedek (DIPF, IDeA, Goethe-Universität), Ulman Lindenberger (Max-Planck-Institut für Bildungsforschung)

## **Über die Vereinbarkeit von inter- und intra-individuellen Modellen in der sozialwissenschaftlichen Forschung**

Die moderne sozial- und bildungswissenschaftliche empirische Forschung beschäftigt sich fast ausschließlich mit der Analyse von Gruppenmittelwerten und Varianz-/Kovarianzstrukturen zwischen Personen, das heißt mit inter-individuellen Modellen. Nahezu unser gesamtes Wissen über grundlegende Persönlichkeitseigenschaften, die Dimensionen affektiven Erlebens, oder die Struktur und prädiktive Validität kognitiver Fähigkeiten geht auf die Analyse von Zusammenhängen zwischen, in einer Personenstichprobe gemessenen, Variablen zurück. Für manch einen Wissenschaftler ist dieses Vorgehen so selbstverständlich, dass es gleichsam als einzig denkbare Form des empirischen Erkenntnisgewinns angesehen wird. Die Frage was all diese Mittelwerte, Kovarianzen, und Korrelationen denn mit dem Individuum, also mit mir, meinem Kind, oder einem ganz bestimmten Schüler/Lehrer zu tun haben ist höchstens dem wissenschaftlichen Laien gestattet (cf. Borsboom, Kievit, Cervone, & Hood, 2009). Bei genauerer Betrachtung heißt die Antwort auf diese Frage jedoch häufig: Nicht viel. Dies ist in keiner Weise lediglich ein Problem der Übertragbarkeit von Forschungsergebnissen in die Praxis, sondern betrifft den Kern der aktuellen sozialwissenschaftlichen Forschung, welche zwar häufig auf Theorien über intra-individuelle (kausale) Mechanismen aufbaut, diese aber in den seltensten Fällen direkt zu untersuchen vermag. Theoretisch würden wir gerne eine Aussage der Form „weil Lisa so intelligent ist, bekommt sie so gute Schulnoten“ treffen können. Praktisch setzt dies jedoch die Beobachtung der kontrafaktischen Situation voraus, dass Lisa keine guten Schulnoten bekommen würde wenn sie nicht so intelligent wäre. Ein gängiger Ansatz mit diesem Fundamentalproblem (der Kausalanalyse) umzugehen ist Personen als austauschbar zu betrachten und sich mit dem Zusammenhang zwischen Intelligenz und Schulnoten auf inter-individueller Ebene zufriedenzugeben. Wie weit jedoch trägt diese Annahme, insbesondere wenn es um komplexere Kovarianzstrukturen wie Intelligenz- oder Persönlichkeitsfaktoren geht? Ziel des Beitrags ist es dieser Frage etwas genauer nachzuspüren. Zu diesem Zweck stellen wir einen theoretischen Bezugsrahmen vor, welcher zwischen Quellen der inter- und intra-individuellen Varianz unterscheidet, und es uns erlaubt Gemeinsamkeiten, sowie das Ausmaß und mögliche Gründe für die Nichtäquivalenz von inter- und intra-individuellen Modellen, zu untersuchen. Wir gehen dabei in fünf Schritten vor: Zu Beginn (1) führen wir das Konzept der konditionalen Äquivalenz von inter- und intra-individuellen Modellen ein und schlagen (2) einen Maximum Likelihood Test zur inferenzstatistischen Untersuchung dieser Äquivalenz vor. Darauf aufbauend (3) diskutieren wir die Konstruktion von Likelihood-Ebenen zur Exploration möglicher Ursachen mangelnder Äquivalenz und (4) illustrieren dies unter Verwendung simulierter Daten anhand von vier gängigen Beispielen: Autoregression, Gruppenunterschiede, lineare Trends und oszillierende Trends. Wir beenden den Vortrag (5) mit einem kurzen empirischen Beispiel. Im Rahmen des Beitrags sollen sowohl grundlegende theoretisch/konzeptuelle Fragen aufgeworfen und diskutiert werden, als auch sich daraus ergebende methodisch/statistische Probleme. Auf diese Weise hoffen wir sowohl das Interesse von primär inhaltlich, als auch primär methodisch orientierten Bildungsforschern wecken zu können.

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 67-97). New York: Springer.

**GEBF** TAGUNG  
2014

A03 c

Symposiums  
beitrag

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G168

Personenorientierte empirische Bildungsforschung?



Florian Schmiedek (DIPF, IDeA, Goethe-Universität), Martin Lövdén (Karolinska Institute and Stockholm University),  
Timo von Oertzen (University of Virginia), Ulman Lindenberger (Max-Planck-Institut für Bildungsforschung)

## ***Was sagt uns die Struktur interindividueller Unterschiede in kognitiven Fähigkeiten über die Struktur intraindividuelle Veränderungen in kognitiven Leistungen?***

Über ein Jahrhundert psychometrischer Forschung zur Struktur interindividueller Unterschiede in verschiedensten kognitiven Leistungen hat zu dem weitgehenden Konsens geführt, dass diese gut durch hierarchische Faktorenmodelle, mit engen, breiten und allgemeinen Fähigkeitsfaktoren, beschrieben werden kann (s. Carroll, 1993). Zu den breiten Fähigkeitsfaktoren gehören dabei unter anderem Faktoren wie Wahrnehmungsgeschwindigkeit, fluide Intelligenz und Merkfähigkeit. Die positive Korrelation dieser Fähigkeitsfaktoren kann wiederum durch einen Generalfaktor allgemeiner kognitiver Fähigkeit erklärt werden.

Strukturen interindividueller Unterschiede sagen jedoch nicht notwendigerweise etwas über die Struktur intraindividuelle Veränderungen einzelner Personen aus (vgl. Borsboom, Mellenbergh, & van Heerden, 2003; Molenaar, 2004; Schmitz, 2000). Die Frage nach (interindividuellen Unterschieden in) intraindividuellen Veränderungsstrukturen ist allerdings von hoher Relevanz im Kontext von pädagogisch-psychologischen Fragestellungen, da für Lernprozesse die jeweils aktuell verfügbaren kognitiven Ressourcen zentral sind. Der Generalfaktor kognitiver Fähigkeiten an der Spitze hierarchischer Strukturmodelle, zum Beispiel, suggeriert, dass die Verfügbarkeit kognitiver Ressourcen wie logischem Schlussfolgern oder Merkleistungen auch innerhalb einzelner Personen über Zeitpunkte hinweg positiv miteinander korrelieren – und das in vergleichbarer Weise für verschiedene Personen. Inwiefern die etablierten Fähigkeitsstrukturen interindividueller Unterschiede tatsächlich solche Aussagen über die Struktur(en) intraindividuelle Prozesse erlauben, ist jedoch eine empirisch völlig offene Frage, zu deren Beantwortung die Untersuchung individueller Zeitreihen von Leistungen auf verschiedenen kognitiven Aufgaben an einer großen Zahl von Zeitpunkten notwendig ist.

In der COGITO-Studie bearbeiteten 101 jüngere Erwachsene im Alter von 20-31 Jahren am Computer eine Batterie von drei Arbeitsgedächtnis-, drei Merkfähigkeits und drei Wahrnehmungsgeschwindigkeitsaufgaben an jeweils 100 Testsitzungen. Interindividuelle Unterschiede in Leistungen bei den gleichen Aufgaben in einer Voruntersuchung dienen zur Etablierung einer Referenzstruktur interindividueller Unterschiede. In Kombination erlauben diese Daten den Vergleich von konfirmatorischen P-Technik-Faktorenstrukturen intraindividuelle Unterschiede zwischen verschiedenen Personen, sowie den direkten Vergleich dieser Strukturen mit der Struktur interindividueller Unterschiede. Erste Ergebnisse zeigen, dass die Personen sich stark darin unterscheiden, (a) für welche Aufgaben und Fähigkeitsfaktoren sich reliable tag-tägliche Schwankungen nachweisen lassen, (b) bei Vorhandensein solch reliabler Schwankungen, diese sich in der Regel zufriedenstellend mit einem hierarchischen Faktorenmodell beschreiben lassen, sowie (c) signifikante Abweichungen der intraindividuellen Strukturen von der interindividuellen Struktur. Die Bedeutung dieser Befunde für pädagogisch-psychologische Fragestellungen wird diskutiert.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110, 203-219.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, 2, 201-218.

Schmitz, B. (2000). Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen. *Psychologische Rundschau*, 51, 83-92.

**GEBF** TAGUNG  
2014

A03 d

Symposiums  
beitrag

Montag,

03.03.2014,

11:15 - 13:00,

PEG-Gebäude

Raum 1.G168

2

Personenorientierte empirische Bildungsforschung?



Chair(s): Janine Buchholz (DIPF), Johannes Hartig (DIPF)

Diskutant/in: Stefan Ufer (Ludwig-Maximilians-Universität München)

Im Zuge der nationalen und internationalen Forschungsinitiativen zur Erfassung des Leistungsstands von Schülerinnen und Schülern ist es von zentraler Bedeutung, mathematische Kompetenzen differenziert und genau messen zu können. Die Modellierung mathematischer Kompetenzen kann dabei als Schnittstelle zwischen Didaktik und psychologischer Diagnostik angesehen werden. In nationalen Forschungsstudien wie beispielsweise dem Projekt Co<sup>2</sup>CA als auch Large-Scale-Assessments zur Erfassung mathematischer Kompetenzen wie dem Test zur Überprüfung des Erreichens der Anforderungen der Bildungsstandards oder PISA wird in der Regel eine Differenzierung mathematischer Kompetenzen vorgenommen, die durch die Mathematikdidaktik fundiert ist. Die Bildungsstandards für Mathematik beispielsweise kategorisieren mathematische Kompetenzen hinsichtlich einer Prozess-, einer Inhalts- und einer Anspruchsdimension (Roppelt, Blum, & Pöhlmann, 2013). Psychometrische Modelle bilden diese theoretische Struktur ab und erlauben so Aussagen über die empirischen Zusammenhänge zwischen einzelnen Teilfähigkeiten. Im Test zu den Bildungsstandards beispielsweise wurden latente Korrelationen zwischen den fünf Leitideen zwischen .75 und .97 (Roppelt, Blum, & Pöhlmann, 2013) gefunden. Diese hohen Zusammenhänge erlauben daher auch eine Rückmeldung der Ergebnisse auf einer Globalskala, die einen über alle Leitideen übergreifenden Messwert darstellt. Eine solche Abwägung zwischen theoretisch angenommener und empirisch ermittelter Dimensionalität kann nur im Dialog zwischen beiden Disziplinen erfolgen. Das Ziel dieses Symposiums ist es daher, Fragen der möglichen Dimensionalität mathematischer Kompetenzen sowohl aus fachdidaktischer als auch aus psychometrischer Sicht vertiefend zu betrachten.

Im ersten Vortrag differenzieren Schukajlow und Krug Mathematikaufgaben hinsichtlich ihres Realitätsbezuges und untersuchen deren Rolle auf motivationale und emotionale Selbsteinschätzungen vor und nach der Aufgabenbearbeitung. Im zweiten Vortrag widmen sich Besser und Leiss Möglichkeiten der empirischen Erfassung und Beschreibung von mathematischen Teilfähigkeiten, und nutzen dafür die Dimensionen prozessbezogene Kompetenzen und mathematische Inhaltsbereiche aus den Studien DISUM und Co<sup>2</sup>CA. Mikolajetz und Frey untersuchen im dritten Beitrag anhand der Daten der Bildungsstandards, ob mit einer Kombination aus theoriebasierter und psychometrisch-explorativer Skalenbildung neben den fünf Leitideen auch die sechs Kompetenzen mit akzeptabler Reliabilität abgebildet werden können. Im vierten Beitrag schlagen Buchholz und Hartig ein Modell der mehrdimensionalen Item Response Theorie vor und demonstrieren anhand von Daten aus PISA 2012, wie dieses unter Berücksichtigung theoretischer Annahmen zur Modellierung mathematischer Antwortprozesse genutzt werden kann. Abschließend wird Stefan Ufer die Vorträge diskutieren.

Roppelt, A., Blum, W. & Pöhlmann, C. (2013). Beschreibung der untersuchten mathematischen Kompetenzen. In Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Münster: Waxmann.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G191



## **Interesse, Freude und Langeweile zu mathematischen Aufgaben mit und ohne Realitätsbezug: Die Rolle der Aufgabenschwierigkeit und der Aufgabenbearbeitung**

Theoretischer Hintergrund und Fragestellungen: Es wird schon lange gefordert, in der Schule realitätsbezogene Aufgaben zu behandeln, um Lernende auf Anforderungen im Beruf und Alltag vorzubereiten. Unklar ist zugleich, welche Wirkung die Bearbeitung von Aufgaben mit Realitätsbezug im Vergleich zum Lösen von traditionellen Aufgabentypen auf lernprozessrelevante Merkmale wie Interesse, Freude und Langeweile von Lernenden haben (Goetz, Frenzel, Hall, & Pekrun, 2008; Hidi & Renninger, 2006; Köller, Baumert, & Schnabel, 2001; Krapp, 2005; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010).

Je nach Grad des Realitätsbezugs können Aufgaben in Modellierungsaufgaben, „eingekleidete“ Textaufgaben und innermathematische Probleme unterteilt werden (Blum, Galbraith, Henn, & Niss, 2007). Bei den Modellierungsaufgaben ist eine problemhaltige Situation aus der Realität gegeben, deren Bearbeitung Konstruktion eines mathematischen Modells verlangt. „Eingekleidete“ Textaufgaben enthalten bereits ein solches mathematisches Modell. Bei den innermathematischen Aufgaben fehlt der Realitätsbezug vollständig.

Die Ergebnisse zum Einfluss des Realitätsbezugs von Aufgaben auf motivationale und emotionale Merkmale sind zum Teil inkonsistent. Während Pekrun et al. (2007) zeigen, dass „eingekleidete“ Aufgaben mehr Freude als innermathematische Aufgaben hervorrufen, wurden in der Studie von Schukajlow et al. (2012) kein Einfluss des Aufgabentyps auf Interesse und Freude festgestellt. Gemäß einer neueren Studie interessierten sich Schüler sogar weniger an der Bearbeitung von Modellierungsaufgaben als am Lösen der anderen Aufgabentypen (Krug & Schukajlow, 2013). Die vorliegende Untersuchung geht der Vermutung nach, dass die Schwierigkeit von Aufgaben Schülereinschätzungen beeinflusst. Zudem kann die Einschätzungen des Interesses und der Emotionen davon abhängig sein, in wie weit Lernende die Aufgaben lediglich gelesen oder auch bearbeitet haben. Es ergeben sich somit folgende Fragestellungen: 1. Wie hängen Interesse, Freude und Langeweile von Schülern mit der Schwierigkeit einer mathematischen Aufgabe zusammen? 2. Unterscheiden sich aufgabenbezogenes Interesse, Freude und Langeweile von Schülern je nach Grad des Realitätsbezugs? 3. Wie beeinflusst die Aufgabenbearbeitung Interesse, Freude und Langeweile zu verschiedenen Aufgabentypen?

Methode: Zur Beantwortung der Forschungsfragen wurden 192 Lernende aus 7 Realschul- und Gymnasialklassen der 9. Jahrgangsstufe befragt. Eine Hälfte hat zuerst ihr Interesse, Freude und Langeweile zu den präsentierten Aufgaben angegeben und anschließend die Aufgaben bearbeitet. Die andere Hälfte hat die Fragen zum Interesse, Freude und Langeweile nach der Aufgabenbearbeitung beantwortet. Die Probanden wurden innerhalb einer Klasse zufällig der jeweiligen Untersuchungsbedingung zugewiesen. Für die Studie wurden 23 Modellierungsaufgaben, „eingekleidete“ Textaufgaben und innermathematische Aufgaben zu Inhaltsbereichen „Satz des Pythagoras“ und „Lineare Funktionen“ verwendet.

Aufgabenschwierigkeiten wurden über die Lösungshäufigkeiten eingeschätzt. Für die Messungen der affektiven Merkmale wurden 5-stufige Likert-Skalen (von „trifft voll zu“ bis „trifft überhaupt nicht zu“) aus Schukajlow et al. (2012) adaptiert. In der Experimentalgruppe 1 war die Aussage zum Interesse: „Es wäre interessant diese Aufgabe zu bearbeiten“ und in der Experimentalgruppe 2: „Es war interessant, diese Aufgabe zu bearbeiten“. Diese Aussagen wurden zu jeweiligen Aufgaben gestellt und so eine Skala zum aufgabenbezogenen Interesse z.B. an Modellierungsaufgaben konstruiert. Analoge Skalen wurden auch zu Freude, Langeweile gebildet. Alle Skalen zeigten gute bis sehr gute Reliabilität.

Ergebnisse und Diskussion: In beiden Experimentalgruppen korreliert die Aufgabenschwierigkeit negativ mit Interesse und Freude sowie positiv mit der Langeweile. Somit soll die Aufgabenschwierigkeit bei den künftigen Studien mit aufgabenbezogenen Messungen stärker beachtet werden. Für die Beantwortung der zweiten und der dritten Forschungsfrage wurde eine ANCOVA mit Messwiederholungen mit den Faktoren Aufgabenbearbeitung, Aufgabentyp und der Kovariate Aufgabenschwierigkeit durchgeführt. Es stellte sich heraus, dass aufgabenbezogenes Interesse nach der Aufgabenbearbeitung deutlich abnimmt. Hingegen ändern sich Freude und Langeweile nicht durch die vorherige Bearbeitung von Aufgaben. Interesse und Emotionen zu verschiedenen Aufgabentypen unterscheiden sich nicht signifikant und nur unwesentlich voneinander. Es wurden keine Interaktionseffekte von Aufgabenbearbeitung und Aufgabentyp gefunden, so dass man von vergleichbaren Änderungen von Interesse, Freude und Langeweile nach der Aufgabenbearbeitung bei Aufgaben mit und ohne Realitätsbezug ausgehen kann.

Blum, W., Galbraith, P. L., Henn, H.-W., & Niss, M. (2007). Modelling and applications in mathematics education. The 14th ICMI study. New York: Springer.

Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, 33(1), 9-33.

Hidi, S., & Renninger, K. A. (2006). The four phase model of interest development. *Educational Psychologist*, 41(2), 111-127.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does Interest Matter? The Relationship Between Academic Interest and Achievement in Mathematics. *Journal for Research in Mathematics Education*, 32(5), 448-470

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, 15, 381-395.

Krug, A., & Schukajlow, S. (2013). Problems with and without connection to reality and students' task-specific interest. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (3), 209-216. Kiel, Germany: PME.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531-549.

Pekrun, R., Vom Hofe, R., Blum, W., Goetz, T., Wartha, S., Frenzel, A. C., & Julien, S. (2007). Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistungen in der Sekundarstufe I. In M. Prenzel & L. Allolio-Naecke (Eds.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*, 21-52. Münster: Waxmann.

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics*, 79(2), 215-237.



## **Möglichkeiten empirischer Erfassung und Beschreibung mathematischer (Teil-)Kompetenzen durch schriftliche Leistungstests**

Im Rahmen der Verabschiedung und Etablierung verbindlicher, kompetenzorientierter Bildungsstandards für das Fach Mathematik für den Mittleren Bildungsabschluss und die Allgemeine Hochschulreife (Kultusministerkonferenz, 2003; Kultusministerkonferenz, 2012) und hiermit einhergehend mit der Durchführung nationaler Vergleichs- und Abschlussarbeiten sowie internationaler Schülerleistungsstudien ist es von zentralem Interesse, mathematische Kompetenzen möglichst detailliert messen zu können, um verlässliche Aussagen über den Leistungsstand von Schülerinnen und Schülern treffen und weiteres Lernen bestmöglich unterstützen zu können. Mit Blick auf in diesen Kontexten herangezogene schriftliche Schülerleistungstests stellt sich dabei insbesondere die Frage nach Möglichkeiten und Grenzen empirischer Erfassung und Beschreibung mathematischer (Teil-)Kompetenzen von Schülerinnen und Schülern (Ehmke, Leiss, Blum, & Prenzel, 2006; Klieme, Leutner & Kenk, 2010).

Im Kontext der beiden interdisziplinären DFG-Forschungsprojekte DISUM (Leiss, Blum, & Messner, 2007; Schukajlow et al., 2009) und Co<sup>2</sup>CA (Bürgermeister et al., 2011) sind schriftliche, kompetenzorientierte Schülerleistungstests zu zwei prozessbezogenen Kompetenzen (mathematisches Modellieren und symbolisch, formales, technisches Arbeiten) sowie zwei mathematischen Inhaltsbereichen (Lineare Gleichungssysteme und Satzgruppe des Pythagoras) entwickelt worden. Ziel dieser Entwicklung war die empirische Erfassung von Schülerleistungen zur Beschreibung von Bedingungsfaktoren für ein Gelingen kompetenzorientierten Mathematikunterrichts. Auf der Basis hier erfolgter Testungen von etwa 300 Schülerinnen und Schülern (DISUM) bzw. etwa 1500 Schülerinnen und Schülern (Co<sup>2</sup>CA) neunter Realschulklassen kann im Kontext einer Auseinandersetzung mit Möglichkeiten empirischer Erfassung und Beschreibung mathematischer (Teil-)Kompetenzen folgenden konkreten Forschungsfragen nachgegangen werden:

1. Inwieweit lassen sich verschiedene mathematische Kompetenzen innerhalb schriftlicher Leistungstests empirisch zuverlässig erfassen und voneinander trennen sowie inhaltlich beschreiben?
2. Inwieweit stellen – aus empirischer Sicht – mathematische Kompetenzen ein vom mathematischen Inhaltsbereich abhängiges Konstrukt dar?
3. Inwieweit können innerhalb schriftlicher Leistungstests Teilkompetenzen einer mathematischen Kompetenz erfolgreich gemessen werden?

Unter Rückgriff auf quantitative Analysen der eingesetzten Schülerleistungstests legen erste Ergebnisse nahe: (1) Sich aus theoretischer Sicht unterscheidende mathematische Kompetenzen sind auch empirisch voneinander trennbar, einzelne Kompetenzdimensionen können dabei erfolgreich gemessen und durch kompetenzorientierte Aufgaben in der Breite inhaltlich beschrieben werden. (2) Mathematische Kompetenzen bilden empirisch betrachtet ein vom Inhaltsbereich keineswegs losgelöstes, sondern durchaus abhängiges Konstrukt. (3) Vor allem durch eine gezielte Variation von Aufgaben- und Antwortformat sowie Kodieranweisungen können einzelne Teilkompetenzen einer mathematischen Kompetenz innerhalb schriftlicher Leistungstests erfolgreich gemessen werden. Insbesondere letztgenannte Ergebnisse zu Forschungsfragen (2) und (3) implizieren dabei Fragen nach möglichen Schlussfolgerungen für eine erfolgsversprechende Umsetzung kompetenzorientierten Mathematikunterrichts in der Schule, welche es zu diskutieren gilt.

Bürgermeister, A., Klimczak, M., Klieme, E., Rakoczy, K., Blum, W., Leiss, D., Harks, B. & Besser, M. (2011). Leistungsbeurteilung im Mathematikunterricht – Eine Darstellung des Projekts „Nutzung und Auswirkungen der Kompetenzmessung in mathematischen Lehr-Lernprozessen“. In A. Fächter & K. Moegling (Hrsg.), Theorie und Praxis der Schulpädagogik. Band 16. Diagnostik und Förderung (S. 28-51). Immenhausen: Prolog-Verlag.

Ehmke, T., Leiss, D., Blum, W. & Prenzel, M. (2006). Entwicklung von Testverfahren für die Bildungsstandards Mathematik. Rahmenkonzeption, Aufgabenentwicklung, Feld- und Haupttests. Unterrichtswissenschaft, 34(3), 220-238.

Klieme, E., Leutner, D. & Kenk, M. (Hrsg.). (2010). Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. Weinheim, Basel: Beltz. Kultusministerkonferenz (Hrsg.). (2003). Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. München: Luchterhand. Kultusministerkonferenz (Hrsg.). (2012). Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife. Online unter [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2012/2012\\_10\\_18-Bildungsstandards-Mathe-Abi.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf) [25.10.13].

Leiss, D., Blum, W. & Messner, R. (2007). Die Förderung selbständigen Lernens im Mathematikunterricht – Problemfelder bei ko-konstruktiven Lösungsprozessen. Journal für Mathematik-Didaktik (JMD), 28(3/4), 224-248.

Schukajlow, S., Blum, W., Messner, R., Pekrun, R., Leiss, D. & Müller, M. (2009). Unterrichtsformen, erlebte Selbständigkeit, Emotionen und Anstrengung als Prädiktoren von Schülerleistungen bei anspruchsvollen mathematischen Modellierungsaufgaben. Unterrichtswissenschaft, 37(2), 164-186.



## **Simultane Messung von mathematischen Leitideen und Kompetenzen**

Im Jahr 2003 wurden von der Kultusministerkonferenz die bundesweit geltenden Bildungsstandards in Mathematik für den mittleren Schulabschluss verabschiedet. Deren Erreichen wird regelmäßig durch Testverfahren empirisch überprüft. Den Bildungsstandards in Mathematik liegt ein mehrdimensionales theoretisches Modell zugrunde, das die mathematische Kompetenz in fünf Inhaltsbereiche, die sogenannten Leitideen, und sechs Kompetenzen gliedert. Psychometrisch wurde bisher lediglich ein fünfdimensionales Item-Response-Theory-Modell (IRT-Modell) zur Messung der Leitideen verwendet (vgl. Pant, Stanat, Schroeders, Roppelt, Siegle, & Pöhlmann, 2013), da eine hinreichend reliable Abbildung der Kompetenzen im Gegensatz zu den Leitideen schwierig ist. Gerade aber die Kompetenzen sind für unterrichtsrelevante Rückmeldungen von hohem Interesse. Diese Problemstellung aufgreifend wird in der vorliegenden Studie der Frage nachgegangen, ob mit einer Kombination aus theoriebasierter und psychometrisch-explorativer Skalenbildung neben den fünf Leitideen auch die sechs Kompetenzen mit akzeptabler Reliabilität abgebildet werden können. Die Belastbarkeit der mit dem resultierenden 11-dimensionalen Modell gewonnenen multidimensionalen Fähigkeitsverteilung wird anhand der Untersuchung von Geschlechtsunterschieden analysiert. Diesbezüglich wird zum einen untersucht, inwieweit bei der gemeinsamen Messung aller 11 Dimensionen mit optimierten Skalen für die sechs Kompetenzen die Ergebnisse der fünf-dimensionalen Skalierung zur Abbildung von Geschlechterunterschieden hinsichtlich der Leitideen repliziert werden können. Zum anderen wird untersucht, ob sich signifikante Unterschiede in den Kompetenzen in Abhängigkeit vom Geschlecht ergeben. Da bezüglich der Kompetenzen keine Referenzwerte existieren, soll die Plausibilität der Ergebnisse mit bislang vorliegenden Befunden zu Geschlechtsunterschieden hinsichtlich mathematischer Kompetenzen diskutiert werden. Die Fragestellung wird auf Basis der Daten aus der Entwicklung eines Testverfahrens zur Überprüfung des Erreichens der Anforderungen der Bildungsstandards in Mathematik für den mittleren Schulabschluss (Prenzel & Blum, 2007; N = 9577) untersucht. Das Antwortverhalten der Schülerinnen und Schüler auf die insgesamt 290 Items wurde mittels eines zweiparametrischen 11-dimensionalen IRT-Modells (M2PL) mit Mehrfachladungen modelliert, wobei die Lösung eines Items in der Regel eine gewisse Fähigkeitsausprägung sowohl auf einer Leitidee als auch mindestens einer Kompetenz erfordert. Im ersten Schritt wurden ausgehend von der bei der Testentwicklung getroffenen theoretischen Zuordnung in einem iterativen Vorgehen psychometrisch tragfähige Ladungen der Items auf den sechs Kompetenzen ermittelt. Die Einschätzung von Ladungen als psychometrisch tragfähig erfolgte sowohl nach statistischen (z.B. Höhe der Diskrimination) als auch für die Testung praktischen Kriterien. In einem zweiten Schritt wurden die Antwortdaten mit der optimierten Kompetenzzuordnung und fixierten Itemparametern mit dem M2PL skaliert, um Schätzungen für die latente 11-dimensionale Fähigkeitsverteilung in Form von Plausible Values (Mislevy & Sheehan, 1987) zu erhalten. Die Skalierung erfolgte dabei mit einem sogenannten Hintergrundmodell, welches Zusammenhänge der 11 Fähigkeitsdimensionen mit Variablen wie Schulform, Bundesland, Geschlecht, Migrationsstatus sowie dem sozialen, ökonomischen und kulturellen Status modelliert. Als Referenzbedingung wurde die Skalierung für das fünfdimensionale Modell mit Hintergrundmodell zur Abbildung der Leitideen durchgeführt. Im Anschluss wurden zur Analyse von Unterschieden in Abhängigkeit vom Geschlecht Mittelwerte für beide Subgruppen auf den fünf Leitideen sowie sechs Kompetenzen berechnet. Den Ergebnissen zufolge lassen sich die konsistent höheren mittleren mathematischen Fähigkeiten der Jungen bezüglich der Leitideen in der fünfdimensionalen Modellierung im 11-dimensionalen Modell größtenteils replizieren. Im Hinblick auf die sechs Kompetenzen weisen Jungen im Mittel eine signifikant höhere Ausprägung auf den Kompetenzen „mathematisch modellieren“ sowie „mathematische Darstellungen verwenden“ auf, während die Mädchen signifikant höhere Ausprägungen bei den Kompetenzen „mit symbolischen, formalen und technischen Elementen der Mathematik umgehen“ sowie „mathematisch kommunizieren“ erreichen. Die Ergebnisse zeigen, dass eine Kombination aus theoriebasierter und psychometrisch-explorativer Skalenbildung die gemeinsame Messung von Leitideen und Kompetenzen ermöglicht. Die Plausibilität der Ergebnisse in Bezug auf die Kompetenzen wird abschließend diskutiert, wobei sowohl empirische Erkenntnisse zu Geschlechterunterschieden in Teilbereichen der mathematischen Kompetenz als auch mathematikdidaktische Expertise einbezogen werden.

Mislevy, R. J., & Sheehan, K. M. (1987). Marginal estimation procedures. In A. E. Beaton (Hrsg.), *The NAEP 1983-1984 Technical Report* (Report No. 15-TR-20). Princeton: Educational Testing Service.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Münster: Waxmann.

Prenzel, M., & Blum, W. (2007). Entwicklung eines Testverfahrens zur Überprüfung der Bildungsstandards in Mathematik für den Mittleren Schulabschluss (Technischer Bericht).





## **Mehrdimensionale Modelle der Item-Response Theorie zur Abbildung nicht-kompensatorischer mathematischer Antwortprozesse**

Um psychologische Konstrukte wie Fähigkeiten oder Kompetenzen überhaupt messen zu können, gilt es zunächst, ein theoretisches Modell zu formulieren. Dieses enthält Annahmen über die Struktur des interessierenden Konstrukts und mögliche Zusammenhänge zwischen einzelnen Bestandteilen. Im einfachsten Fall ist dieses Modell eindimensional, häufig sind jedoch komplexere Modelle wünschenswert oder notwendig. Auch bei der Messung mathematischer Kompetenzen werden zugrunde liegende Teilfähigkeiten definiert, häufig bezogen auf Kriterien wie Inhalte oder kognitive Prozesse. Die Bildungsstandards für Mathematik beispielsweise beschreiben mathematische Kompetenz als einen „Verbund von Kompetenzen“, die hinsichtlich einer Prozess-, einer Inhalts- und einer Anspruchsdimension kategorisiert werden können (Roppelt, Blum, & Pöhlmann, 2013). Ist ein solches theoretisches Modell entwickelt, werden Testitems konstruiert, die jeweils einzelne Dimensionen messen, das heißt, für deren Lösung jene Teilfähigkeit in einem bestimmten Ausmaß vorhanden sein muss. In diesem Kontext ist es jedoch auch möglich, dass die Lösungswahrscheinlichkeit für ein einzelnes Item von verschiedenen Teilfähigkeiten gleichzeitig abhängt. Die Bildungsstandards beispielsweise ordnen jedes einzelne Item (mindestens) je einer der fünf Inhaltsbereiche und einer der sechs kognitiven Prozesse zu. Jedes Item misst somit mehrere Dimensionen gleichzeitig. Diese Struktur wird auch als „within-Item-Mehrdimensionalität“ bezeichnet. Bei der statistischen Modellierung gilt es nun zu überlegen, wie diese Dimensionen verknüpft werden sollen, also zur Lösung des Items beitragen. Das meist verbreitete Modell zur Abbildung von within-Item-Mehrdimensionalität, das mehrdimensionale IRT (Item Response Theorie) Modell, wurde in der Vergangenheit als „kompensatorisch“ beschrieben, da die Ausprägung der beteiligten Teilfähigkeiten additiv zur modellierten Lösungswahrscheinlichkeit beitragen. Hohe Ausprägungen einer Teilfähigkeit können somit niedrige Ausprägungen in anderen Teilfähigkeiten ausgleichen. Diese Eigenschaft ist jedoch dann nicht plausibel, sobald das gemeinsame Vorhandensein mehrerer Teilfähigkeiten angenommen wird. So sollen zur Lösung eines Items der Bildungsstandards beispielsweise sowohl das Können des Inhaltsbereichs als auch des beteiligten kognitiven Prozesses vorhanden sein. Sobald eine der beiden Dimensionen eine niedrige Ausprägung aufweist, ist eine korrekte Lösung, selbst bei hoher Ausprägung der anderen Fähigkeit, nicht zu erwarten. Ein psychometrisches Modell sollte jene theoretischen Annahmen reflektieren. Das Ziel dieser Studie ist es, einen alternativen Modellierungsansatz vorzustellen und anhand empirischer Daten aus der Mathematik zu illustrieren. Dafür erweitern wir das additive Modell um die Interaktion der beteiligten Dimensionen. Als Folge der nicht-kompensatorischen Modellierung werden hohe Lösungswahrscheinlichkeiten erst dann vorhergesagt, wenn alle beteiligten Dimensionen zunehmend hoch ausgeprägt sind. Als Grundlage dienen Daten aus PISA 2012, in denen Mathematik das Schwerpunktthema bildet. In einer vorangegangenen Studie (Buchholz & Hartig, 2013) konnten wir die Überlegenheit des hier vorgeschlagenen Interaktionsmodells gegenüber einem kompensatorischen Modell anhand von Daten aus PISA 2009 zeigen. In Anlehnung an Walker und Beretvas (2003) wurden Mathematik-Items hinsichtlich ihres Antwortformats (offen, geschlossen) unterschieden. Während die geschlossenen (Multiple Choice) Items ausschließlich das interessierende Konstrukt, mathematische Fähigkeit, messen, wurde im Modell angenommen, dass die offenen Items zusätzlich eine produktiv-kommunizierende Fähigkeit erfordern, da für deren Lösung eine längere Antwort erforderlich war. Anhand der Ergebnisse konnte gefolgert werden, dass zur Lösung offener Items tatsächlich beide Fähigkeiten vorhanden sein müssen. Die Ergebnisse unterstreichen die Notwendigkeit, dass psychometrische Modelle theoretische Annahmen über das Zusammenspiel der beteiligten Dimensionen berücksichtigen müssen.

Buchholz, J. & Hartig, J. (2013). Modellierung des Antwortformats als genesteten Faktor in einem MIRT-Modell mit latenter Interaktion. Vortrag auf der 11. Tagung der Fachgruppe „Methoden und Evaluation“ der Deutschen Gesellschaft für Psychologie. Klagenfurt, Österreich.

Roppelt, A., Blum, W. & Pöhlmann, C. (2013). Beschreibung der untersuchten mathematischen Kompetenzen. In Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Münster: Waxmann.

Walker, C.M. & Beretvas, S.N. (2003). Comparing Multidimensional and Unidimensional Proficiency Classifications: Multidimensional IRT as a Diagnostic Aid. *Journal of Educational Measurement*, 40 (3), 255-275.



Chair(s): Dominik Leiss (Leuphana Universität Lüneburg), Tina Seidel (TUM School of Education München)

Diskutant/in: Cornelia Gräsel (Universität Wuppertal)

Die Expertise von Lehrpersonen spielt für das Lernen im Unterricht eine entscheidende Rolle (Baumert & Kunter, 2006). Da Lerninhalte primär über Sprache vermittelt werden, nimmt die Qualität der schriftlichen wie mündlichen Kommunikation im Klassenzimmer, aber auch zwischen allen Beteiligten in Bildungsprozessen (Lehrkräfte, Eltern, Kollegien) einen zentralen Stellenwert ein. Betrachtet man die Ergebnisse einer Zusammenfassung mehrerer Meta-Analysen zur Schul- und Unterrichtseffektivität (Hattie, 2009), rangieren Aspekte der Interaktionen zwischen Lehrenden und Lernenden, z.B. Feedback, kognitiv aktivierende Fragen oder formatives Assessment, auf den „vordersten Plätzen“ der berichteten Effektstärken. Trotz dieses Wissens um die Bedeutung von Kommunikation und Feedback weisen die existierenden Studien darauf hin, dass bezüglich der unterrichtlichen Umsetzung einer lernförderlichen Kommunikation noch Optimierungsbedarf besteht. So zeigen sich in der Regel Befunde von eng geführten und von der Lehrkraft kontrollierten Klassen- (Seidel & Prenzel, 2006) bzw. Gruppengesprächen (Leiss, 2010) sowie eine deutlich eingeschränkte Umsetzung von lernförderlichem Feedback an die Lernenden (Besser et al., 2010); beides verbunden mit negativen Effekten auf die motivationale und kognitive Entwicklung der Lernenden.

Vor diesem Hintergrund zielen einige Studien darauf ab, auf der Basis kontrollierter Interventionen Prozesse der Veränderung von Handlungspraktiken und die Stärkung der Lehrerexpertise in Bezug auf Kommunikation und Feedback zu untersuchen. Ziel des Symposiums ist es, Befunde aus diesen Studien zu präsentieren und darzulegen, unter welchen Bedingungen sowohl in der universitären Lehrerbildung als auch in Lehrerfortbildungen gezielte Weiterentwicklung von Expertise im kommunikativen Bereich erzielt werden kann. Um die Breite der unterschiedlichen Aspekte abzudecken, kommen die Beiträge aus der fachdidaktischen, erziehungswissenschaftlichen und pädagogisch-psychologischen Forschung.

Dazu wird im ersten Beitrag die Frage beantwortet, ob Lehramtsstudierende im Rahmen einer Intervention die Fähigkeit entwickeln, relevante Aspekte von Kommunikation und Feedback im Unterrichtsprozess zu identifizieren und auf der Basis ihres professionellen Wissens zu interpretieren (Krammer). Im zweiten Beitrag werden Aspekte der Kommunikation und des Feedbacks auf die Situation der Beratung in Elterngesprächen übertragen. Auf der Basis einer Intervention wurde hier überprüft, ob bereits Studierende grundlegende Fähigkeiten für die Kommunikation in solchen Gesprächssituationen erwerben können (Gartmeier). Der dritte Beitrag untersucht aus einer fachdidaktischen Perspektive die Wirkungen einer Lehrerfortbildung zu formativen Rückmeldungen auf die diesbezügliche Expertise der Lehrpersonen sowie auf die von den Lernenden wahrgenommenen Unterrichtsveränderungen (Leiss). Im vierten Beitrag wurde im Rahmen einer Intervention zur Förderung produktiver Klassengespräche erforscht, ob eine videobasierte Lehrerfortbildung positive Auswirkungen auf kognitiv aktivierende Gesprächselemente hat und welche Schüler davon am stärksten profitieren (Pehmer). Abschließend werden alle Beiträge aus einer übergreifenden Perspektive auf Innovation in der Lehrerbildung diskutiert.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G161

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469–520.

Besser, M., Leiss, D., Harks, B., Rakoczy, K., Klieme, E., & Blum, W. (2010). Kompetenzorientiertes Feedback im Mathematikunterricht: Entwicklung und empirische Erprobung prozessbezogener, aufgabenbasierter Rückmeldesituationen. Empirische Pädagogik, 24(4), 404-432.

Hattie, J. (2009). Visible Learning. A synthesis of over 800 meta-analyses relating to achievement. London & New York: Routledge.

Leiss, D. (2010). Adaptive Lehrerinterventionen beim mathematischen Modellieren – empirische Befunde einer vergleichenden Labor- und Unterrichtsstudie. JMD, 31 (2), 197-226.

Seidel, T., & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: Findings from a video study. Learning And Instruction, 16(3), 228-240.



Kathrin Krammer (Pädagogische Hochschule Luzern), Isabelle Hugener (Pädagogische Hochschule Luzern), Sandro Biaggi (Pädagogische Hochschule Luzern), Gabriela Furrer (Pädagogische Hochschule Luzern)

## **Förderung der professionellen Unterrichtswahrnehmung von angehenden Lehrpersonen anhand von eigenen und fremden Videos**

Theoretischer Hintergrund: Die professionelle Unterrichtswahrnehmung gilt als bedeutsam für erfolgreiches Handeln im Unterricht (Davis, 2006; Seidel & Prenzel, 2007). Ergebnisse zum Zusammenhang der unterrichtsbezogenen Analysekompetenz mit dem erfolgreichen Unterrichten bestätigen die Bedeutung dieser Fähigkeit (Kersting et al., 2012; Sherin & van Es, 2009). Voraussetzung für das Erkennen und Interpretieren von lernrelevanten Situationen sind das Wissen über die Bedingungen eines lernwirksamen Unterrichts und die Fähigkeit, dieses Wissen auf konkrete Situationen im Unterricht zu übertragen. Basierend auf Befunden der Expertiseforschung werden die Ebenen des Beschreibens (bedeutsame Merkmale identifizieren), Erklärens (theoriebasierte Begründung des identifizierten Merkmals) und Vorhersagens (vermutete Wirkung des Merkmals auf Lern-, Verstehens- und Motivationsprozesse der Lernenden) unterschieden (Seidel et al., 2010). Videoaufnahmen von Unterrichtssituationen stellen ein hilfreiches Werkzeug zur Verbindung von theoretischen Kenntnissen und konkreten Situationen dar (z.B. Brophy, 2004). Untersuchungen in der Aus- und Weiterbildung von Lehrpersonen verweisen auf das Potenzial des fallbasierten Lernens mit Videos zur Förderung der Fähigkeit, Unterricht hinsichtlich der Umsetzung lernrelevanter Merkmale zu interpretieren (Borko et al., 2008; Santagata & Guarino, 2011; Seidel et al., 2013; Stürmer et al., 2013). In der Lehrerbildung noch kaum erforscht ist jedoch die Frage der Entwicklung der professionellen Unterrichtswahrnehmung in Abhängigkeit vom videobasierten Nachdenken über eigenen versus fremden Unterricht. Während Videos aus dem eigenen Unterricht die Reflexion über eigene Erfahrungen unterstützen, eröffnen Videos von fremden Lehrpersonen die Möglichkeit, gezielt die Umsetzung von ausgewählten Merkmalen der Unterrichtsqualität der Beobachtung zugänglich zu machen (Baechler et al., 2013). Vorliegende Ergebnisse zeigen, dass die Auseinandersetzung mit eigenen Unterrichtsvideos als authentischer und motivierender wahrgenommen wird. Demgegenüber werden Unterrichtsvideos von fremden Lehrpersonen differenzierter, aber auch kritischer kommentiert (Kleinknecht & Schneider, 2013; Seidel et al., 2011; Zhang et al., 2011).

Fragestellung: Im Projekt VideA (Videos in der Ausbildung von Lehrpersonen) werden die differenziellen Effekte der fallbasierten Arbeit mit eigenen und fremden Videos auf die Entwicklung der professionellen Unterrichtswahrnehmung von Kernmerkmalen eines lernwirksamen Unterrichts (Zielorientierung, Lernbegleitung, Lernatmosphäre) untersucht. Während eines Semesters arbeiteten die Studierenden in kleinen Gruppen entweder mit Videobeispielen von fremden Lehrpersonen oder mit Ausschnitten aus dem eigenen Unterricht und dem von Kolleginnen und Kollegen. Die Diskussion der Videos wurde von Dozierenden entlang eines Vorgehens nach Santagata und Guarino (2011) moderiert, welches die Interpretation der Unterrichtsausschnitte hinsichtlich der Umsetzung lernrelevanter Merkmale unterstützte. In der Kontrollgruppe analysierten die Studierenden Unterrichtsbeispiele anhand von schriftlichen Materialien und Beschreibungen von Unterrichtssituationen nach dem gleichen Vorgehen.

Methode: Zu jeder Bedingung (fremde Videos/eigene Videos/Kontrollgruppe) waren ca. 60 Studierende zufällig zugeteilt. Die Entwicklung der professionellen Unterrichtswahrnehmung wurde mittels Prä-Post-Design mit dem von Seidel und Mitarbeitenden entwickelten Instrument Observer (Seidel et al., 2010; Seidel & Stürmer; eingereicht) untersucht, welches die situationsnahe Erhebung der Fähigkeit zur Interpretation von lernrelevanten Situationen im Unterricht ermöglicht. Das Online-Tool verbindet Unterrichtsvideos mit standardisierten Rating-Items. Da die professionelle Unterrichtswahrnehmung ein wissensbasierter Prozess ist, wurde zudem das Wissen über die fokussierten Unterrichtsmerkmale zu beiden Zeitpunkten erhoben (Stürmer et al., 2013).

Ergebnisse: Die Ergebnisse zu den Wirkungen des fallbasierten Lernens mit Videos in der Ausbildung auf die professionelle Unterrichtswahrnehmung zeigen, dass die professionelle Unterrichtswahrnehmung der Gruppen, welche mit Videos arbeiten, signifikant zunimmt, sowohl insgesamt als auch in den drei Dimensionen Beschreiben, Erklären und Begründen. Demgegenüber ist der Zuwachs der professionellen Unterrichtswahrnehmung in der Kontrollgruppe nicht signifikant. Der Zuwachs des Wissens über die Unterrichtsmerkmale ist in allen drei Gruppen bedeutsam. Die Ergebnisse deuten damit darauf hin, dass das Wissen über Merkmale eines lernwirksamen Unterrichts eine wichtige, jedoch nicht hinreichende Bedingung für die Entwicklung der professionellen Unterrichtswahrnehmung darstellt und das fallbasierte Lernen mit Videos zur Förderung der professionellen Unterrichtswahrnehmung beitragen kann. Die Arbeit mit eigenen oder fremden Videos zeigt keine differenziellen Effekte, beide Typen scheinen sich bereits zu Beginn der Ausbildung zu eignen.

Baechler, L., Kung, S.-C., Jewkes, A. M., & Rosalia, C. (2013). The role of video for self-evaluation in early field experiences. *Teaching and Teacher Education*, 36, 189-197.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Videos as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, 24(2), 417-436.

Brophy, J. (2004). *Using video in teacher education*. Oxford: Elsevier.

Davis, E. A. (2006). Characterizing productive reflection among preservice elementary teachers: Seeing what matters. *Teaching and Teacher Education*, 22, 281-301.

Kersting, N.B., Givvin, K.B., Thompson, B.J., Santagata, R., & W. Stigler, J.W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, 49(3), 568-589.

Kleinknecht, M., & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teaching and Teacher Education*, 33, 13-23.

Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM The International Journal on Mathematics Education*, 43(1), 133-145.

Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching and Teacher Education*, 34, 56-65.

Seidel, T., Blomberg, G., & Stürmer, K. (2010). "Observer" – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Zeitschrift für Pädagogik, Beiheft* 56(296-306).

Seidel, T., & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. *Zeitschrift für Erziehungswissenschaft*, 10, Sonderheft 8, 201-216.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, 27, 259-267.

Sherin, M.G., & van Es, E.A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1), 20-37.

Stürmer, K., Könings, K. D., & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: Effect of courses in teaching and learning. *British Journal of Educational Psychology*, 83, 467-483.

Zhang, M., Lundeberg, M., Koehler, M.J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education*, 27, 454-462.



Martin Gartmeier (TUM School of Education München), Johannes Bauer (TUM School of Education München),  
Anne Wiesbeck (TUM School of Education München), Manfred Prenzel (TUM School of Education München)

## **Förderung von Expertise im Bereich Kommunikation: Differenzielle Effekte der Implementation eines E-Learning Moduls in ein interdisziplinäres Trainingsprogramm**

Hintergrund: Lehrpersonen müssen ihrem beruflichen Alltag vielfältige kommunikative Situationen meistern (Hertel, Bruder, & Schmitz, 2009), u.a. im Kontakt mit Eltern (Sacher, 2008). Ähnliches gilt für den Bereich der Medizin, bzw. den Austausch zwischen ÄrztInnen und PatientInnen (Beck & Daughtridge, 2002). Im Gegensatz zur Lehramtsausbildung sind in der Medizin Gesprächsführungstrainings häufig Bestandteil des Studiums, wobei primär auf klassische Trainingsformen zurückgegriffen wird, z.B. auf Rollenspiele. Im Gegensatz dazu ist der Einsatz von E-Learning Formaten für Gesprächsführungstrainings noch weniger verbreitet, obwohl sich hier relevante Inhalte über Videos anschaulich darstellen ließen. Im Rahmen des vorliegenden Projekts wurde deshalb der Versuch unternommen, bereits im Rahmen der universitären Ausbildung einen Grundstein für die Entwicklung von Gesprächsführungskompetenz in beiden Domänen, dem Lehramt sowie der Medizin, zu legen und hierzu auch E-Learning Bausteine einzusetzen. In Kooperation mit KollegInnen aus der Medizindidaktik wurde dazu ein parallel aufgebautes, aber domänenspezifisch ausgearbeitetes Trainingsprogramm entwickelt, das Trainingsmaterialien in zwei äquivalenten Versionen für beide Bereiche umfasst. Beide Versionen des Trainings rekurren auf das Münchner Modell der Gesprächsführungskompetenz (MMG) (Gartmeier et al., 2011).

Fragestellungen: Vor allem im Hinblick auf den zukünftigen Einsatz des Lernmoduls in der Hochschullehre wirft die vorliegende Studie die Frage auf, welche Variante des Einsatzes des E-Learning Moduls am vielversprechendsten ist hinsichtlich des Interesses der Teilnehmenden an den jeweiligen Lerninhalten sowie ihrer inhaltspezifischen Anstrengungsbereitschaft (Ramm et al., 2006). Außerdem fragen wir nach Effekten des Trainings bzgl. der selbst eingeschätzten Gesprächsführungskompetenz im Führen von Eltern, bzw. Patientengesprächen. Der vorliegende Beitrag untersucht differenzielle Effekte dreier Einbettungsvarianten eines E-Learning Moduls in das Trainingsprogramm. Das Lernmodul bestand aus drei Teilen: (1) Grundlagen der Gesprächsführung, (2) Umsetzung eines Beratungsgesprächs zur Entscheidungsfindung (PEF) und (3) videobasierte Reflexion zu PEF-Gesprächen. Im Rahmen der Teile 1 und 3 wurden jeweils aufwendig produzierte Lehrvideos eingesetzt. In Variante A hatten die Trainees die volle Trainingszeit (300 Minuten) um die drei Teile des E-Learning Moduls zu durchlaufen. In Variante B wurde das E-Learning Modul in 150 Minuten absolviert, anschließend wurde in weiteren 150 Minuten ein rollenspielbasiertes Trainingsmodul durchlaufen. In Variante C führten die TeilnehmerInnen zuerst ein simuliertes Eltern-, bzw. Patientengespräch und absolvierten dann das E-Learning Modul in 150 Minuten mit anschließenden Rollenspielen.

Methoden: Im Rahmen der experimentellen Studie wurden N=125 Personen aus den Domänen Medizin (nMedizin=54) und Lehramt (nLehramt=71) trainiert. Die Trainees wurden randomisiert den Trainingsbedingungen zugewiesen (nKombination=43; ne-learning=42; nKontrolle=40). Jeweils nach Absolvierung der Teile 1-3 des E-Learning Moduls wurde das Interesse der Teilnehmenden an den Lerninhalten sowie ihre inhaltspezifische Anstrengungsbereitschaft (Ramm et al., 2006) mithilfe eines Fragebogens erhoben. Außerdem wurde vor Beginn sowie am Ende des Trainings die Selbsteinschätzung der Gesprächsführungskompetenz der Studierenden erfasst.

Ergebnisse: Die Ergebnisse zeigen, dass sich das Interesse an den Inhalten über die verschiedenen Trainingsvarianten und Domänen hinweg tendenziell positiv von Modul 1 (Grundlagen) hin zu Modul 2 (Partizipative Entscheidungsfindung) und Modul 3 (Anwendung) entwickelte. Die jeweils höchsten Interessenwerte zeigten sich dabei in der Kombinationsbedingung, in der für das E-Learning 150 Minuten zur Verfügung standen und die von einem Rollenspielmodul gefolgt wurde. Bzgl. der prä-post Messung der Selbsteinschätzung der Gesprächsführung zeigten sich bedeutsame, positive Effekte über die verschiedenen Trainingsbedingungen hinweg (alle  $p < .01$  bei R2-Werten um .10). Lediglich Variante 1 des Trainings (reines E-Learning) im Bereich Medizin fiel hier aus dem Rahmen, da sich nur eine marginale, nicht signifikante Verbesserung zeigte.

Diskussion: Die Absolvierung des E-Learning-Moduls zeigt positive Effekte bzgl. der Selbsteinschätzung der Gesprächsführungskompetenz, besonders in Kombination mit Rollenspielen. So scheint z.B. die Nutzung des E-Learning-Moduls als Vorbereitung für ein Präsenzseminar vielversprechend. Eine für uns wichtige Diskussionsfrage ist, wie sich die kognitiven Prozesse im Kontext des E-Learning im Rahmen zukünftiger Studien genauer beleuchten lassen, auch im Kontrast zu anderen Trainingsformaten, wie etwa Rollenspielen.

- Beck, R. & Daughtridge, R. (2002). Physician-Patient Communication in the Primary Care Office: A Systematic Review. *Journal of the American Board of Family Medicine*, 15(1), 25-38.
- Gartmeier, M., Bauer, J., Fischer, M. R., Karsten, G. & Prenzel, M. (2011). Modellierung und Assessment professioneller Gesprächsführungskompetenz von Lehrpersonen im Lehrer-Elterngespräch. In O. Zlatkin-Troitschanskaia (Hrsg.), *Stationen empirischer Bildungsforschung. Traditionslinien und Perspektiven*. (S. 412-424). Wiesbaden: VS.
- Hertel, S., Bruder, S. & Schmitz, B. (2009). Beratungs- und Gesprächsführungskompetenz von Lehrkräften. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung*. Weinheim: Beltz.
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, A. C., Leutner, D., et al. (2006). PISA 2003. Dokumentation der Erhebungsinstrumente. Münster: Waxmann.
- Sacher, W. (2008). *Elternarbeit; Gestaltungsmöglichkeiten und Grundlagen für alle Schularten*. Bad Heilbrunn: Klinkhardt.

## **Professionswissen von Lehrkräften in einem kompetenzorientierten Mathematikunterricht: Lehrerfortbildungen zum formativen Assessment**

Theoretischer Hintergrund: Im Rahmen der allgemeinen Diskussion darüber, wie kompetenzorientierter (Mathematik-)Unterricht in Schule verbessert bzw. implementiert werden kann (u.a. Blomhøj & Jensen, 2007; Niss, 2003), stellt sich auf Seite der Lehrpersonen die handlungspraktische Frage, wie insbesondere in solch einem Unterricht eine lernförderliche Leistungsbeurteilung und -rückmeldung zu gestalten ist. Während üblicherweise am Ende einer Lerneinheit den Lernenden ihre Leistung in Form einer Note, die nur einen geringen inhaltlichen Einfluss auf den weiteren Lernprozess hat (Sadler, 1989, S. 120), zurückgemeldet wird, erscheint eine Form des formativen Assessments – also das lernprozessbegleitende Erfassen und lernförderliche Rückmelden von Schülerleistungen – besonders geeignet, um die Kompetenzentwicklung von Schülerinnen und Schülern positiv zu beeinflussen (Baker, 2007; Black & William, 2009; Shepard, 2000). Dabei gehen Hattie & Timperley (2007) davon aus, dass eine derartige Leistungsrückmeldung den Lernenden insbesondere auf folgende Fragen antworten liefern muss: “Where am I going? How am I going? and Where to next?”. Inwiefern die unterrichtliche Umsetzung eines solchen formativen Assessments gelingt, hängt, wie generell die Qualität von unterrichtlichen Lehr-Lernprozessen (Baumert et al., 2010), mit dem (spezifischen) fachlichen und fachdidaktischen Wissen der Lehrpersonen zusammen. Fragestellungen: (1) Ist es möglich, eine Lehrerfortbildung zu entwickeln und durchzuführen, die fachdidaktisches (pck) sowie allgemeinpädagogisches Wissen (pk) über formatives Assessment an Lehrpersonen vermittelt? (2) Ist es möglich, einen reliablen und validen Test (pck & pk) zu entwickeln, der geeignet ist, um eine derartige Lehrerfortbildung zu evaluieren? (3) Hat eine Lehrerfortbildung, in der die Lehrpersonen einen Zuwachs an (fachspezifischem) Wissen zum Thema formatives Assessment erworben haben, einen Einfluss auf den Mathematikunterricht?

Methode: Im Rahmen der hier vorgestellten Co2CA-Studie nahmen 69 Lehrkräfte von Februar 2013 bis Dezember 2013 an Lehrerfortbildungen teil (siehe Abbildung 2). Die teilnehmenden Lehrkräfte wurden dabei einer von zwei Untersuchungsbedingungen zugeteilt. Im Zentrum der Untersuchungsbedingung A standen Ideen zu Möglichkeiten und Grenzen formativen Assessments in einem kompetenzorientierten Mathematikunterricht und in Untersuchungsbedingung B wurden allgemein-fachdidaktische Grundlagen zu einem kompetenzorientierten Mathematikunterricht vermittelt. Aus organisatorischen Gründen wurden die Fortbildungen beider Untersuchungsbedingungen in zwei Erhebungswellen durchgeführt werden.

Die Wirksamkeit der Fortbildungen wird durch mehrfach stattfindende schriftliche Erhebungen sowohl auf Lehrer- als auch auf Schülerebene erfasst. So wird (1) das Vorwissen der Lehrkräfte durch einen fachdidaktischen Wissenstests des Forschungsprojekts COACTIV (Krauss et al., 2008) sowie durch bezüglich der Fortbildungsinhalte sensitive CK-, PCK- und PK-Items kontrolliert. Kontinuierlich begleitet werden die Lehrerfortbildungen (2) durch Lehrerfragebögen u. a. zur Wahrnehmung der Fortbildungen sowie (3) durch Schülerfragebögen u. a. zur Wahrnehmung der Unterrichtsqualität.

Ergebnisse: Analysen vorliegender Erhebungsdaten zu den Lehrerfortbildungen von Februar bis Mai 2013 erlauben erste Antworten auf Möglichkeiten und Grenzen von Aufbau und empirischer Erfassung fachdidaktischen Wissens (pck) von Lehrkräften zu zentralen Ideen formativer Leistungsmessung und -rückmeldung (Forschungsfrage 1 und 2): So ist es (1) gelungen, einen Professionswissenstest zu entwickeln, welcher fachdidaktisches Wissen zu formativem Assessment reliabel erfasst (11 Items,  $\alpha = 0.75$ ) und welcher (2) differenzierte Ergebnisse für die beiden Untersuchungsbedingungen UB A1 und UB B1 am Ende der Fortbildungen liefert (siehe Tabelle 1). Im Rahmen manifester Regressionsanalysen wird schließlich unter Berücksichtigung des ebenfalls erhobenen allgemein fachdidaktischen Wissens (Professionswissenstest aus COACTIV, siehe Krauss et al., 2008) als Kovariat deutlich, dass aufgezeigte Leistungsunterschiede zwischen den beiden Untersuchungsbedingungen nicht allein auf das fachdidaktische Wissen der Lehrkräfte zu Beginn der Fortbildungen zurückzuführen ist (standardisierter Regressionskoeffizient = 0.32), sondern insbesondere durch die Fortbildungsbedingung selbst hervorgerufen werden (standardisierter Regressionskoeffizient = 0.62).

N m SD Differenz

Bedingung UB A1 20 4.35 2.23 1.28  $t(38) = 5.99$ ;  $p < 0.01$

Bedingung UB B1 20 1.05 0.95

Tabelle 1: Erste Ergebnisse des fachdidaktischen Professionswissenstests zu formativer Leistungsmessung und -rückmeldung

Baker, E. L. (2007). The end(s) of testing. *Educational Researcher*, 36, 309-317.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133-180.

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.

Blomhøj, M., & Jensen, T. H. (2007). What's all the fuss about competencies? Experiences with using a competence perspective on mathematics education to develop the teaching of mathematical modelling. In W. Blum, P.L.

Galbraith, H.W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education. The 14th ICMI Study* (pp. 45-56). New York: Springer.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.

Kraus, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29(3/4), 223-258.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29, 4-14.





## Wer profitiert vom produktiven Klassengespräch?

### Die Wahrnehmung von Schülern mit unterschiedlichem Fähigkeitsselbstkonzept

Theoretischer Hintergrund: Das Klassengespräch stellt die dominierende Unterrichtsform dar (Seidel, 2011). Die bisherige Forschung zeigt, dass das Klassengespräch dabei von einer hohen Lehrerzentrierung und geringen Schülerbeteiligung geprägt ist (Jurik, Gröschner, & Seidel, 2013). Daraus ergeben sich häufig negative Folgen für das Lernen von Schülern (u.a. Pauli et al., 2008). Für eine produktive Klassengesprächsführung werden insbesondere die Komponenten Schüleraktivierung und Begleitung von Schülerlernprozessen als zentral angesehen (Walshaw & Anthony, 2008). Durch die aktiviere Einbindung von Schülern sind positive Auswirkungen auf situative Lernprozesse (Nachvollziehen, Elaborieren, Organisieren) sowie Elaborationsstrategien von Schülern zu erwarten (Weinstein & Mayer, 1986). Bisherige Befunde zeigen außerdem die hohe Bedeutung von individuellen Unterschieden von Schülern (Corno & Snow, 1986). Das Fähigkeitsselbstkonzept (FSK) ist ein wesentliches individuelles Merkmal, das Verhaltensweisen und Wahrnehmungen von Lernenden steuert (Bandura, 1986). Bezüglich der Wahrnehmung von Lernprozessen scheint es deshalb wesentlich, das Fähigkeitsselbstkonzept von Schülern zu berücksichtigen.

Fragestellungen: (1) Nehmen Schüler einer Interventionsgruppe (IG) ihre situativen Lernprozesse und Elaborationsstrategien im Verlauf eines Schuljahres im Vergleich zu einer Kontrollgruppe (KG) verändert wahr? (2) Hat die Intervention Effekte auf die Wahrnehmung der situativen Lernprozesse und Elaborationsstrategien für Schüler mit einem niedrigem vs. einem hohem FSK?

Methode: Die Intervention (IG) bestand aus dem Angebot einer Fortbildung zum produktiven Klassengespräch im Rahmen von zwei Videozirkeln („Dialogic Video Cycle“). Zunächst wurden Elemente des produktiven Klassengesprächs in einen eigenen Unterrichtsplan integriert. In zwei weiteren Sitzungen reflektierten die Lehrkräfte (n=6) im kollegialen Austausch die produktive Klassengesprächsführung (Schüleraktivierung; Begleitung von Schülerlernprozessen) am eigenen Video. Die Lehrkräfte (n=4) in der Kontrollgruppe (KG) reflektierten ihre Erfahrungen aus herkömmlichen Fortbildungen zum Thema Lehrer-Schüler-Kommunikation zweimal in „Runden Tischen“.

In beiden Gruppen wurden die Schüler (N=226) zu Beginn (Prä) und zu Ende des Schuljahres (Post) zur Wahrnehmung ihrer situativen Lernprozesse, dem Einsatz ihrer Elaborationsstrategien und ihrem Fähigkeitsselbstkonzept auf einer 4-stufigen Likert-Skala (0= „trifft nicht zu“ bis 3= „trifft zu“) befragt.

Für die Forschungsfrage 1 wurden ANCOVAs für beide abhängige Variablen (situative Lernprozesse; Elaborationsstrategien) gerechnet. Für Forschungsfrage 2 kam das gleiche Verfahren jeweils für die Kohorte der Schüler mit geringem und hohem FSK zur Anwendung. Die Unterteilung nach dem FSK wurde mithilfe eines Mediansplits vollzogen.

Aufgrund von gerichteten Hypothesen wurde ein Signifikanzniveau von  $p < .10$  angenommen.

Ergebnisse: Die Befunde zeigen für die erste Forschungsfrage positive Effekte in der Schülerwahrnehmung für die IG sowohl in der Veränderung der situativen Lernprozesse ( $F(1,174) = 4.14, p = .04, \eta^2 = .02$ ) als auch der Elaborationsstrategien ( $F(1,177) = 5.65, p = .02, \eta^2 = .03$ ). Schüler der IG berichteten keine Veränderung ihrer situativen Lernprozesse ( $M\Delta$  Pre/Post = .00,  $SD = .52$ ), wohingegen die KG eine negative Veränderung wahrnahm ( $M\Delta$  Pre/Post = -.18,  $SD = .38$ ). Ihre Elaborationsstrategien stuften Schüler der IG am Ende des Schuljahres höher ein ( $M\Delta$  Pre/Post = .19,  $SD = .72$ ), während Schüler der KG eine negative Veränderung berichteten ( $M\Delta$  Pre/Post = -.04,  $SD = .56$ ). Nach Unterteilung der Schüler bezüglich ihres FSK auf der Basis des Prä-Tests, zeigte die Intervention für Lernende mit einem niedrigen FSK hinsichtlich der situativen Lernprozesse einen positiven Effekt ( $F(1,93) = 5.56, p = .02, \eta^2 = .06$ ). Schüler der IG schätzten ihre situativen Lernprozesse zu Ende des Schuljahres höher als zu Beginn ein ( $M\Delta$  Pre/Post = .05,  $SD = .47$ ). Für die Schüler in der KG zeigte sich eine negative Entwicklung ( $M\Delta$  Pre/Post = -.21,  $SD = .36$ ). Für Schüler mit hohem FSK wirkte sich die Intervention positiv hinsichtlich der Elaborationsstrategien aus ( $F(1,76) = 3.24, p = .08, \eta^2 = .04$ ). Sie berichteten einen Anstieg in ihren Elaborationsstrategien ( $M\Delta$  Pre/Post = .26,  $SD = .81$ ), wohingegen die Schüler in der KG eine leichte Verschlechterung wahrnahmen ( $M\Delta$  Pre/Post = -.02,  $SD = .54$ ).

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Corno, L., & Snow, E. R. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.), Handbook of research on teaching (pp. 695-629). New York: Macmillan.

Jurik, V., Gröschner, A., & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. Learning and Instruction(23), 33-42. doi: 10.1016/j.learninstruc.2012.09.002

Pauli, C., Drollinger-Vetter, B., Hugener, I., & Lipowsky, F. (2008). Kognitive Aktivierung im Mathematikunterricht. Zeitschrift für Pädagogische Psychologie, 22(2), 127-133. doi: 10.1024/1010-0652.22.2.127

Seidel, T. (2011). Lehrerhandeln im Unterricht. In E. Terhart, H. Bennewitz & M. Rothland (Eds.), Handbuch der Forschung zum Lehrerberuf (pp. 605-629). Münster: Waxmann.

Walshaw, M., & Anthony, G. (2008). The Teacher's Role in Classroom Discourse: A Review of Recent Research Into Mathematics Classrooms. Review Of Educational Research, 78(3), 516-551. doi: 10.3102/0034654308320292

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), Handbook of research on teaching (pp. 315-327). New York.



**Perspektiven der Förderung des selbstregulierten Lernens:  
Zielgruppen, Förderkonzepte und Evaluationsmethoden**

Chair(s): Sabine Ogrin (Technische Universität Darmstadt)

Diskutant/in: Tina Seufert (Universität Ulm)

Selbstreguliertes Lernen (SRL) hat als wichtige Schlüsselkompetenz einen positiven Einfluss auf die schulische und akademische Leistung (Dignath & Buettner 2008; Richardson, Abraham, & Bond, 2012). Um eine möglichst hohe Zielgruppenpassung von Fördermaßnahmen zu ermöglichen, ist die stetige Weiterentwicklung von Interventionen wie auch Evaluationsmethoden unerlässlich. Im vorgestellten Symposium werden verschiedene Perspektiven auf die Förderung selbstregulierten Lernens dargestellt. Gemeinsamer theoretischer Hintergrund aller Beiträge ist das Selbstregulationsmodell von Zimmerman (2000) bzw. dessen Weiterentwicklung durch Schmitz und Wiese (2006). Dabei werden jedoch in den einzelnen Beiträgen unterschiedliche Zielgruppen betrachtet (Grundschüler der Klassen 2 bis 4, Hauptschüler der Klasse 5 sowie Studierende des Lehramts und Studierende in einem Mathematikvorkurs), sowie unterschiedliche Förderkonzepte durchgeführt (Förderung durch Tagebücher mit Kurzinstruktionen, durch ein web-basiertes Training oder durch Schulung von Lehrkräften, welche wiederum die Schüler trainieren). Nicht zuletzt werden auch unterschiedliche Methoden zur Evaluation der verschiedenen Maßnahmen eingesetzt (Prä-Post-Analysen von Fragebögen und Leistungstests, Prozessanalysen von Tagebüchern sowie Logfile-Analysen einer web-basierten Lernplattform). Im Beitrag von Perels, Bieck, Jesch und Leidinger wird ein Förderkonzept des SRL in der Grundschule vorgestellt. Dabei wurde das Lehrpersonal geschult, welches dann Materialien der Schulung im regulären Schulunterricht verwendete. Anhand von Prä-Post-Analysen von Fragebogen und objektiven Leistungstests auf Schüler- wie auch auf Lehrerebene konnten positive Effekte der Schulung nachgewiesen werden. Im Beitrag von Ogrin und Schmitz wurden ebenso Lehrkräfte geschult, eine Förderintervention im regulären Mathematikunterricht der 5. Klasse der Hauptschule durchzuführen. Die Prozessuale Auswertung der eingesetzten Lerntagebücher zeigt Verbesserungen im SRL bei der Hausaufgabenbearbeitung auf. Im Beitrag von Fabriz und Dignath wurden Lehramtsstudierende über den Verlauf eines Semesters mit einem standardisierten Lerntagebuch gefördert, das zugleich als Messinstrument verwendet wurde. Mit Hilfe von prozessanalytischen Methoden und einer Prä-Post Befragung konnten positive Effekte der Förderung auf unterschiedlichen Variablen nachgewiesen werden. Im Beitrag von Bellhäuser und Schmitz wird ein web-basierter Trainingsansatz dargestellt, der zur Förderung des selbstregulierten Lernens von Studierenden eingesetzt wurde. Die Evaluation erfolgt anhand von Logfiles, die auf einer mathematischen Lernplattform aufgezeichnet wurden und umfasst Aussagen über die Dauer des Lernens, die Regelmäßigkeit, die präferierten Uhrzeiten sowie das Erreichen von selbstdefinierten Zielen. Es werden Aspekte der Optimierung von Zielgruppenpassung der Interventionen wie auch der Evaluationsmethoden diskutiert.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231–264.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387.

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, 31(1), 64–96.

Zimmerman, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, M. Zeidner, & M. Zeider (Eds.), *Handbook of Self-Regulation* (pp. 13–39). San Diego, Calif: Academic Press.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G165

Franziska Perels (Universität des Saarlandes), Jenifer Bieck (Hessisches Landesschulamt und Lehrkräfteakademie), Heike Jesch (Hessisches Landesschulamt und Lehrkräfteakademie), Manuela Leidinger (Universität des Saarlandes)

## ***Transfer einer Interventionsmaßnahme zur Förderung selbstregulierten Lernens in den regulären Unterricht hessischer Grundschulen***

Zielsetzung der Studie war es, die Effekte einer Intervention zur Förderung selbstregulierten Lernens in der vierten Grundschulklasse (Leidinger & Perels, 2012) in den regulären Grundschulunterricht der Klassen 2 bis 4 zu transferieren und damit zu untersuchen, in welcher Jahrgangsstufe welche Selbstregulationsstrategie die größten Effekte sowohl auf das selbstregulierte Lernen als auch auf die Leistung der Schülerinnen und Schüler in Mathematik und Deutsch erzielt.

Theoretische Grundlage der Studie bildete dabei das Selbstregulationsmodell von Zimmerman (2000), das das Lernen in drei Phasen unterteilt, innerhalb derer spezifische Strategien postuliert werden (Planungsphase: Zielsetzung, Strategieplanung, Selbstmotivation, Handlungsphase: Volition und Konzentration, Lernstrategien; Reflexionsphase: Selbstreflexion).

Aufbauend auf dem Modell wurde eine Intervention zur Förderung selbstregulierten Lernens in der Grundschule für Lehrkräfte entwickelt und sowohl auf Ebene der Lehrkräfte als auch auf Ebene der Schülerinnen und Schüler längsschnittlich evaluiert.

Insgesamt nahmen 59 Lehrkräfte und 938 Schülerinnen und Schüler (340 aus Klasse 2, 298 aus Klasse 3 und 300 aus Klasse 4) aus sechs hessischen Grundschulen an der Evaluation teil. Sowohl die Lehrkräfte als auch die Schülerinnen und Schüler bearbeiteten vor und nach der Intervention einen Selbstregulationsfragebogen, der entsprechend des Modells nach Zimmerman konstruiert wurde. Der Fragebogen der Schülerinnen und Schüler enthielt die reliablen Skalen Zielsetzung, Strategieplanung, Selbstmotivation, Konzentration und Selbstreflexion (Cronbachs  $\alpha$  zwischen .60 und .86); der Fragebogen der Lehrkräfte die Skalen Zielsetzung, Strategieplanung, Selbstmotivation, Konzentration, Selbstreflexion und Bewusstsein über Modellfunktion (Cronbachs zwischen .65 und  $\alpha = .90$ ). Die Schülerinnen und Schüler bearbeiteten zusätzlich standardisierte Leistungstests für Deutsch (ELFE 1-6; Lenhard & Schneider, 2006) und Mathematik (DIRG; Grube, Weberschock, Blum, & Hasselhorn, 2010). Aufgrund der Normierung der Tests war so ein Vergleich über die Jahrgangsstufen hinweg möglich.

Die Intervention bestand aus drei Lehrerworkshops, die in Anlehnung an das Selbstregulationsmodell von Zimmerman (2000) entwickelt wurden und zudem Methoden zur Förderung selbstregulierten Lernens im Grundschulunterricht thematisierten. Die im Workshop verwendeten Materialien waren Adaptationen des Arbeitsmaterials, das in der Ausgangsstudie von Leidinger und Perels (2012) entwickelt und erfolgreich evaluiert wurden. Sie fokussierten die folgenden Strategien: Zielsetzung, Strategieplanung, Selbstmotivation, Konzentration und Selbstreflexion.

Zur Evaluation des Projektes wurde geprüft, ob sich innerhalb jeder Jahrgangsstufe signifikante Veränderungen vom Vortest zum Nachtest ergaben. Um Aussagen zum Zusammenhang von Selbstregulation und Leistung treffen zu können, wurden zusätzlich Korrelationsanalysen durchgeführt.

Die Ergebnisse des längsschnittlichen Vergleichs zeigen, dass sich die Schülerinnen und Schüler aller Klassenstufen sowie die Lehrkräfte im Bereich Zielsetzung vom Vor- zum Nachtest verbessert haben (Klasse 4:  $t = -2,689$ ,  $df = 299$ ,  $p = ,008$ ). Zusätzlich zeigte sich ein Anstieg der Schülerinnen und Schüler der Klasse 2 bei der Skala Strategieplanung, in Klasse 3 bei der Skala Selbstreflexion und in Klasse 4 bei der Skala Konzentration. Bezogen auf die Leistungsveränderung der Schülerinnen und Schüler zeigt sich ein uneinheitliches Bild. Bei den Lehrkräften zeigte sich der deutlichste Anstieg bei der Skala „Bewusstsein über Modellfunktion“ ( $t = -1,768$ ,  $df = 47$ ,  $p = ,084$ ). Außerdem verbesserten sie sich in den Skalen Strategieplanung, Selbstmotivation und Selbstreflexion.

Die Korrelationsanalysen zeigten, dass im Jahrgang 2 die Selbstregulationsskalen Zielsetzung und Konzentration mit der Deutschleistung zusammenhängen. Ein ähnliches Ergebnis zeigte sich im Jahrgang 3. Im Jahrgang 4 hingegen zeigen sich keine signifikanten Zusammenhänge im Fach Deutsch.

Im Fach Mathematik ergeben die Analysen im Jahrgang 2 einen Zusammenhang des DIRG mit den Skalen Motivation und Konzentration. In Jahrgang 3 zeigt sich hingegen ein Zusammenhang mit den Skalen Selbstreflexion und Strategieplanung. Zusammenhänge in Jahrgang 4 zeigten sich mit den Skalen Motivation und Konzentration. Die Ergebnisse werden vor dem Hintergrund der Interventionsforschung zum selbstregulierten Lernen diskutiert, um daraus Konsequenzen für die weitere Forschung und Praxis abzuleiten.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), Handbook of selfregulation (pp.13-39). San Diego, CA: Academic Press.

GEBF TAGUNG  
2014

A06 a

Symposiums  
beitrag

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G165

2

Perspektiven der Förderung des selbstregulierten Lernens

## ***Selbstreguliertes Lernen als Prozess:***

### ***Zeitreihenanalytische Effekte des Projekts „Lernen mit Plan“ in der 5. Klasse der Hauptschule***

Der Fähigkeit zum selbstregulierten Lernen kommt im Laufe des Lebens eine immer größere Bedeutung zu. Gerade während und nach der Schulzeit bestimmen Fähigkeit und Bereitschaft, sich neues Wissen anzueignen über selbstständiges Arbeiten und beruflichen Erfolg (Brunstein & Spörer, 2001). Metaanalysen konnten zeigen, dass besonders Aspekte des selbstregulierten Lernens wie das Setzen von Zielen und die Reflexion des Lernprozesses sowie die Selbstwirksamkeitserwartung großen Einfluss auf gute Leistungen haben (Richardson, Abraham, & Bond, 2012). Kern des selbstregulierten Lernens (SRL) ist ein systematisches Herangehen an die Lernhandlung unter Anwendung metakognitiver, motivationaler und kognitiver Lernstrategien zur Erhöhung der Effektivität des Handelns wie auch der eigenen Wahrnehmung ebendieser (Zimmerman, Bonner, & Kovach, 1996).

Im internationalen Vergleich sieht man jedoch, dass deutsche Schüler, insbesondere leistungsschwache Schüler, Schwierigkeiten aufweisen, selbstreguliert zu lernen. (Baumert et al., 2003). Gerade die Hausaufgabenbearbeitung stellt hier eine große Herausforderung an die Selbstständigkeit dar und kann überdies auch emotionale Hürden bergen (Trautwein & Köller, 2003). Ziel der Interventionsstudie „Lernen mit Plan“ war es deshalb, das SRL gezielt bei leistungsschwachen Schülern für die Anwendung bei den Hausaufgaben zu fördern und die Entwicklung des Strategieeinsatzes prozessual über ein Lerntagebuch zu betrachten.

Insgesamt nahmen 75 Lehrer der Haupt- und Werkrealschulen in Baden-Württemberg mit ihren insgesamt 1.408 Schülern der 5. Klasse an der Studie teil. 47 dieser Lehrer (801 Schüler) wurden geschult, über einen Zeitraum von sechs Wochen das SRL im in Kombination mit Mathematischem Modellieren (Maaß, 2007) zu fördern. 28 Lehrer (527 Schüler) fungierten mit ihren Klassen als Wartekontrollgruppe. Die Hälfte der Schüler war männlich. Das mittlere Alter der Schüler betrug elf Jahre.

Das Lerntagebuch wurde ausschließlich von der Experimentalgruppe ausgefüllt. Die Schüler wurden angehalten, es an 42 Tagen während der sechswöchigen Intervention auszufüllen. Das selbstregulierte Lernverhalten war im Lerntagebuch an Otto (2007) angelehnt und wurde mit acht Items erfasst, die auf einer vierstufigen Likert-Skala zu bewerten waren. Die Items deckten die Bereiche Selbstwirksamkeit, Motivation, Planung, Aufschiebeverhalten, Anstrengung, Konzentration, Self-Monitoring und Reflexion ab. Der Mittelwert dieser acht Items stellt den SRL-Gesamtscore dar. Für die Tagebuchbearbeitung an mind. 28 Tagen über den Interventionszeitraum erhielten die Schüler für jeden ausgefüllten Tag 0,50€ (max. 20 €).

Insgesamt füllten 379 Schüler aus 36 Klassen das Lerntagebuch in einer für weitere Analysen angemessenen Qualität aus (Rücklaufquote von 43%). Erste Trendanalysen zeigen einen positiven quadratischen Trend für den SRL-Gesamtscore, der in der Mitte des Interventionszeitraums einen Tiefpunkt aufweist. Weiterhin sind solche positiven quadratischen Verläufe für Motivation, Planung, Anstrengung, Self-Monitoring und Reflexion beobachtbar. Ein positiver linearer Verlauf zeigt sich für das Aufschiebeverhalten, was in diesem Zusammenhang jedoch eine Häufung des Aufschiebens darstellt. Für Selbstwirksamkeit und Konzentration können keine signifikanten Trends nachgewiesen werden. Neben Trendanalysen erfolgt eine Zeitreihenanalyse nach Schmitz, Klug, & Schmidt (2011). Die Ergebnisse werden vor dem Hintergrund der Interventionsforschung zum SRL diskutiert. Es werden Konsequenzen für die Forschung und die Praxis abgeleitet.

Brunstein, J.C. & Spörer, N. (2001). Selbstgesteuertes Lernen. In D.H. Rost (Hrsg.) Handwörterbuch Pädagogische Psychologie (622-629). Weinheim: Beltz.

Maaß, K. (2007). Mathematisches Modellieren. Berlin: Cornelson Scriptor.

Otto, B. (2007). SELVES – Schüler-, Eltern- und Lehrertrainings zur Vermittlung effektiver Selbstregulation. Dissertation. Berlin: Logos.

Richardson, M, Abraham, C. & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological Bulletin, 138 (2), 353–387.

Schmitz, B., Klug, J. & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B. Zimmerman & D. Schunk (Eds.), Handbook of Self-Regulation of Learning and Performance (pp. 251-266). New York, NY: Routledge.

Schmitz, B. & Wiese, B. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. Contemporary Educational Psychology, 31, 64-96.

Trautwein, U. & Köller, O. (2003). The Relationship Between Homework and Achievement - Still Much of a Mystery. Educational Psychology Review, 15, 115–145.

Zimmerman, B.J., Bonner, S. & Kovach, R. (1996). Developing Self-Regulated Learners. Beyond Achievement to Self-Efficacy. Washington (DC): American Psychological Association.



## ***Selbstreguliertes Lernen bei Studierenden fördern: Ergebnisse eines Experiments mit elektronischen Lerntagebüchern***

Nach gängigen Modellen des selbstregulierten Lernens (SRL) ist Lernen definiert als eine Reihe selbstbezogener Prozesse und darauf bezogener Selbsteinschätzungen, durch die Lernen in spezifischen Situationen initiiert, reguliert und aufrechterhalten wird (Zimmerman, 2000). Diese Prozesse umfassen kognitive, metakognitive und motivationale Aspekte gleichermaßen. Dabei ist SRL als bedeutsame Schlüsselkompetenz für die unterschiedlichsten Bildungskontexte von Bedeutung (EU Council, 2002), besonders auch für den universitären Bereich, in welchem Lernende einen hohen selbstständigen Lernanteil haben. Es ist jedoch bekannt, dass Lerner nicht in allen Fällen effektiv mit dieser Autonomie umgehen (vgl. z.B. Sageder, 1994).

Obwohl vielfältig in Studienprogramme eingebunden, besteht häufig ein Defizit zwischen der Bedeutung, die selbstregulatorischen Kompetenzen in der Universität zugemessen wird und der stattfindenden Förderung (Wild, 2000). Die gezielte Förderung von SRL im Kontext von regulären Veranstaltungen bedarf dabei Konzepten, die nicht nur extracurricular verankert sind um einen Ertrag für das Lernen direkt erfahrbar zu machen. Ein vielversprechender Ansatz ist ein Zugang über die Fähigkeit zur Selbstüberwachung, dem im selbstregulierten Lernen eine entscheidende Rolle zukommt. Durch den reflexiven Charakter kann diese Fähigkeit besonders gut durch den Einsatz von Lerntagebüchern (z.B. Klug, Ogrin, Keller, Ihringer, & Schmitz, 2011) gefördert und mit Aspekten zur Lernplanung kombiniert werden. Darüber hinaus liefern Lerntagebücher auch einen Zugang zu Prozessdaten studentischen Lernens. Die hier vorgestellte Studie untersucht den selbstregulationsfördernden Effekt eines standardisierten Lerntagebuchs in elektronischer Form („Studiplaner“), das zusammen mit wöchentlichen Kurzhinweisen zum Einsatz von Strategien im akademischen Lernen eingesetzt wurde.

In einem Prä-Post Kontrollgruppendesign bearbeiteten Lehramtsstudierende im Rahmen einer Veranstaltung im Fach Pädagogische Psychologie (N = 65) ein Lerntagebuch über den Verlauf eines Semesters. Die Akzeptanz des Instruments und der selbstberichtete Nutzen von Seiten der Studierenden können als positiv gewertet werden. In der Prä-Post Befragung wurden Selbstberichtsinventare zu Strategieeinsatz, metakognitiver Kompetenz, volitionaler Kompetenz und Selbstwirksamkeit eingesetzt. Darüber hinaus wurden die Lerntagebuchdaten prozessanalytisch ausgewertet. Die Analyse der Prozessdaten ergab einen durchgehenden bedeutsamen Anstieg von lernzeitbezogenen Variablen über den Verlauf des Semesters. Aus der Prä-Post Befragung ergaben sich bedeutsame Ergebnisse zugunsten der Experimentalgruppe für einzelne Variablen, wie zum Beispiel der selbstberichteten metakognitiven Erfahrung, ein durchgehender Vorteil konnte nicht gefunden werden. Trotz der insgesamt positiven Einschätzung des Instruments durch die Studierenden liegen korrumpierende motivationale und volitionale Effekte nahe, die zur Erklärung der nicht durchgehend positiven Ergebnisse heran gezogen werden können (Zimmerman & Schunk, 2008). Insgesamt scheint es notwendig, den Einsatz von Lerntagebüchern noch stärker in Lehrveranstaltungen einzuweben, um einen größtmöglichen Nutzen wahrscheinlich zu machen.

EU Council (July 2002). Council resolution of 27 June 2002 on lifelong learning. Official Journal of the European Communities, 9.

Klug, J., Ogrin, S., Keller, S., Ihringer, A., & Schmitz, B. (2011). A plea for self-regulated learning as a process: Modelling, measuring and intervening. *Psychological Test and Assessment Modeling*, 53, 51–72.

Sageder, J. (1994). Lernmotivation, Attributionstendenzen und Lernmethoden von Studienanfängern, *Psychologie in Erziehung und Unterricht*, 41, 120–133.

Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster- Waxmann.

Zimmerman, B. J. (2000). Attaining Self-Regulation. A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeider (Eds.), *Handbook of Self-Regulation* (pp. 13–39). San Diego: Academic Press.

Zimmerman, B. J. & Schunk, D. (2008). Motivation, an essential dimension of self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research and applications* (pp. 1–30). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.





## **Effekte eines web-basierten Trainings zum selbstregulierten Lernen auf objektive Verhaltensmaße: Logfile-Analysen**

Selbstreguliertes Lernen (SRL) ist eine Kompetenz, die insbesondere im Studium von hoher Relevanz ist, in der Studierende aber dennoch häufig Defizite aufweisen. Trainingsmaßnahmen sind effektiv in der Förderung der SRL-Kompetenz (Schmitz & Wiese, 2006), erfordern aber einen hohen personellen und zeitlichen Aufwand. Web-basierte Trainings (WBT) können diesen Aufwand reduzieren, indem sie über das Internet distribuiert und von den Nutzern zeitlich flexibel eingesetzt werden können. Ein weiterer Vorteil von computerbasierten Lernumgebungen ist die Dokumentation des Lernverhaltens in Form von Logfiles. Diese bieten die Möglichkeit, die Evaluation von Trainingsmaßnahmen neben subjektiven Selbstreportdaten auch objektiv zu überprüfen. Bisherige Forschung zu Logfiledaten im Forschungsgebiet des selbstregulierten Lernens verwendete allerdings speziell konstruierte Lernumgebungen, z.B. gStudy (Hadwin, Nesbit, Jamieson Noel, Code, & Winne, 2007), in denen selbstreguliertes Lernen durch eigens programmierte Funktionen zur Annotation oder zur Verlinkung von Inhalten dokumentiert wurde. In solchen zu Forschungszwecken erstellten, „künstlichen“ Lernumgebungen wird beispielsweise die Häufigkeit oder Qualität der Nutzung der angebotenen Funktionen als Maß für selbstreguliertes Lernen verwendet. Weitgehend unerforscht sind hingegen Logfiles aus „natürlichen“ Lernumgebungen wie z.B. Moodle, die in realen Lernszenarien eingesetzt werden. Hier stellt sich die Frage, welche Parameter aus den Logfiles extrahiert werden können, um Rückschlüsse auf das selbstregulierte Lernen der Probanden zu ermöglichen. In unserem Projekt wurde ein WBT entwickelt, das die SRL-Kompetenz angehender Studierender fördern soll. Das Training basiert auf dem Prozessmodell des Lernens nach Schmitz (2001) und vermittelt in drei Lektionen insbesondere metakognitive, aber auch kognitive, volitionale und motivationale Strategien. In einer Evaluationsstudie wurden N=175 angehende Studierende untersucht, die in Vorbereitung auf ihr Studium den Mathematikvorkurs der Technischen Universität Darmstadt besuchen. Dieser vierwöchige Kurs findet ohne Präsenzanteile vollständig web-basiert statt und umfasst in mehr als 50 Kapiteln die mathematischen Schulkenntnisse, die von Studienanfängern in Informatik, Maschinenbau, Bauingenieurwesen und Mathematik erwartet werden. Nach einer randomisierten Einteilung in vier Gruppen besuchte die Kontrollgruppe den Kurs ohne weitere Unterstützung. Die Tagebuchgruppe führte ein standardisiertes Lerntagebuch; die Trainingsgruppe erhielt zusätzlich zum Lerntagebuch Zugang zum WBT. Die Coachinggruppe erhielt zusätzlich zum Lerntagebuch und zum Training darüberhinaus noch Instruktionen für ein Peer-Coaching. In einem Prä-Post-Design wurde ein SRL-Fragebogen sowie ein Mathematiktest eingesetzt. Die Wirksamkeit der drei Interventionen zur Förderung des selbstregulierten Lernens (Tagebuch, Training und Peer-Coaching) konnte anhand von Selbstberichtsdaten (Fragebogen und Tagebuch) und Leistungstests (Mathematiktest) nachgewiesen werden (Bellhäuser, Lösch, Winter & Schmitz, submitted). Im vorliegenden Beitrag sollen nun Logfile-Analysen diese Ergebnisse ergänzen. Erste Analysen konnten eine signifikant erhöhte Time-on-Task (Dauer der Nutzung der Lernplattform) der unterstützten Gruppen gegenüber der Kontrollgruppe zeigen ( $\chi^2$  df=3= 12.255;  $p = .007$ ). So verbrachte die Kontrollgruppe im Durchschnitt M=16:04:50 (SD= 13:10:25) Stunden auf der Lernplattform, die Trainingsgruppe hingegen im Durchschnitt M=21:47:35 (SD= 14:58:14) Stunden. In weiteren Analysen soll u.a. die Regelmäßigkeit der Nutzung der Lernplattform untersucht werden, z.B. ob in allen vier Wochen des Mathematikurses gleich viel Zeit auf das Lernen verwendet wurde oder ob das Lernen aufgeschoben wurde und daher insbesondere in der letzten Woche stattfand. Ebenso sollen die individuellen Lernzeiten analysiert werden, da beispielsweise eine Präferenz für späte Uhrzeiten oft mit Prokrastination assoziiert ist (Digdon & Howell, 2008). Darüber hinaus soll das Erreichen selbstdefinierter Ziele im Lerntagebuch objektiv anhand der Logfiles überprüft werden.

Bellhäuser, H., Lösch, T., Winter, C., & Schmitz, B. (submitted). Applying a Web-Based Training to Foster Self-Regulated Learning.

Digdon, N. L., & Howell, A. J. (2008). College students who have an eveningness preference report lower self-control and greater procrastination. *Chronobiology International*, Vol.25(6), 1029–1046.

Hadwin, A. F., Nesbit, J. C., Jamieson Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, Vol.2(2-3), 107–124.

Schmitz, B., & Wiese, B. S. (2006). New Perspectives for the Evaluation of Training Sessions in Self-Regulated Learning: Time-Series Analyses of Diary Data. *Contemporary Educational Psychology*, v31 n1 p64, 33.



Chair(s): Katja Hartig (Goethe-Universität Frankfurt), Johannes Naumann (DIPF)

Diskutant/in: Nina Jude (DIPF)

Dieses Symposium bringt Beiträge aus vier Projekten zusammen, die vorlaufende, begleitende oder nachlaufende Forschung zu Large-Scale-Assessments (LSAs) betreiben und im Rahmen der Initiative des Bundesministeriums für Bildung und Forschung (BMBF) zur Förderung von Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments gefördert werden. Die Beiträge repräsentieren dabei eine große Bandbreite von in LSAs erfassten Kompetenzkonstrukten (Lesen, Mathematik, komplexes Problemlösen, ICT-Skills) und methodischer Zugänge. Die integrierende Klammer für alle Beiträge ist dabei die Frage, wie Leistungen von Schüler/innen bei der Bearbeitung von Aufgaben in LSAs im Einzelnen zustande kommen, und wie die Validität von LSAs gesichert werden kann. In diesem Sinne wird in Beitrag 1 (Hartig et al.) die Entwicklung eines Tests zur Messung von ICT-Skills vorgestellt. Die Gültigkeit wird zunächst durch theoriegeleitete Aufgabenkonstruktion und systematische Variation von Aufgabenmerkmalen anhand eines Frameworks sichergestellt, das bisherige Frameworks für ICT-Skills integriert und weiterentwickelt. Darüber hinaus sollen Prozessanalysen aus Cognitive Laboratory (CogLab) Untersuchungen Aufschluss geben über die Validität der Testaufgaben. Komplementär hierzu wird in Beitrag 2 (Naumann et al.) anhand der Domäne „Lesen Digitaler Texte“ gezeigt, dass ein Prozessmaß – die aufgewendete Bearbeitungszeit – über das Fähigkeitsniveau der Personen und das Schwierigkeitsniveau der Aufgaben hinaus die erfolgreiche Bearbeitung von LSA-Aufgaben erklärt, wobei der Zusammenhang theoriekonform durch Merkmale der Aufgaben und der Testpersonen moderiert wird. Der Beitrag ist damit auch ein Beispiel für die Verwendung von Daten aus LSAs zur Klärung grundlagenwissenschaftlicher Fragen, die wiederum relevant für die Beschreibung von in LSAs gemessenen Konstrukten sind. Beitrag 3 (Wüstenberg et al.) zeigt exemplarisch, wie ein neu in LSAs eingeführtes Konstrukt – Komplexes Problemlösen – in ein nomologisches Netzwerk so eingepasst wird, dass man von einer gültigen Messung ausgehen und ggf. praktische und bildungspolitische Folgerungen ableiten kann. Während Beitrag 3 den Zusammenhang zwischen einem Konstrukt (Komplexes Problemlösen) und relevanten Prädiktoren und Kriterien in den Blick nimmt, analysiert Beitrag 4 (van den Ham et al.) die Äquivalenz von unterschiedlichen Messungen desselben Konstrukts. Dabei wird gezeigt, dass die Messungen mathematischer Kompetenz über unterschiedliche LSAs hinweg trotz divergenter theoretischer Grundlagen und Frameworks zu vergleichbaren Ergebnissen führt. Insgesamt belegen die Beiträge des Symposiums damit die Fruchtbarkeit begleitender Forschung zu Large-Scale-Assessments für die Weiterentwicklung des Feldes sowohl durch vorlaufende Testentwicklung, Validierung und Grundlagenforschung, die ihrerseits an künftige Assessments rückgebunden werden kann.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G131

## Zur Erfassung des Konstrukts ICT-Literacy

Der kompetente Umgang mit Informations- und Kommunikationstechnologie (ICT-Literacy) ist für eine erfolgreiche Teilhabe in modernen Wissensgesellschaften eine zentrale Voraussetzung (ETS, 2002). Bisher wurden computerbezogene Kompetenzen in Large-Scale-Assessments wie PISA anhand von subjektiven Selbsteinschätzungen der Kompetenz im Umgang mit Computeranwendungen erfasst. Vor dem Hintergrund einer valideren und effizienteren Messung von ICT-Skills hat diese Arbeit das Ziel, ein Messinstrument zur computerbasierten verhaltensnahen Erfassung von ICT-Skills zu entwickeln.

Der vorliegende Beitrag setzt damit an, das Konstrukt ICT-Skills mithilfe einer geeigneten Definition und Beschreibung der internen Struktur theoretisch zu fundieren. Dabei wird ICT-Skills wie folgt definiert: „ICT-Skills sind Wissen und Fertigkeiten zur Nutzung von Informations- und Kommunikationstechnologien, die für die erfolgreiche Teilhabe in modernen Wissensgesellschaften relevant sind. ICT-Skills umfassen die Rezeption, Anwendung und Produktion multimedialer Information zur erfolgreichen Bearbeitung informationsbezogener Aufgaben.“ Die Beschreibung der Struktur wurde entwickelt unter Hinzunahme verschiedener Theorien und Modelle (z.B. Ferrari, 2012; van Deursen & van Dijk, 2009; International ICT Literacy Panel, 2002) und umfasst vier verschiedene Dimensionen. Die Dimension kognitive Operationen ist in Anlehnung an das ETS-Framework formuliert und umfasst die Facetten access, manage, integrate, evaluate und create. Eine weitere Dimension beschreibt die Situationen, in denen ICT-Skills zur Anwendung kommen können (Senkbeil & Wittwer, 2007; Richter, Naumann, & Horz, 2010). Dabei werden persönliche, berufliche und bildungsbezogene Situationen unterschieden. Im Hinblick auf die Dimension soziale Interaktionen wird zwischen individuell und kollektiv differenziert. Anders als in früheren Konzeptionen (z. B. ETS, 2003; Eshet-Alkalai, 2004) ist die Unterscheidung zwischen der individuellen und kollektiven Bearbeitung von ICT-Aufgaben orthogonal zu den beanspruchten kognitiven Operationen. Schließlich wird zwischen der Verarbeitung von Inhalten in Bezug auf Modalität und Kodalität unterschieden, wobei wir auf das Modell der integrativen Text- und Bildverarbeitung von Schnotz (2005) zurückgreifen.

Auf Grundlage des theoretischen Frameworks wurde ein Assessmentframework entwickelt aus dem abgeleitet wird, welche Art von Aufgaben die konstruktbestimmenden Fertigkeiten und Fähigkeiten repräsentieren. Es können so für die Bearbeitung unterschiedlicher Aufgaben bestimmte notwendige Prozesse abgeleitet werden. Aus den Dimensionen und Kombinationen dieser Dimensionen ergeben sich Aufgabenklassen, die sich im Hinblick auf die Fähigkeiten und Prozesse unterscheiden, die für die Lösung der Aufgaben benötigt werden. Dabei wurden aus Gründen der Praktikabilität nur drei der im theoretischen Framework beschriebenen Dimensionen aufgenommen, auf die Unterscheidung zwischen Modalitäten wurde verzichtet. Bei dieser Art der systematischen theoriegeleiteten Aufgabenkonstruktion findet für jedes Item eine Einordnung im Hinblick auf die Dimensionen statt sowie eine Beschreibung kritischer Aufgabenaspekte und schwierigkeitsbestimmender Merkmale, wie z. B. der Anzahl und Güte von Distraktoren, Anzahl der Teilhandlungen zur Erreichung der Aufgabenlösung und verfügbare Funktionalitäten innerhalb der simulierten Oberfläche.

Im vorliegenden Beitrag wird der Aufgabenentwicklungsprozess exemplarisch dargestellt, von der Entwicklung eines Items über die Implementierung mit einer speziellen Software bis hin zur qualitativen Analyse im Cognitive Laboratory. Die systematische Prüfung von prototypischen Aufgaben im CogLab ist für November 2013 geplant. Hierzu werden ca. 15-20 Schüler/innen der 9. Klasse aus allen Schulformen (Hauptschule, Realschule, Gymnasium) untersucht. Die Testungen finden als Einzeluntersuchungen statt. Hierzu soll das Verhalten beim Bearbeiten der Aufgaben mit Videoaufzeichnungen erfasst werden, zusätzlich sollen die Schüler/innen ihr Bearbeitungsverhalten im Nachhinein anhand der Videoaufzeichnungen erläutern. Diese retrospektiven verbalen Protokolle sollen Aufschluss darüber geben, welche kognitiven Prozesse während der Bearbeitung der Aufgaben stattfinden. Neben qualitativen Variablen sollen auch die Bearbeitungszeit und die Performanz in der Aufgabe erfasst werden. Außerdem kommt in den Untersuchungen ein eigens entwickelter Hintergrundfragebogen zum Einsatz, der sowohl demografische Daten erfasst als auch Informationen zu Erfahrungen und Nutzungsverhalten im Hinblick auf Informations- und Kommunikationstechnologien. Die quantitative Überprüfung des theoretischen Frameworks im Large-Scale-Assessment ist für Frühjahr/Sommer 2014 geplant, Zielstichprobe sind dabei 15-jährige Schüler/innen.

Eshet-Alkalai, Y. (2004). Digital literacy: A framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, 13, 93-106.

ETS (Educational Testing Service; 2002). Digital transformation. A framework for ICT literacy. Princeton, NJ: ETS.

ETS (Educational Testing Service, 2003). Succeeding in the 21st Century: What higher education must do to address the gap in information and communication technology proficiencies. Retrieved from [http://www.ets.org/Media/Tests/Information\\_and\\_Communication\\_Technology\\_Literacy/ICTwhitepaperfinal.pdf](http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ICTwhitepaperfinal.pdf)

Ferrari, A. (2012). Digital competence in practice: An analysis of frameworks. Luxembourg: Publications of the European Union.

International ICT Literacy Panel (2002). Digital transformation: A framework for ICT literacy (A report of the International ICT Literacy Panel). Princeton, NJ: Educational Testing Service.

Heruntergeladen am 21. Februar 2013, [http://www.ets.org/research/policy\\_research\\_reports/publications/report/2002/cjik](http://www.ets.org/research/policy_research_reports/publications/report/2002/cjik)

Richter, T., Naumann, J. & Horz, H. (2010). Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R). *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology*, 24, 23-37.

Schnotz, W. (2005). An Integrated Model of Text and Picture Comprehension. In R.E. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (pp. 49-69). Cambridge: Cambridge University Press.

Senkbeil, M. & Wittwer, J. (2007). Die Computervertrautheit von Jugendlichen und Wirkungen der Computernutzung auf den fachlichen Kompetenzerwerb. In PISA-Konsortium Deutschland (Hrsg.),

PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie (S. 277-308). Münster: Waxmann.

Van Deursen, A. J., & Van Dijk, J. A. (2009). Using the Internet: Skill related problems in users' online behavior. *Interacting with Computers*, 21(5), 393-402.



## Die Beziehung zwischen Bearbeitungszeit und Bearbeitungsergebnis bei Digitalen Leseaufgaben

Bei kompetenten Leser/innen beruht Lesen auf einer Kombination automatischer und kontrollierter Prozesse. Prozesse wie Worterkennung, syntaktisches Parsing und semantische Integration sind bei kompetenten Lesern automatisiert (Gernsbacher & Faust, 1991; Perfetti, 1985; 1994). Die Geschwindigkeit dieser Prozesse hat enge Zusammenhänge mit dem Leseverständnis auf Textebene (Naumann, Richter, Flender, Christmann, & Groeben, 2007; Richter, Isberner, Naumann, & Neeb, 2012; 2013). Gleichzeitig erfordern komplexe Leseaufgaben zeitintensive Regulationsaktivitäten. Entsprechend finden sich positive Zusammenhänge zwischen der Zeit, die mit relevanten oder komplexen Teilen eines Textes verbracht wird und der Qualität des Bearbeitungsergebnisses (z. B. Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012). Vor diesem Hintergrund ist zu vermuten, dass die Beziehung zwischen der Bearbeitungszeit und dem Bearbeitungsergebnis durch kognitive Anforderungen und die Aufgabenschwierigkeit moderiert wird. Bei leichten Aufgaben, die sich durch Routineoperationen lösen lassen, sollte sich ein negativer Zusammenhang zwischen Bearbeitungszeit und Bearbeitungsergebnis finden. Bei schwierigen Aufgaben, die den Einsatz kognitiver Strategien erfordern, sollte sich ein positiver Zusammenhang finden. Diesen Überlegungen folgend sollte der Zusammenhang zusätzlich von Merkmalen der Testperson moderiert werden. Insbesondere bei weniger kompetenten Leser/innen ist zu erwarten, dass sie von zusätzlich in die Aufgabenbearbeitung investierter Zeit profitieren.

Methode: Datengrundlage bildet das PISA 2009 Digital Reading Assessment, bei dem in 19 Bildungssystemen computergestützt die Fähigkeit von 15-Jährigen beim Verständnis digitaler Texte geprüft wurde. Die Aufgaben unterscheiden sich in Bezug auf Navigationskomplexität und Textkomplexität. Zur Prüfung der Hypothesen wurden generalisierte gemischte lineare Modelle geschätzt. In Modell 1 wurden ein Zufallseffekt für die Testperson ( $\sim$ Fähigkeit), ein Zufallseffekt für die Items ( $\sim$ Leichtigkeit), ein fester Effekt der Bearbeitungszeit, sowie ein Zufallseffekt der Bearbeitungszeit über Items und über Personen geschätzt. Die Moderation des Effekts der Bearbeitungszeit durch die Itemschwierigkeit findet sich in diesem Modell als Kovarianz zwischen dem Zufallseffekt für Items und dem Zufallseffekt der Bearbeitungszeit über Items. Die Moderation des Effekts der Bearbeitungszeit über Personen findet sich als Kovarianz zwischen dem Zufallseffekt für Personen und dem Zufallseffekt der Bearbeitungszeit über Personen. In Modell 2 wurden zusätzlich als schwierigkeitsbestimmende Itemmerkmale Navigationskomplexität und Textkomplexität, sowie die Interaktion dieser Variablen mit der Bearbeitungszeit modelliert. Auf Personenseite wurde als Moderator der Beziehung zwischen Bearbeitungszeit und Bearbeitungsergebnis die Lesefähigkeit der Testpersonen im Printmedium angesetzt. Alle Analysen wurden getrennt nach Bildungssystemen durchgeführt.

Ergebnisse: Modell 1: Die Parameterschätzungen stützen die Hypothesen. Für 18 von 19 Bildungssystemen fand sich eine signifikante negative Kovarianz zwischen dem Zufallseffekt für die Items ( $\sim$ Leichtigkeit) und dem Zufallseffekt der Bearbeitungszeit über die Items ( $-.68 < r < -.42$ ). Inhaltlich bedeutet dies, dass für schwere Items positive und für leichte Items negative Effekte der Bearbeitungszeit resultierten. Diese Ergebnisse stützen die Annahme, dass der Bearbeitungszeiteffekt bei leichten Leseaufgaben den Grad der Automatisierung relevanter Prozesse widerspiegelt, bei schweren Leseaufgaben dagegen die Allokation kognitiver Ressourcen. Gestützt wird diese Interpretation durch eine negative Kovarianz zwischen dem Zufallseffekt über Personen ( $\sim$ Fähigkeit) und dem Zufallseffekt der Bearbeitungszeit über Personen ( $-.90 < r < -.19$ ). Diese Kovarianz bedeutet, dass insbesondere bei schwächeren Personen positive Effekte der aufgewendeten Zeit resultieren.

Modell 2: Die Parameterschätzungen stützen die Hypothesen weitgehend. Wie erwartet finden sich Interaktionen zwischen der Anzahl benötigter Navigationsschritte und der Bearbeitungszeit ( $0.45 < \beta < 0.23$ ). Bei navigationsintensiven Aufgaben resultiert ein positiver Effekt der Bearbeitungszeit, der bei abnehmender Navigationsintensität verschwindet. Entgegen den Annahmen wird der Effekt der Bearbeitungszeit jedoch nicht durch die Textkomplexität moderiert ( $-0.05 < \beta < 0.12$ ). Konsistent mit den Hypothesen ergibt sich eine Interaktion zwischen Bearbeitungszeit und Print-Lesefähigkeit ( $-1.23 < \beta < -0.89$ ). Im Printmedium schwächere Leser/innen profitieren demnach bei der Bearbeitung digitaler Leseaufgaben von zusätzlich aufgewendeter Bearbeitungszeit.

Die Ergebnisse zeigen, dass die für Leseaufgaben aufgewendete Zeit keine uniforme Beziehung zum Bearbeitungsergebnis aufweist. Insbesondere bei schwachen Testpersonen und schwierigen Aufgaben hängt die Bearbeitungszeit positiv mit dem Bearbeitungsergebnis zusammen. Implikationen dieses Befundes für die Validität von Lesetests mit selbst gewählter Aufgabenbearbeitungszeit werden diskutiert.

- Gernsbacher, M.A., & Faust, M.E. (1991). Less-skilled comprehenders have less-efficient suppression mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 245-262.
- Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. M. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, 47, 356-381.
- Naumann, J., Richter, T., Flender, J., Christmann, U., & Groeben, N. (2007). Signaling in expository hypertexts compensates for deficits in reading skill. *Journal of Educational Psychology*, 99, 791-807.
- Perfetti, C.A. (1985). *Reading ability*. New York: Academic Press.
- Perfetti, C.A. (1994). Psycholinguistics and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849-894). San Diego, CA: Academic Press.
- Richter, T., Isberner, M.-B., Naumann, J., & Neeb, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern. *Zeitschrift für Pädagogische Psychologie*, 26, 313-313.
- Richter, T., Isberner, M.-B., Naumann, J., & Neeb, Y. (2013). Lexical quality and reading comprehension in primary school children. *Scientific Studies of Reading*. Advance online publication. doi:10.1080/10888438.2013.764879





## **Konstruktvalidierung von Komplexem Problemlösen mit Hilfe von und für Large-Scale Assessments**

In den letzten Jahren ist im Rahmen internationaler Large Scale Assessments (LSA) wie PISA ein verstärktes Fokussieren auf fachübergreifende Kompetenzen zu beobachten. Komplexes Problemlösen (KPL) stellt eine solche breit gefächerte und weitestgehend fachunspezifische Kompetenz dar, deren Erfassung in die PISA 2012 Studie integriert wurde (Leutner, Klieme, Meyer, & Wirth, 2004; OECD, 2013). Eine Voraussetzung für die Interpretation von in LSAs festgestellten Leistungsunterschieden ist jedoch eine adäquate Konstruktvalidierung, die für ein innovatives Konstrukt wie KPL als ausbaufähig zu betrachten ist, insbesondere wenn bildungspolitische Entscheidungen aus den Ergebnissen abgeleitet werden sollten. In diesem Beitrag wurden deshalb die Daten von Sechstklässlern eines in Finnland durchgeführten internetbasierten LSAs verwendet um zu überprüfen, inwiefern KPL durch kognitive Fähigkeiten sowie durch externe Faktoren wie der schulischen Lernumgebung beeinflusst wird und ob KPL inkrementelle Validität in der Vorhersage relevanter Leistungskriterien aufweist. Ein komplexes Problem zeichnet aus, dass zu Beginn des Problemlösungsvorgangs nicht alle Informationen zur Verfügung stehen und sich die Problemsituation dynamisch verändern kann (Funke, 2001). Demzufolge sind die Hauptanforderungen des KPL an den Problemlöser sich mittels interaktiver Auseinandersetzung mit der Problemstellung neues Wissen anzueignen und dieses anzuwenden (Novick & Bassok, 2005). Bisherige empirische Befunde zeigten, dass KPL mit dem Kurzzeitgedächtnis (Schweizer, Wüstenberg, & Greiff, 2012) und fluidem Reasoning (Wüstenberg, Greiff, & Funke, 2012) moderat assoziiert ist. Auch konnte KPL über die genannten kognitiven Fähigkeiten hinaus Leistung in Außenkriterien wie Schulnoten (Wüstenberg et al., 2012) und Berufserfolg (Danner, Hagemann, Schankin, Hager, & Funke, 2011) vorhersagen. Dieser Beitrag geht über bestehende Arbeiten hinaus, indem erstmalig die Beziehung von KPL zu mehreren kognitiven Konstrukten und der Lernumgebung als externen Faktor kombiniert untersucht wurden. In Hypothese 1 wurde erwartet, dass sowohl kognitive Faktoren wie fluides Reasoning und Kurzzeitgedächtnis, als auch externe Faktoren wie die Lernumgebung einen signifikanten Anteil an Varianz in KPL erklären. In Hypothese 2 wurde formuliert, dass KPL inkrementelle Validität in der Vorhersage von Schulnoten und einer Lehrereinschätzung des akademischen Potentials über fluides Reasoning und Kurzzeitgedächtnis hinaus aufweist. Bei der Stichprobe handelte es sich um N=2029 Finnische Sechstklässler. Zur Erfassung komplexer Problemlösefähigkeit wurden 9 MicroDYN-Aufgaben vorgegeben (s. Greiff, 2012). Alle weiteren Konstrukte wurden mit bewährten Standardtests erfasst (z.B. fluides Reasoning mit Bond's Test for Logical Operations; Bond, 1995). Die Lernumgebung wurde mittels Lehrereinschätzungen erfragt und umfasste 5 Items zum Klassenklima. Zur Untersuchung der Hypothesen wurden latente Strukturgleichungsmodelle unter Berücksichtigung der Multi-Level Struktur der Daten gerechnet (d.h. zwei Ebenen: Schüler geschachtelt in Klassen). Die Ergebnisse zu Hypothese 1 zeigten, dass lediglich fluides Reasoning (Ebene 1:  $\beta = .562$ ,  $SE = .113$ ,  $p < .001$ ;  $R^2 = .312$ ) und die Lernumgebung (Ebene 2:  $\beta = .528$ ,  $SE = .119$ ,  $p < .001$ ;  $R^2 = .279$ ) Varianz in KPL erklärten, nicht jedoch das Kurzzeitgedächtnis. In Hypothese 2 war fluides Reasoning der stärkste Prädiktor der Kriterien (Schulnoten:  $\beta = .447$ ,  $SE = .107$ ,  $p < .001$ ;  $R^2 = .200$ ; akademisches Potential:  $\beta = .312$ ,  $SE = .101$ ,  $p < .001$ ;  $R^2 = .097$ ), jedoch erklärte KPL Varianz über fluides Reasoning und Kurzzeitgedächtnis hinaus (Schulnoten:  $\beta = .173$ ,  $SE = .051$ ,  $p < .001$ ;  $\Delta R^2 = .03$ ; akademisches Potential:  $\beta = .111$ ,  $SE = .045$ ,  $p = .014$ ;  $\Delta R^2 = .01$ ). Zusammengefasst lässt sich festhalten, dass neben fluidem Reasoning auch eine förderliche Lernumgebung Auswirkungen auf die Leistung in KPL hat. Ferner zeigte KPL inkrementelle Validität in der Vorhersage von Schulnoten und einer Einschätzung akademischen Potentials über fluides Reasoning und Kurzzeitgedächtnis hinaus. Die Ergebnisse dieser Arbeit sollen aufzeigen, dass eine Überprüfung der Determinanten und Auswirkungen von KPL in LSAs wichtige Informationen über die bloße Berichterstattung von Leistungsunterschieden hinaus liefert, aus welchen wiederum wissenschaftlich begründete Ansatzpunkte für etwaige bildungspolitische Schlussfolgerungen gezogen werden können.

Bond, T. G. (1995). Piaget and measurement II: Empirical validation of the Piagetian model. *Archives de Psychologie*, 63, 155–185.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7, 69–89.

Greiff, S. (2012). Individualdiagnostik der komplexen Problemlösefähigkeit. Münster: Waxmann.

Leutner, D., Klieme, E., Meyer, K., & Wirth, J. (2004). Problemlösen. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H. G. Rolf, J. Rost & U. Schiefele (Eds.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (pp. 147–176). Münster: Waxmann.

Novick, L. R., & Bassok, M. (2005). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (p. 321–349). Cambridge, NY: University Press.

OECD. (2013). *PISA 2012 assessment and analytical framework*. Paris: OECD Publishing.

Schweizer, F., Wüstenberg, S., & Greiff, S. (2013). Do Complex Problem Solving dimensions measure something over and above working memory capacity? *Learning and Individual Differences*, 24, 42–52.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex Problem Solving. More than reasoning? *Intelligence*, 40, 1–14.





## **PISA, NEPS und BiSta – Sind die Kompetenzmessungen in Mathematik vergleichbar?**

Die mathematische Kompetenz von Jugendlichen am Ende der Sekundarschulzeit wird in Deutschland derzeit in mehreren nationalen und internationalen Large Scale Assessments gemessen. Hierzu zählen etwa das Programme for International Student Assessment (PISA), der nationale Ländervergleich der Bildungsstandards (BiSta) und die Erhebungen im Nationalen Bildungspanel (NEPS). Inwieweit sind diese Kompetenzmessungen aber vergleichbar und erfassen das gleiche Konstrukt von mathematischer Kompetenz? Obwohl alle drei Studien mathematische Kompetenz erfassen, lassen sich die Studienergebnisse nicht direkt in Beziehung setzen, da sie auf unterschiedlichen Rahmenkonzeptionen und Berichtsskalen beruhen. So ist PISA eine internationale Schulleistungsstudie, die seit 2000 in einem dreijährlichen Rhythmus von der OECD durchgeführt wird und das Ziel verfolgt, mathematische Kompetenzen im Sinne von Mathematical Literacy zu messen (OECD 2003). Dabei geht es vor allem um die funktionale Anwendung von Mathematik in alltagsnahen Kontexten. Der nationale Ländervergleich untersucht für die 9. Jahrgangsstufe die mathematischen Kompetenzen der Schülerinnen und Schüler mit dem Ziel, festzustellen, inwieweit in den Bundesländern die deutschlandweit verbindlichen Bildungsstandards erreicht werden (Pant, Stanat, Pöhlmann, & Böhme, 2013). Die Definition der Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss weist zwar eine strukturelle Ähnlichkeit zur PISA-Rahmenkonzeption in Mathematik auf, ist aber inhaltlich stärker auf die unterrichtliche Situation in Deutschland angepasst (KMK, 2003). Das nationale Bildungspanel (NEPS) hingegen ist eine Längsschnittstudie, die die Bildungsprozesse und unter anderem die Kompetenzentwicklung in Mathematik über die Lebensspanne analysiert. Die Rahmenkonzeption für die Erhebungen in Mathematik orientiert sich an den Kompetenzstrukturmodellen aus PISA und den Bildungsstandards, überträgt diese Struktur jedoch auch auf andere Alterskohorten, um die gesamte Lebensspanne abzudecken (Blossfeld, von Maurice, & Schneider, 2011). Um die Äquivalenz der Kompetenzmessungen in Mathematik zwischen diesen drei LSA zu überprüfen, wurde eine Validierungsstudie durchgeführt. Diese Studie basiert auf einer Stichprobe von 80 Schulen mit  $N = 1965$  Schülerinnen und Schülern aus der neunten Klassenstufe. Alle Schülerinnen und Schüler haben in einem Rotationsdesign die Mathematiktests vom Ländervergleich, aus NEPS und PISA an zwei Testtagen bearbeitet. Anhand dieses Single-Group-Designs lässt sich die Äquivalenz der drei Kompetenzmessungen analysieren, indem in drei Schritten die konzeptionelle, die dimensionale und die skalenbezogene Vergleichbarkeit der Studien geprüft wird (Kolen & Brennan, 2010). Die Ergebnisse des konzeptionellen Vergleichs zeigen, dass die Kompetenzmessung in NEPS, trotz einiger Abweichungen im Framework und den Itemformaten, konzeptionell eine hohe Ähnlichkeit zu PISA und zum Ländervergleich aufweist. Aus Ergebnissen eines Expertenreviews ergibt sich weiterhin, dass sich alle Aufgaben des NEPS-Mathematiktests in die Testkonzeptionen vom Ländervergleich und von PISA übertragen lassen. Der zweite Schritt der Untersuchung beinhaltet die Analyse der dimensional Äquivalenz. Mit Hilfe von multidimensionalen IRT-Modellen wurden die latenten Korrelationen innerhalb der Subskalen der Testinstrumente und zwischen den Studien berechnet. Dabei sind die latenten Korrelationen mit  $r = .86$  zwischen NEPS und PISA sowie  $r = .90$  zwischen NEPS und Ländervergleich als hoch zu bezeichnen. Insgesamt kann die dimensionale Struktur als durchaus vergleichbar eingeschätzt werden. In Bezug auf die Skalenäquivalenz soll abschließend diskutiert werden, in wie weit sich die Skalenmetriken und Kompetenzstufenmodelle vom Ländervergleich und von PISA auf den NEPS-Mathematiktest übertragen lassen.

- Blossfeld, H., von Maurice, J., & Schneider, T. (2011). 1 the national educational panel study: Need, main features, and research potential das Nationale Bildungspanel: Notwendigkeit, Grundzüge und Analysepotential. Zeitschrift Für Erziehungswissenschaft, 14(2), 5-17. doi:10.1007/s11618-011-0178-3
- Kolen, M. J. & Brennan, R. L. (2010): Test Equating, Scaling, and Linking. Methods and Practices. Springer.
- Kultusministerkonferenz (2003). Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss vom 04.12.2003. Neuwied: Luchterhand.
- OECD (2003). The PISA 2003 Assessment framework – Mathematics, reading, science and problem solving knowledge and skills. Paris: OECD.
- Pant, H. A., Stanat, P., Pöhlmann, C. & Böhme, K (2013): Die Bildungsstandards im allgemeinbildenden Schulsystem, In Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.): IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I, Münster: Waxmann, S. 13-22.

Miriam Lotz (Otto-Friedrich-Universität Bamberg, Universität Kassel) Ann-Katrin Denn (Universität Kassel),  
Frank Lipowsky (Universität Kassel)

## ***Allen das Gleiche oder Jedem das Seine? Zur Verteilung von Hilfestellungen im Leseunterricht des ersten Schuljahres in Abhängigkeit von Lehrer- und Schülermerkmalen***

Chair(s): Miriam Lotz (Otto-Friedrich-Universität Bamberg, Universität Kassel)

Insbesondere im Grundschulunterricht stehen Lehrpersonen vor der Herausforderung trotz heterogener Klassenzusammensetzung jeden Schüler individuell zu fördern (Pauli & Reusser, 2006). Schülerarbeitsphasen ermöglichen dabei grundsätzlich eine dem eigenen Lerntempo entsprechende aktive Auseinandersetzung mit dem Unterrichtsgegenstand und schaffen für die Lehrperson gleichzeitig zahlreiche Gelegenheiten, einzelne Schüler individuell zu unterstützen (vgl. Krammer, 2009). Bislang ist allerdings wenig darüber bekannt, wie Lehrpersonen im Deutschunterricht der Grundschule Hilfestellungen geben und ob sie dabei eher versuchen, alle Schüler gleichermaßen zu berücksichtigen oder aber die Schülerarbeitsphasen zur gezielten Förderung einzelner Schüler nutzen. Studien, die sich auf den öffentlichen Unterricht, also lehrerzentrierte Unterrichtsphasen beziehen, zeigen, dass hier eher leistungsstarke Schüler und Schüler mit einem höheren Selbstkonzept mehr Beitragsmöglichkeiten im Unterricht haben (Brophy & Good, 1976; Gage & Berliner, 1986; Jurik et al., 2013; Koblitz, 1981). Für Schülerarbeitsphasen im Mathematikunterricht der Sekundarstufe konnten Lipowsky und Kollegen (2008) eine etwas häufigere Interaktion der Lehrpersonen mit schwächeren Schülern nachweisen. Das Ziel der im Vortrag vorgestellten Studie besteht daher darin, zu untersuchen, wie Lehrpersonen in Schülerarbeitsphasen des Leseunterrichts im ersten Schuljahr ihre Hilfestellungen an die Schüler verteilen und welche Unterschiede sich hierbei zwischen den Lehrpersonen zeigen. Im Fokus steht dabei die Frage, ob bestimmte Schülermerkmale (Leseleistung, Geschlecht und Selbstkonzept) die Anzahl der Hilfestellungen, die die Lehrpersonen den Schülern zukommen lassen, beeinflussen. Zum anderen wird geprüft, inwiefern Einstellungen und Merkmale der Lehrpersonen (wie beispielsweise Selbstwirksamkeitserwartungen und Zielorientierungen) die gleichmäßige oder ungleichmäßige Verteilung der Hilfestellungen beeinflussen. Zur Beantwortung dieser Fragen wurden 46 Unterrichtsvideos aus dem Projekt PERLE – Persönlichkeits- und Lernentwicklung von Grundschulkindern – analysiert (vgl. Lotz, Lipowsky & Faust, 2013). Die Längsschnittstudie umfasste die gesamte Grundschulzeit, in der sowohl Schüler als auch Lehrpersonen mehrfach getestet und befragt wurden, sodass die Daten aus der Videostudie mit Daten aus Leistungstests und Fragebögen verknüpft werden können. Im Rahmen der Videostudie im Fach Deutsch wurden die Lehrpersonen gebeten, eine Leseübung mit ihren Schülern des ersten Schuljahres durchzuführen. In den Schülerarbeitsphasen wurde jede individuelle Lehrer-Schüler-Interaktion im Event-Sampling-Verfahren identifiziert. Anschließend wurde kodiert, ob diese Interaktionen lehrer- oder schülerinitiiert stattfinden und an welchen Schüler sie sich richten. Weiterhin wurde jede einzelne Hilfestellung identifiziert und anhand unterschiedlicher Kategorien klassifiziert. Für jede Hilfestellung wurde ebenfalls kodiert, an welchen Schüler sie sich richtet. Alle Kodierungen wurden auf Basis selbst entwickelter Manuale von trainierten Beobachtern vorgenommen, wobei die Übereinstimmungen mehrfach getestet wurden und dabei immer über den geforderten Mindestwerten lagen (prozentuale Übereinstimmung > 85%; Cohens Kappa > .70). Die Ergebnisse der ersten Analysen zeigen, dass die Anzahl der Hilfestellungen mit der Leistungsstärke der Klasse negativ zusammenhängt: Je geringer die Leistungsstärke der Klasse, desto häufiger geben die Lehrpersonen insgesamt Hilfestellungen (Lotz & Lipowsky, in Druck). Die Berechnung des individuellen Egalitätsideals nach Sacher (1995) deutet an, dass es bei der Verteilung der Hilfestellungen auf einzelne Schüler sowohl Lehrpersonen gibt, die ihre Hilfen nahezu gleich zwischen den Schülern verteilen als auch Klassen mit starker Ungleichverteilung. Für das gruppenbezogene Egalitätsideal zeigt sich in Regressionsanalysen, dass schwächere Schüler häufiger Hilfestellungen erhalten als stärkere Schüler und Jungen auch unter Kontrolle der Leistung mehr Hilfestellungen bekommen als Mädchen. Ein Einfluss des lesebezogenen Selbstkonzepts auf die Verteilung der Hilfestellungen durch die Lehrpersonen kann nicht nachgewiesen werden. Im Vortrag werden diese Ergebnisse detailliert dargestellt. Zusätzlich wird darauf eingegangen, ob bestimmte Lehrermerkmale und -überzeugungen mit der Verteilung von Hilfestellungen auf einzelne Schüler und Schülergruppen zusammenhängen.

Brophy, J. E., Good, T. L. & Ulich, D. (1976). Die Lehrer-Schüler-Interaktion. München: Urban & Schwarzenberg.

Gage, N.L. & Berliner, D.C. (1986). Pädagogische Psychologie. Weinheim: Beltz.

Jurik, V., Gröschner, A. & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. *Learning and Instruction*, 23, 33-42.

Koblitz, J. (1981). Leistungserwartungen von Lehrern und die Lehrer-Schüler-Interaktion. Eine empirische Untersuchung aus der Sicht sozial-kognitiver Lerntheorien. Frankfurt am Main: Haag + Herchen.

Krammer, K. (2009). Individuelle Lernunterstützung in Schülerarbeitsphasen. Eine videobasierte Analyse des Unterstützungsverhaltens von Lehrpersonen im Mathematikunterricht. Münster u.a.: Waxmann.

Lipowsky, F., Pauli, C. & Rakoczy, K. (2008). Schülerbeteiligung und Unterrichtsqualität. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise. Analyse und Bedeutung unterrichtlichen Handelns* (S. 67–90). Münster, München [u.a.]: Waxmann.

Lotz, M. & Lipowsky, F. (in Druck). Jedem das Seine oder allen das Gleiche? – Individuelle Lernunterstützung in Schülerarbeitsphasen im Leseunterricht der Grundschule. In B. Kopp, S. Martschinke, M. Munser-Kiefer, M. Haider, E. Kirschhock, G. Ranger & G. Renner (Hrsg.), *Individuelle Förderung und Lernen in der Gemeinschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lotz, M., Lipowsky, F. & Faust, G. (Hrsg.). (2013). Technischer Bericht zu den PERLE-Videostudien. In F. Lipowsky & G. Faust: *Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulkindern“ (PERLE)* (Materialien zur Bildungsforschung, 23/3). Frankfurt am Main: Gesellschaft zur Förderung Pädagogischer Forschung (GFPF).

Pauli, C. & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. *Zeitschrift für Pädagogik*, 52 (6), 774–798.

Sacher, W. (1995). *Meldungen und Aufrufe im Unterrichtsgespräch: Theoretische Grundlagen, Forschungsergebnisse, Trainingselemente und Diagnoseverfahren*. Augsburg: Wißner.



## **Kausalattributionen von Leseerfolgen und -misserfolgen bei Grundschülerinnen und -schülern im Zusammenhang mit Leistungsrückmeldungen ihrer Lehrkräfte**

Ein angemessener Umgang mit erlebten Erfolgen und Misserfolgen stellt eine zentrale Voraussetzung für ein gesundes Selbstbild und das Vertrauen in die eigenen Fähigkeiten dar. Für Kinder im Grundschulalter ist dies in Lern- und Leistungssituationen von besonderer Bedeutung, wenn sie dazu aufgefordert sind, Gründe für Erfolge oder Misserfolge während ihrer Lernprozesse zu hinterfragen und sich selbst beim Lernen zu regulieren (Ziegler & Schober, 2001). Folgt man Weiner (1986), so werden verschiedene Ursachenzuschreibungen unterschieden, zum einen internale (von der Person abhängige) und zum anderen externale (von der Person unabhängige) Ursachenzuschreibungen. Zu den internalen Kausalattributionen werden das eigene Anstrengungsverhalten oder die eigenen Fähigkeiten gezählt, zu den externalen Kausalattributionen gehören hingegen der Schwierigkeitsgrad der Aufgaben oder Unterstützungen signifikanter Anderer (z.B. Lehrkräfte, Eltern). Ursachenzuschreibungen können stabilen oder variablen Charakter haben (Buff, 2004): Stabile Faktoren sind nicht veränderbar, variable Faktoren, etwa das eigene Anstrengungsverhalten, können prinzipiell durch eigenes Zutun beeinflusst werden. Kausalattributionen von Erfolgen und Misserfolgen stehen in einem engen Zusammenhang zu Lernprozessen von Kindern: So konnte Faber (2007) zeigen, dass leistungsschwächere Kinder Erfolge in einem Diktat auf günstige Umstände attribuieren (external, variabel), Misserfolge hingegen werden – weniger selbstwertdienlich – mit mangelnden Fähigkeiten begründet (internal, stabil). Befunde von Hosenfeld legen den Schluss nahe, dass Kausalattributionen im Laufe eines Schuljahres durch internale und externale Umstände variieren können. Das Lehrerverhalten hat dabei einen erheblichen Einfluss auf diese Veränderungen (Hosenfeld, 2002). Theoretisch naheliegend ist, dass Kausalattributionen von Erfolgen und Misserfolgen durch die Art der Rückmeldungen seitens der Lehrkräfte erklärt werden können (vgl. Hattie & Timperley, 2007). In diesem Zusammenhang liegen auch Befunde von Dweck, Davidson, Nelson und Enna (1978) vor, die darauf hindeuten, dass sich Kinder von Lehrerinnen bzw. Lehrern implizit geäußerte Kausalattributionen aneignen und sich Fremdattributionen seitens der Lehrkräfte häufig in ungünstigen Rückmeldungen bemerkbar machen. Ungeklärt ist vor dem Hintergrund der dargestellten Befunde bisweilen jedoch, ob und inwiefern Lehrkräfte Kausalattributionen von Erfolgen bzw. Misserfolgen bei Kindern im Grundschulalter angemessen deuten und – mit dem Blick auf die Entwicklung selbstwertdienlicher Kausalattributionen – entsprechende Rückmeldungen bei Lehr-Lernprozessen senden. An dieser Stelle setzt die vorliegende Studie an: Am Beispiel des Leseunterrichts in der Grundschule wurden Schülerinnen und Schüler dritter und vierter Klassen (N=693) zu ihren Kausalattributionen bei Leseerfolgen und -misserfolgen befragt. Der Fokus lag hierbei auf der Erfassung internaler Attributionen, d.h. Ursachenzuschreibungen auf eigene Anstrengungen oder eigene Fähigkeiten. Zusätzlich bearbeiteten die Kinder Aufgabenstellungen aus dem Leseverständnistest ELFE 1-6 („Ein Leseverständnistest für Erst- bis Sechstklässler“; Lenhard & Schneider, 2006). Die Lehrerinnen und Lehrer (N=34) wurden gebeten, Kausalattributionen von Leseerfolgen bzw. -misserfolgen leistungsschwächerer und -stärkerer Kinder einzuschätzen. Zusätzlich hierzu nahmen sie Stellung zu ihrem Rückmeldeverhalten (Anstrengungs- versus Fähigkeitsfeedback) im Leseunterricht. Bei den Analysen wurden zwei Gruppen von Kindern anhand der ermittelten Ergebnisse aus dem Lesetest unterschieden. Leistungsstärkere Schülerinnen und Schüler attribuieren ihre Leseerfolge gleichermaßen auf vorhandene Fähigkeiten und das eigene Anstrengungsverhalten. Die leistungsschwächeren Schülerinnen und Schüler hingegen erklären nicht-selbstwertdienliche Kausalattributionen, indem sie ihre Misserfolge signifikant häufiger auf mangelnde Fähigkeiten zurückführen als auf das eigene Anstrengungsverhalten. Die Fremdeinschätzungen der Lehrerinnen und Lehrer korrespondieren dabei mit den Ursachenzuschreibungen der Kinder. Sie erklären, dass leistungsstärkere Kinder ihre Leseerfolge selbstwertdienlich attribuieren und Erfolge als Ergebnis des eigenen Fähigkeitspotentials und eigener Anstrengungen deuten. Leistungsschwächere Schülerinnen und Schüler würden aus ihrer Sicht signifikant häufiger Misserfolge als mangelnde Talente im Sinne feststehender Fähigkeiten interpretieren als über das Ausbleiben eigener Anstrengungen. Pädagogisch-didaktisch wünschenswert bekunden die Lehrkräfte, dass sie bei der Erteilung der Rückmeldungen nicht zwischen leistungsstärkeren und -schwächeren Kindern differenzieren und beiden Gruppen signifikant häufiger anstrengungsbezogene Rückmeldungen geben als solche, die sich auf ihre Fähigkeiten und Fertigkeiten beziehen. Während anstrengungsbezogenes Feedback aus der Sicht der Lehrkräfte überdurchschnittlich häufig zum Einsatz kommt, werden fähigkeitsbezogene Feedbacks eher selten verwendet.

- Buff, A. (2004). Sind selbst- und fremdbezogene Kausalüberzeugungen austauschbar? Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 36(1), 10-18.
- Dweck, C. S., Davidson, W., Nelson, S. & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental Psychology*, 14, 268-276.
- Faber, G. (2007). Selbstkonzept, Kausalattributionen und Leistungsangst im Rechtschreiben. Schulfachspezifische Analysen zu den Selbsteinschätzungen von Grundschulkindern. Verfügbar über: <http://www.psychologie.uni-hannover.de/fileadmin/psychologie/Dateien/Institut-Downloads/Faber/volltext.pdf> (Datum des Zugriffs: 08.10.2013)
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hosenfeld, I. (2002). Kausalitätsüberzeugungen und Schulleistungen. Münster: Waxmann.
- Lenhard, W. & Schneider, W. (2006). ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.
- Ziegler, A. & Schober, B. (2001). *Theoretische Grundlagen und praktische Anwendung von Reattributionstrainings*. Regensburg: Roderer.



## ***Effekte des computerbasierten Trainingsprogramms „Lautarium“ auf die phonologische Verarbeitung und die Lese- und Rechtschreibleistungen von Grundschulkindern***

Bei der Lese-Rechtschreibstörung (LRS) handelt es sich um eine umschriebene Entwicklungsstörung schulischer Fertigkeiten, die durch massive und persistierende Schwierigkeiten im Erwerb des Lesens und Rechtschreibens gekennzeichnet ist. Diese Schwierigkeiten sind nicht durch mangelnde Intelligenz, fehlende Motivation oder unzureichende schulische Unterweisung erklärbar. Als Folge dieser Beeinträchtigung erreichen die Betroffenen in den meisten Fällen nicht die ihren intellektuellen Fähigkeiten entsprechende Ausbildungsabschlüsse und Beschäftigungsniveaus (Esser, Wyschkon & Schmidt, 2002). Bezüglich der Ursachen besteht weitgehend Einigkeit darüber, dass die LRS aus einem „Kerndefizit“ im Bereich der Repräsentation und Verarbeitung von Sprachlauten, d.h. aus einem phonologischen Defizit, resultiert (Vellutino, Fletcher, Snowling & Scanlon, 2004). Trainingsstudien haben gezeigt, dass die bei LRS beeinträchtigten phonologischen Fertigkeiten trainierbar sind, und dass bei bestimmten phonologischen Trainingsaufgaben ein Transfer auf die Lese- und Rechtschreibleistungen erreicht werden kann. Als besonders effektiv zur Förderung der schriftsprachlichen Leistungen von Kindern mit LRS erwiesen sich Trainingsprogramme, die sowohl phonologische Übungen als auch eine schrittweise Vermittlung der Graphem-Phonem-Zuordnungen beinhalten (vgl. Torgerson, Brooks & Hall, 2006). Ausgehend von diesen Forschungserkenntnissen wurde im Rahmen des BMBF Förderschwerpunkts „Entwicklungsstörungen schulischer Fertigkeiten“ ein computerbasiertes Trainingsprogramm für Grundschul Kinder mit LRS konstruiert, welches Übungen zur Phonemwahrnehmung und phonologischen Bewusstheit mit einer systematischen Vermittlung der Graphem-Phonem-Korrespondenzen und Übungen zum Lesen und Schreiben lautgetreuer Wörter kombiniert (Klatte, Steinbrink, Bergström & Lachmann, 2013). Das adaptive Programm kann von den Kindern weitgehend selbstständig durchgearbeitet werden. Das Training soll fünfmal wöchentlich für etwa 20 Minuten über einen Zeitraum von acht Wochen erfolgen, da sich ein so strukturiertes Intensivtraining in zahlreichen Trainingsstudien bewährt hat. Zur Aufrechterhaltung der Motivation wurde ein virtuelles Belohnungssystem implementiert, welches einen Anreiz zur zügigen und fehlerfreien Bearbeitung der Aufgaben schafft. Im Rahmen einer ersten empirischen Prüfung des Programms (Klatte et al., 2013) konnte bestätigt werden, dass die implementierten Aufgaben die phonologischen Defizite von Kindern mit LRS tatsächlich abbilden. Leseschwache Grundschul Kinder zeigten in allen Übungen schlechtere Leistungen als Kontrollkinder mit mindestens durchschnittlichen Leseleistungen. Im nächsten Schritt wurden die Auswirkungen des Trainings mit dem Programm auf die phonologischen und schriftsprachlichen Leistungen in zwei Studien anhand eines Kontrollgruppendesigns mit drei Messzeitpunkten untersucht (Vortest, Nachttest 1 unmittelbar nach Trainingsende, Nachttest 2 zehn Wochen nach Trainingsende) (Klatte et al., eingereicht). Bei allen Analysen wurden die Vortestleistungen berücksichtigt. An Studie 1 nahmen 41 von LRS betroffene Drittklässler teil. Die Kinder, die mit dem Programm trainierten (N=20) zeigten in beiden Nachttests im Vergleich zur Kontrollgruppe (N=21) bessere Leistungen in Untertests zur Phonemwahrnehmung, phonologischen Bewusstheit und lautgetreuen Rechtschreibung. Weiterhin zeigten die Kinder der Trainingsgruppe im Nachttest 2 signifikant bessere Leistungen im Vorlesen von Wörtern und Pseudowörtern. An Studie 2 nahmen Erstklässler teil, da geprüft werden sollte, ob das Programm prinzipiell auch zur Unterstützung des beginnenden Schriftspracherwerbs geeignet ist. Im Vergleich zur Kontrollgruppe (N=35) zeigten die Kinder der Trainingsgruppe (N=41) in beiden Nachttests bessere Leistungen in Untertests zur phonologischen Verarbeitung und zum Lesen und Rechtschreiben. Insgesamt bestätigen die Befunde die Wirksamkeit des Programms. Allerdings zeigten sich in beiden Studien bei einigen Untertests auch keine oder nur kurzfristig anhaltende Trainingseffekte. Anhand der Evaluationsergebnisse sowie systematischer Analysen der individuellen Trainingsverläufe sollen in der nächsten Projektphase verschiedene, für die beiden Zielgruppen optimierte Programmversionen konstruiert werden.

Esser, G., Wyschkon, A. & Schmidt, M. H. (2002): Was wird aus Achtjährigen mit einer Lese- und Rechtschreibstörung. Ergebnisse im Alter von 25 Jahren. In: Zeitschrift für Klinische Psychologie und Psychotherapie 31 (4), S. 235–242.

Klatte, M., Steinbrink, C., Bergström, K. & Lachmann, T. (2013). Phonologische Verarbeitung bei Grundschulkindern mit schwacher Lesefähigkeit. Lernen und Lernstörungen, 2, 199-215.

Klatte, M., Steinbrink, C., Pröhl, A., Estner, B., Christmann, C. & Lachmann, T. Effekte des computerbasierten Trainingsprogramms "Lautarium" auf die phonologische Verarbeitung und die Lese-Rechtschreibleistungen von Grundschulkindern. Erscheint in: Schulte-Körne, G. (Hrsg.). Legasthenie und Dyskalkulie - Übergänge gestalten von der frühen Kindheit bis ins Erwachsenenalter. Bochum: Winkler.

Torgerson, C.; Brooks, G. & Hall, J. (2006). A systematic review of the research literature on the use of phonics in the teaching of reading and spelling. Nottingham: DfES Publications.

Vellutino, F. R., Fletcher, J. M., Snowling, M.J. & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? Journal of Child Psychology and Psychiatry, 45 (1), 2-40.



## ***Komplexe Evaluationsdesigns in der Leseförderung – am Beispiel der Evaluation des BMBF-Programms „Lesestart – Drei Meilensteine für das Lesen“***

Die Evaluation von bildungspolitischen Programmen und Innovationen ist in Deutschland bislang weder theoretisch-methodisch noch praktisch-politisch angemessen entwickelt, wir erachten Evaluationsforschung darum als eines der großen Desiderate empirischer Bildungsforschung und praktischer Bildungspolitik. Der Beitrag möchte in diesem Zusammenhang ein laufendes Evaluationsprojekt im Bereich der Leseförderung vorstellen, in dem das Tagungsmotto „Die Perspektiven verbinden“ in mehrfacher Hinsicht eingelöst ist: Verbunden werden hier universitäre und außeruniversitäre Forschung (Kooperation von Universität Köln und Universitätsklinikum Hamburg mit einem privatwirtschaftlich arbeitenden Evaluationsinstitut in Berlin), unterschiedliche Disziplinen (Kinder- und Jugendpsychiatrie / Epidemiologie, Sozialwissenschaften, Pädagogik, Leseforschung, Deutschdidaktik) sowie quantitative und qualitative Methoden der Sozialforschung. Interessant ist ferner die Verbindung von Evaluationsteam, Auftraggebern (BMBF) und ausführender Einrichtung (Stiftung Lesen); hier haben sich in den zwei Jahren seit Beginn des Projektes produktive Formen der Kooperation entwickelt, die auch im Hinblick auf die Praxis von Evaluationsprojekten lehrreich sind. Es handelt sich um die Evaluation des größten und längsten Leseförder-Projektes der Bundesrepublik, vom BMBF mit rund 26 Mio. Euro gefördert, von der „Stiftung Lesen“ durchgeführt, Laufzeit: 2011 – 2017 (mit geplanter Verlängerung um eine dritte Phase). Das Leseförderprogramm „Lesestart – Drei Meilensteine für das Lesen“ adressiert Eltern von ein-, drei- und (bei Verlängerung auch) sechsjährigen Kindern (über Kinderärzte, Bibliotheken und Grundschulen) und bietet Materialien zur Unterstützung einer frühen Sprach- und Leseförderung der Kinder. Insgesamt werden zunächst 50 % und in der dritten Phase 100 % aller Kinder eines Jahrganges (bzw. deren Eltern) erreicht. (Weitere Informationen zum Programm finden sich unter [www.lesestart.de](http://www.lesestart.de).) Das BMBF hat mit rund 500.000 Euro Budget eine begleitende Evaluation für den gesamten Zeitraum ausgeschrieben, die einerseits die Qualität der Materialien (Lesestart-Sets) und der Implementierung laufend überprüfen und ggf. nachjustieren soll (im Sinne einer formative Evaluation) und andererseits durch ein Interventions-Kontrollgruppen-Design die Wirksamkeit des Programms überprüfen soll, indem Effekte auf das (Vorlese-)Verhalten der Eltern und die sprachliche Entwicklung der Kinder erhoben und ausgewertet werden. Ein besonderer Evaluationsauftrag besteht darin zu prüfen, ob Eltern und Kinder aus Familien mit Migrationshintergrund und aus schrift- bzw. bildungsfernen Lebenswelten mit dieser Maßnahme tatsächlich und wie intendiert auch überproportional erreicht werden. Auch aus diesem Grunde wurden qualitative Fallstudien mit Familien dieser Schichten in das Evaluationsdesign integriert. Der Auftrag wurde an die InterVal GmbH Berlin (in Kooperation mit Prof. Garbe, Universität Köln und Dr. Barkmann, UKE Hamburg) vergeben. Das Evaluationsteam hat gemeinsam ein theoretisches Wirkungsmodell entwickelt, das Grundlage für die quantitativen und qualitativen Wirkungsanalysen ist. Dazu werden folgende Verfahren eingesetzt: Fortlaufende Fragebogen-Erhebungen und Telefon-Befragungen großer Stichproben von Akteuren (z.B. Kinderärzten) und der Zielgruppe Eltern (repräsentative Stichprobe, die im Panel über mehrere Jahre verfolgt wird), wissenschaftliche Beratung bei der Erstellung der Materialien zur Leseförderung für die Eltern von Ein-, Drei- und Sechsjährigen (z.B. Elternratgeber, Demo-Videos, Bilderbuch-Empfehlungen für verschiedene Altersgruppen), qualitative Fallstudien in Familien speziell bildungsferner Schichten und Migranten-Familien, Sprachtests mit älteren Kindern. Eine umfangreiche Kontrollgruppe (Cati 1; N = 630) wurde bereits vor Beginn des Programms konstituiert und wird bei den sequenziell organisierten Elternbefragungen mit in die Längsschnittuntersuchung einbezogen. Die Evaluation hat im Herbst 2011 begonnen und läuft bis 2017; bereits jetzt liegen umfangreiche Datensätze und jährliche Zwischenberichte für das BMBF vor, so dass außer dem methodischen Design auch die interessantesten Zwischenergebnisse präsentiert werden können. Von besonderem Interesse ist dabei für das Evaluationsteam und die Auftraggeber, ob mit dem dreimaligen Verteilen von Materialien zur Leseförderung tatsächlich eine Verhaltensänderung bei den Eltern aus bildungsfernen Schichten erreicht werden kann; wir werden in der Vorstellung von Ergebnissen darum auf diese Problematik fokussieren.



## **Qualität von Sprachförderung in (vor-)schulischen Institutionen**

Chair(s): Christine Beckerle (Leibniz Universität Hannover)

Diskutant/in: Ulrich Mehlem (Goethe-Universität Frankfurt)

Aufgrund der großen Bedeutung des Themas „Qualität in Erziehungs- und Bildungsinstitutionen“ und der aktuellen Brisanz einer effektiver Sprachförderung für Kinder werden in diesem Symposium beide Forschungsstränge unter den Fragestellungen miteinander verbunden: Wie lässt sich Qualität von Sprachförderung in (vor-)schulischen Situationen beschreiben? Wie sieht diese Qualität aktuell in deutschen Kindergärten und Schulen aus?

Die Qualität von Sprachförderung in (vor-)schulischen Institutionen wird hierbei aus unterschiedlichen Perspektiven untersucht. Diese Perspektivenvielfalt begründet sich mehrfach, u.a. durch:

1. vier verschiedene Sprachförderprojekte, die mit ihren Konzeptionen und Evaluationsergebnissen vorgestellt werden: EvaniK (Koch & Hofmann), Fellbach (Beckerle, Kucharz, & Mackowiak), FörMig (May, Hildenbrandt, & Heckt) sowie TransKIGs (Fried).
2. die Forschenden, die verschiedene Disziplinen vertreten: die Allgemeine Erziehungswissenschaft, Kindheits-, Grundschul- und Sonderpädagogik sowie Entwicklungspsychologie;
3. die Berücksichtigung verschiedener Institutionen und Zielgruppen: Kindertageseinrichtungen und Grundschulen auf der einen Seite, Kinder, pädagogische Fachkräfte, Leitungskräfte und auch Eltern auf der anderen Seite;
4. verschiedene (strukturelle und prozessuale) Qualitätsmerkmale von Sprachförderung: Merkmale der Erwachsenen-Kind-Kommunikation (z.B. Redeanteile, Sprachfördertechniken/ -strategien, sprachliche Korrektheit), Einstellungen und Zufriedenheit der Pädagog/innen im Bereich Sprachförderung, Kooperationen im Rahmen von Sprachförderung, kontextuelle Rahmenbedingungen und sprachliche Leistungen (auf den verschiedenen Sprachebenen, aber auch Erzählkompetenz) der Kinder;
5. verschiedene Forschungsmethoden, die in den Projekten angewandt werden; vorrangig werden Videoauswertungen dargestellt, die durch Fragebogen- und Sprachtestergebnisse ergänzt werden.

Montag,

03.03.2014,

11:15 - 13:00,

PEG-Gebäude

Raum 1.G147

***Sprachförderung sichtbar machen: Erste Ergebnisse zu Sprachförderstrategien in den Videografien des Projekts „EvaniK“***  
***(Evaluation unterschiedlicher Sprachfördermaßnahmen in niedersächsischen Kindertagesstätten)***

Die Vorstellung, dass die Beherrschung der Unterrichtssprache Deutsch eine wichtige Voraussetzung für den schulischen Erfolg aller Kinder sei, hat in den letzten Jahren zu verstärkten Förderbemühungen und grundlegender empirischer Forschung geführt. Dabei dominiert in der deutschsprachigen Forschung die Frage, unter welchen Bedingungen Kinder mit einer anderen Herkunftssprache die Zielsprache Deutsch besonders gut lernen und welche Auswirkungen dies auf ihre zukünftige Bildungskarriere hat. In der linguistisch akzentuierten Forschung wird vor diesem Hintergrund seit längerem bereits auf zielgerichtete Sprachlehrstrategien verwiesen (vgl. z.B. Yoder, Davis, Bishop, & Munson, 1994). Im Gegensatz zur eher „intuitiven“ Verwendung entwicklungsförderlicher sprachlicher Mittel, z.B. in der Mutter-Kind-Kommunikation, unterscheiden sich Sprachlehrstrategien durch ihre zweckorientierte Vermittlung und reflektierte Anwendung. Zudem deuten Studien daraufhin, dass durch systematische Interventionen der Einsatz gezielter Sprachlehrstrategien zur Anregung und Optimierung sprachlicher Interaktionen mit Kindern grundsätzlich geschult werden kann (u.a. Buschmann & Jooss, 2011). Im Rahmen des vom niedersächsischen Forschungsverbundes frühkindliche Bildung und Entwicklung (FBE) geförderten Forschungsprojekts EvaniK - „Sprachförderung für Migrantenkinder im Elementarbereich-Evaluation unterschiedlicher Sprachförderkonzepte in niedersächsischen Kindertagesstätten“ wurden Sprachfördererinnen von Erzieherinnen des Kindergartens und von Lehrkräften, die im Rahmen von „Fit in Deutsch“ ein Jahr vor der Einschulung Sprachförderunterricht im Kindergarten erteilten, mit dem Ziel videografiert, die in diesen Fördererinnen aufscheinenden Sprachförder- und Sprachlehrstrategien zu dokumentieren und zu analysieren (Horrman & Koch, 2011). Eine theoretische Grundlage für die Analyse von sprachlichen Interaktionsprozessen bieten dabei Studien aus dem angloamerikanischen Sprachraum, die im Rückgriff auf soziokulturelle Lerntheorien die direkten Kommunikationen zwischen Kindern und fördernden Personen als Schlüssel für gelingende Sprachentwicklung herausstellen. Soziokulturelle Erklärungsmodelle konzipieren Bildungsprozesse - und so auch die sprachbezogene Entwicklung von Kindern - als Ergebnis eines gemeinsamen Konstruktionsprozesses, in dem Denkprozesse der Beteiligten über Kommunikationen intersubjektiv miteinander verbunden sind (Stamm, 2010). In dieser Perspektive scheinen Kommunikationen für die Sprachentwicklung besonders günstig, in denen Kinder häufig Gelegenheit zu wechselseitig aufeinander bezogenen – kontingenten – Dialogen erhalten (Rogoff, 1990). In Anlehnung an Alexander (2008) werden "echte" dialogische Interaktionen als Möglichkeiten verstanden, die kindliche Teilhabe zu fördern sowie die Beiträge anderer kommentieren können. Demnach ist ein essentieller Bestandteil von Dialogen ein Austausch von Ideen, welche weitere Fragen hervorlocken. Kinder mit einer anderen Herkunftssprache brauchen dabei die Unterstützung von Erwachsenen ("teacher-guided-participation"), die ihre Gesprächsinitiativen wahrnehmen und die ihre Aufgabe darin sehen, Kontinuität und Reziprozität in den Gesprächen sowie peer-to-peer-Interaktionen zu ermöglichen (u. a. Rogoff, 1990; Myhill, 2006). Anhand videografischer Materialien aus unterschiedlichen Sprachfördererheiten sollen im Vortrag Zusammenhänge zwischen den Merkmalen der sprachlichen Erwachsenen Kind-Interaktion und den kontextuellen Rahmenbedingungen dargestellt werden. Außerdem werden erste Teilergebnisse zur Anwendung spezifischer Sprachlehrstrategien innerhalb der betrachteten Kommunikationen diskutiert.

Yoder, P., Davies, B., Bishop, K. & Munson, L. (1994). Effect of adult continuing why-questions on conversational participation in children with developmental disabilities. *Journal of Speech and Hearing Research*, 37, 193-204.

Buschmann, A., Jooss, B., Simon, S. & Sachse, S. (2010). Alltagsintegrierte Sprachförderung in Krippe und Kindergarten. Das "Heidelberger Trainingsprogramm". Ein sprachbasiertes Interaktionstraining für den Frühbereich. *L.O.G.O.S. INTERDISZIPLINÄR*, 2, 84-95

Hormann, O./ Koch, K. (2011). Elementare Sprachförderung in quantitativer und qualitativer (Mehrebenen-)Perspektive: Ergebnisse aus dem Projekt „Sprachförderung für Migrantenkinder-Evaluation unterschiedlicher Sprachförderkonzepte in niedersächsischen Kindertagesstätten“. In: *Empirische Pädagogik*, 25,4. 406-422.

Stamm, M. (2010). *Frühkindliche Bildung, Betreuung und Erziehung*. Stuttgart: UTB

Rogoff, B. (1990). *Apprenticeship in Thinking. Cognitive Development in Social Context*. New York: Oxford University Press.

Alexander, R. (2008). Culture, dialogue and learning: notes on an emerging pedagogy. In N. Mercer & S. Hodgkinson (Hrsg.), *Exploring talk in school* (S. 91–114). London: Sage

Myhill, D. (2006). Talk, talk, talk: Teaching and learning in whole class discourse. *Research Papers in Education*, 21(1), 19-41.



## **Qualitätsaspekte des Sprachförderhandelns von Erzieher/innen und Grundschullehrer/innen: Ergebnisse aus dem Fellbach-Projekt**

Zur Wirksamkeit von Sprachförderung in (vor-)schulischen Institutionen liegen derzeit nur einzelne Befunde in Deutschland vor (Gasteiger-Klicpera, Knapp, & Kucharz, 2010; Roos, Polotzek, & Schöler, 2010; Gretsche & Fröhlich-Gildhoff, 2012). Größtenteils weisen diese darauf hin, dass es noch an effektiven Maßnahmen mangelt, von denen alle Kinder profitieren, und es wird deutlich, dass die sprachfördernden Pädagog/innen zwar über einzelne Wissens- und Handlungsaspekte im Bereich Sprachförderung verfügen, es den meisten aber an professionellen Kompetenzen fehlt (vgl. z.B. Fried, 2010; Redder et al., 2010; Roux & Kammermeyer, 2011; Schneider et al., 2012). Auf Grundlage dieses unbefriedigenden Forschungsstandes entstand das Projekt „Durchgängige Sprachförderung im Kindergarten und in der Grundschule“ (Kucharz & Mackowiak, 2010). In diesem wurde ein Konzept zur alltagsintegrierten durchgängigen Sprachförderung in Kindergarten und Grundschule entwickelt, welches eine Weiterqualifizierung für Erzieher/innen und Grundschullehrer/innen beinhaltet. Im Jahr 2010/11 wurde die einjährige Weiterqualifizierung bestehend aus Fortbildungen und einem Coaching mit 14 Pädagog/innen aus vier Einrichtungen der Stadt Fellbach erprobt. Zur Überprüfung der Wirksamkeit der Weiterqualifizierung und der damit einhergehenden Sprachförderung wurden in einem Prä-Post-Test vor und nach der Intervention Daten zum Wissen und zur Einstellung im Bereich Sprachförderung (Interviews) sowie zum Sprachförderhandeln (Videos) der Pädagog/innen und zu den sprachlichen Leistungen der Kinder (Sprachtests) erhoben. Eine Warte-Kontrollgruppe mit sechs Pädagog/innen aus zwei Fellbacher Einrichtungen, die nicht an der Weiterqualifizierung teilnahmen, diente als Vergleichsgruppe. In diesem Beitrag sollen Ergebnisse der Videoanalysen vorgestellt werden, die Aspekte der Quantität und Qualität des Sprachförderhandelns fokussierten: Zuerst wurden die Menge der Redebeiträge der Pädagog/innen und Kinder ermittelt. Dabei zeigte sich ein gegenläufiger Effekt: Sowohl in der Modell- als auch in der Kontrollgruppe nahm die Redemenge der Pädagog/innen zu; trotzdem sanken zum zweiten Messzeitpunkt – im Vergleich zu den Kindern – die Redeanteile in beiden Gruppen. Erwachsene und Kinder sprechen mehr; die Kinder kommen mehr zu Wort. Im Weiteren wurden die angewandten Sprachfördertechniken in den Videos identifiziert (deduktiv entwickeltes und induktiv ergänztes Kategoriensystem). Während es im Prätest keine signifikanten Unterschiede in der Modell- und Kontrollgruppe gab, nutzte die Modellgruppe im Posttest signifikant mehr Techniken. Hinsichtlich der sprachlichen Korrektheit der Pädagog/innen zeigten sich dagegen im Prä- und Posttest keine Unterschiede zwischen beiden Gruppen. Zur Überprüfung der Wirksamkeit der Sprachförderung durch die Pädagog/innen wurden die Sprachtestergebnisse der Kinder herangezogen: Dabei zeigten die Kinder der Modellgruppe im Posttest tendenziell bessere Leistungen als die der Kontrollgruppe. Die verschiedenen Ergebnisse auf der Ebene der Pädagog/innen und der Kinder sollen in diesem Beitrag diskutiert werden, um Schlussfolgerungen für zukünftige Weiterqualifizierungen im Bereich alltagsintegrierter (vor-)schulischer Sprachförderung zu erarbeiten.

- Fried, L. (2010). Wie steht es um die Sprachförderkompetenz von deutschen Kindergartenerzieherinnen? Ausgewählte Ergebnisse einer empirischen Studie. In: H.-J. Fischer, P. Gansen & K. Michalik (Hrsg.). (2010). Sachunterricht und frühe Bildung. Forschungen zur Didaktik des Sachunterrichts, 9 (S. 205-218). Bad Heilbrunn: Klinkhardt.
- Gasteiger-Klicpera, B., Knapp, W., & Kucharz, D. (2010). Abschlussbericht der Wissenschaftlichen Begleitung des Programms »Sag' mal was - Sprachförderung für Vorschulkinder«. [http://www.sagmalwas-bw.de/media/WiBe%201/pdf/PH-Weingarten\\_Abschlussbericht\\_2010.pdf](http://www.sagmalwas-bw.de/media/WiBe%201/pdf/PH-Weingarten_Abschlussbericht_2010.pdf) [05.09.2013].
- Gretsche, P. & Fröhlich-Gildhoff, K. (2012). Evaluation der Sprachfördermaßnahmen für 3-5jährige Kinder in der Stadt Freiburg. In: K. Fröhlich-Gildhoff, I. Nentwig-Gesemann & H. Wedekind, H. (Hrsg.) (2012). Forschung in der Frühpädagogik V. Schwerpunkt: Naturwissenschaftliche Bildung (S. 275-304). Freiburg: FEL.
- Knapp, W., Kucharz, D., & Gasteiger-Klicpera, B. (2010). Sprache fördern im Kindergarten: Umsetzung wissenschaftlicher Erkenntnisse in die Praxis. Weinheim: Beltz.
- Kucharz, D., & Mackowiak, K. (2010). Konzeption einer Sprachförderung für die Stadt Fellbach. Unveröffentlicht.
- Redder, A., Schwippert, K., Hasselhorn, M., Forschner, S., Fickermann, D., & Ehlich, K. (2010). ZUSE-Diskussionspapier 1: Grundzüge eines nationalen Bildungsprogramms zu „Sprachdiagnostik und Sprachförderung“. Hamburg: ZUSE.
- Roos, J., Polotzek, S., & Schöler, H. (2010). EVAS Evaluationsstudie zur Sprachförderung von Vorschulkindern. Abschlussbericht der Wissenschaftlichen Begleitung der Sprachfördermaßnahmen im Programm »Sag' mal was – Sprachförderung für Vorschulkinder«. Unmittelbare und längerfristige Wirkungen von Sprachförderungen in Mannheim und Heidelberg. [http://www.sagmalwas-bw.de/media/WiBe%201/pdf/EVAS\\_Abschlussbericht\\_Januar2010.pdf](http://www.sagmalwas-bw.de/media/WiBe%201/pdf/EVAS_Abschlussbericht_Januar2010.pdf) [05.09.2013].
- Roux, S. & Kammermeyer, G. (2011). Editorial. Empirische Pädagogik, 25 (4), 283-285.
- Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T., Roßbach, H.-G., Roth, H.-J., Rothweiler, M. & Stanat, P. (2012). Expertise „Bildung durch Sprache und Schrift (BIS)“: Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. [http://www.bmbf.de/pubRD/BIS\\_Expertise.pdf](http://www.bmbf.de/pubRD/BIS_Expertise.pdf) [05.09.2013].



## **Evaluation des Projekts FÖRMIG-Transfer Hamburg: Durchgängige Sprachbildung an der Schnittstelle zwischen Elementar- und Primarbereich**

Sprachliche Kompetenzen gelten heute unumstritten als Schlüsselkompetenz zur schulischen Leistungsfähigkeit sowie zur gesamten gesellschaftlichen Teilhabe. In Hamburg wurde jedoch bei 13% aller knapp fünfjährigen Kinder ein ausgeprägter Sprachförderbedarf festgestellt, d.h. sie haben essentielle Verständnis- und Ausdrucksschwierigkeiten. Unter Kindern mit Migrationshintergrund liegt dieser Anteil gar bei 23,5% (IfBQ, 2013). Ohne gezielte sprachbildende Maßnahmen laufen diese Kinder Gefahr, sich nicht ausreichend in der deutschen Sprache ausdrücken und die bildungssprachlichen Anforderungen des Schulalltags nicht erfüllen zu können. Sowohl in Kindertagesstätten, als auch in Grundschulen gelten daher die gezielte Förderung der Sprachentwicklung sowie bildungssprachlicher Kompetenzen als Kernaufgaben (Freie und Hansestadt Hamburg, 2009; May & Berger, in Druck). Dies kann jedoch nachhaltig nur gelingen, wenn die Sprachbildung systematisch, koordiniert und kontinuierlich umgesetzt wird und sich auf verschiedene Bildungsbereiche erstreckt (Behörde für Schule und Berufsbildung, 2010). Dieses Konzept der durchgängigen Sprachbildung ist das Hauptmerkmal des Projekts FÖRMIG-Transfer, welches im Schuljahr 2010/11 startete und eine Laufzeit von drei Jahren hatte. Ziel des Projekts war die Entwicklung und Erprobung von diagnosebasierten Förderstrategien an der Schnittstelle zwischen Elementar- und Primarbereich. Dabei sollten die sprachlichen Kompetenzen in allen Lern- und Bildungsbereichen in den Blick genommen werden. Gleichzeitig sollte eine stärkere Einbeziehung von Eltern in die Förderung ihrer Kinder gewährleistet und die Kooperation zwischen Kindertagesstätten und Grundschulen gestärkt werden. Zur Umsetzung des Transfervorhabens wurden Entwicklungspartnerschaften gebildet, die sich aus jeweils einer Grundschule in einem sozial besonders belasteten Stadtteil Hamburgs und bis zu drei Kindertagesstätten zusammensetzten. Die Evaluation des Projekts verfolgt folgende Fragen, welche Veränderungen der pädagogischen Praxis durch das Transferprojekt erkennbar werden, welche Effekte der Projektarbeit auf die sprachliche Entwicklung feststellbar sind und ob sich Zusammenhänge zwischen unterschiedlichen Maßnahmen und der Lernentwicklung der Kinder zeigen. Zur Untersuchung dieser Fragestellungen wurde ein längsschnittliches Multi-Kohortendesign mit Kontrollgruppe umgesetzt. Neben der Erhebung der sprachlichen Kompetenzen mit dem Test „KEKS – Kompetenzerfassung in Kindergarten und Schule“ (Bennöhr & May, 2013) wurden jährlich die Leitungen und pädagogischen Fachkräften mittels teilstandardisierter Fragebögen befragt. Die Testungen wurden im Jahr vor der Einschulung sowie in der ersten und zweiten Klasse, jeweils zum Schuljahresanfang, durchgeführt. An dem Projekt waren insgesamt 17 Grundschulen und 28 Kindertagesstätten der Programmgruppe beteiligt, aus denen längsschnittliche Daten von 1617 Kindern (NKohorte1 = 479, NKohorte2 = 548, NKohorte3 = 590) vorliegen. Die Kontrollgruppe besteht aus 429 Kindern aus vier Grundschulen und sieben Kindertagesstätten. Erste varianzanalytische Auswertungen zeigen, dass sich der Lernzuwachs der Programm-gruppenkinder zwischen dem ersten und zweiten Erhebungszeitpunkt von Kohorte zu Kohorte signifikant erhöht hat. Bei den Kindern der ersten Kohorte ist ein leichter Rückgang der erzielten Punkte festzustellen, wohingegen sich die Kinder der zweiten und dritten Kohorte verbessern konnten und die Kinder der dritten Kohorte die höchsten Differenzwerte erzielen. Im Vergleich mit der Kontrollgruppe verzeichnet die erste Kohorte der Programmgruppe einen signifikant niedrigeren Differenzwert, wohingegen sich die Differenzwerte zwischen der zweiten und dritten Kohorte sowie der Kontrollgruppe nicht unterscheiden. Nach der Umsetzung der Projektziele während der Projektlaufzeit gefragt, zeigten sich die Leitungen der beteiligten Einrichtungen am zufriedensten mit der Gestaltung des Übergangs von der Kindertagesstätte in die Schule sowie mit dem Austausch über Sprachfördermethoden und -materialien und der Ausarbeitung eines gemeinsamen Konzepts zur Sprachbildung in der Entwicklungspartnerschaft. Positiv verändert habe sich darüber hinaus auch der Stellenwert der Sprachbildung im Bewusstsein des Kollegiums. Am schwierigsten erwies sich nach Aussage den Leitungen der Einbezug der Eltern in die Sprachförderarbeit. Insgesamt geben die ersten Analysen einen Hinweis darauf, dass die Effekte des Transferprojekts auf die Leistungsentwicklung der Kinder als eher gering einzustufen sind, wohingegen sich jedoch die Zusammenarbeit der Einrichtungen aus Sicht der Beteiligten deutlich verbessert hat. Weitere differentielle Analysen hinsichtlich der Merkmale für eine gelingende Kooperation und Sprachbildung werden durchgeführt und sollen präsentiert werden.

Behörde für Schule und Berufsbildung, Universität Hamburg, Institut für International und Interkulturell Vergleichende Erziehungswissenschaft (2010). Expertise zum FÖRMIG-Transfer Projekt. Diagnosegestützte durchgängige Sprachbildung an der Schnittstelle zwischen Elementar- und Primarbereich. Verfügbar unter: [cosmea.erzwiss.uni-hamburg.de/cosmea/core/corebase/mediabase/foermig/materialien/Kompetenzzentrum/Foermig\\_Expertise\\_doppel.pdf](http://cosmea.erzwiss.uni-hamburg.de/cosmea/core/corebase/mediabase/foermig/materialien/Kompetenzzentrum/Foermig_Expertise_doppel.pdf) [Zugriff: 13.10.2013].  
Freie und Hansestadt Hamburg, Behörde für Soziales, Familie, Gesundheit und Verbraucherschutz (2009). Landesrahmenvertrag Kinderbetreuung in Tageseinrichtungen. Verfügbar unter: <http://www.hamburg.de/contentblob/110036/data/landesrahmenvertrag.pdf> [Zugriff: 13.10.2013].  
Institut für Bildungsmonitoring und Qualitätsentwicklung (2013). Vorstellungsverfahren Viereinhalbjähriger. Bericht über die Auswertung der Ergebnisse im Schuljahr 2012/13. Unveröffentlicht.  
May, P. & Berger, C. (in Druck). Monitoring des Hamburger Sprachförderkonzepts. In: R. Valtin & B. Hofmann (Hrsg.), IGLU - Wie lässt sich eine nachhaltige Verbesserung der Lesekompetenz erreichen? Berlin: DGSL.

## ***Häusliche und institutionelle Kontexteffekte auf das Erzählen von Kindergartenkindern: Ergebnisse aus dem TransKiGs-Projekt***

Die Transition vom Kindergarten in die Grundschule stellt für den kindlichen Kompetenzerwerb eine exponierte Entwicklungsphase dar (Griebel & Niesel, 2004), die in einem Spannungsfeld zwischen „Problempotential“ und „Entwicklungspotenz“ (Schneider, 2001) liegt. In dieser Phase werden die Entwicklungs- und Lernprozesse des Kindes durch ein komplexes Geflecht kontextueller Faktoren beeinflusst. So unterstreichen die Befunde der Transitionsforschung, dass die Art und Weise, wie Eltern, pädagogische Fachkräfte und Lehrkräfte junge Kinder in dieser Phase begleiten bzw. unterstützen und herausfordern, deren Kompetenzentwicklung bedeutsam moderiert. Dabei spielen die häuslichen Kontexte eine mindestens ebenso bedeutsame Rolle wie die institutionellen Kontexte. Vor diesem Hintergrund war es u.a. Ziel der wissenschaftlichen Flankierung von TransKiGs, diejenigen kontextuelle Faktoren zu identifizieren, welche sich in dieser Phase am bedeutsamsten auf das Erzählen von Kindern auswirken; also einer Fähigkeit, welche einen wichtigen Grundstein für spätere schulische Literacy legt, zu der u.a. Lese- und Textproduktionsfähigkeiten zählen (Becker, 2005; Quasthoff et al., 2011). Für dieses Projekt mussten eigens Verfahren entwickelt bzw. adaptiert werden. Konkret eingesetzt wurden: a) ein Test zur Erfassung der Erzählkompetenz; b) ein Einschätzverfahren zur Erfassung lernprozessbezogener Kompetenzen (Arbeitsverhalten, Resilienz etc.); c) ein Fragebogen zur Erfassung des elterlichen Involvements sowie d) ein Fragebogen für pädagogische Fach- und Lehrkräfte zur Erfassung der institutionellen Struktur- und Orientierungsqualität. In der berichteten Teilstudie wird speziell der Frage nachgegangen, ob und gegebenenfalls wieweit sich das häusliche oder/und institutionelle Modell sowie die häusliche und/oder institutionelle Sprachförderdidaktik bedeutsam im Status und in der Entwicklung der Erzählkompetenz von Kindergartenkindern niederschlagen. Für die Erfassung der Erzählkompetenz musste ein neues Verfahren konstruiert werden. Das Narrationsentwicklungsinstrument (NEI-KiGs) erfasst die Erzählkompetenz Fünf- bis Siebenjähriger in den Erzählformen Fantasie-, Bilder- und Nacherzählung (Briedigkeit, Stude, & Fried, 2011). Die Testdurchführung erfolgte handpuppengestützt in einer Einzelsituation mit dem Kind. Neben diesem „harten“, standardisierten Kompetenztest wurde auch ein „weiches“ Assessmentverfahren eingesetzt. So wurden die pädagogischen Fachkräfte der Kinder gebeten, einen Einschätzbogen auszufüllen, welcher prozessgebundene Lernkompetenzen ermittelt. Zur Erfassung des häuslichen Kontexts wurde ein Fragebogen (EI-KiGs) eingesetzt, welcher auf das Elterninvolvement zielt und eine Adaption eines angloamerikanischen Verfahrens darstellt (Fried, Isele, & Stude, 2011; vgl. Fantuzzo et al., 2002). Die institutionelle Kontextqualität wurde zum einen mit einem ebenfalls neu entwickelten Fragebogen (Struktur-, Orientierungsqualität), zum anderen mit einem bereits erprobten Einschätzverfahren zur sprachförderrelevanten Prozessqualität (DO-RESI; Fried & Briedigkeit, 2008) erfasst. Die Messgüte der Kind- und Kontexterfassungsinstrumente ist mehrfach geprüft und bestätigt worden (Objektivität, Reliabilität, Konstrukt- und Kriteriumsvalidität). Die Stichprobe setzte sich aus rund 100 Kindern, deren Müttern und/oder Vätern sowie deren pädagogische Fachkräfte zusammen. Der so gewonnene Datensatz wurde mittels Regressionsanalysen ausgewertet. Die Ergebnisse zeigen auf, dass es Effekte gibt. Abschließend wird über pädagogische Folgerungen nachgedacht (vgl. Fried, 2012; Fried et al., 2009; Fried et al., 2011).

Becker, T. (2005). Mündliche Vorstufen literaler Textentwicklung: vier Erzählformen im Vergleich. In H. Feilke & R. Schmidlin (Hrsg.), *Literale Textentwicklung. Untersuchungen zum Erwerb von Textkompetenz* (S. 19-42). Frankfurt: Peter Lang.

Briedigkeit, E., Stude, J. & Fried, L. (2010). Narrative Kompetenzen: Erste Ergebnisse zu ihrer Modellierung und diagnostischen Erfassung im Vorschulalter. In B. Schwarz, P. Nenniger, P. & R.S. Jäger (Hrsg.), *Erziehungswissenschaftliche Forschung – nachhaltige Bildung. Beiträge zur 5. DGfE-Sektionstagung "Empirische Bildungsforschung"/ AEPF-KBBB im Frühjahr 2009* (S. 269-278). Landau: Empirische Pädagogik.

Fantuzzo, J.W., Tighe, E., McWayne, C.M., Davis, G. & Childs, S. (2002). Parent involvement in early childhood education and children's peer play competencies: An examination of multivariate relationships. *NHSA Dialog: A Research-To-Practice Journal for the Early Intervention Field*, 6, 3-21.

Fried, L. (2012). Sprachlich-kognitive Entwicklung in der Phase der Transition vom Kindergarten zur Grundschule – ausgewählte Ergebnisse der TransKiGs-Längsschnittstudie. In: *Sprache – Stimme – Gehör*, 36, S. 1-4.

Fried, L. & Briedigkeit, E. (2008). *Sprachförderkompetenz*. Berlin: Cornelsen Scriptor.

Fried, L., Hoefl, M., Isele, P., Stude, J. & Wexler, W. (2009). Flankierende Begleitforschung. In: *Lenkungsgruppe TransKiGs u.a. (Hrsg.): Übergang Kita – Schule. Zwischen Kontinuität und Herausforderung. Materialien, Instrumente und Ergebnisse des TransKiGs-Verbundprojekts* (S. 56-58). Weimar: das netz.

Fried, L., Hoefl, M., Isele, P., Stude, J. & Wexler, W. (2011). *Wissenschaftliche Flankierung von TransKiGs. BMBF-Abschlussbericht*. Dortmund: TU, Fak 12.

Fried, L., Isele, P. & Stude, J. (2011). Väter-Involvement während der Phase des Übergangs vom Elementar- in den Primarbereich – Erste Befunde der Wissenschaftlichen Flankierung von TransKiGs. In: S. Oehlmann, Y. Manning-Chlechowitz & M. Sitter (Hrsg.), *Frühpädagogische Übergangsforschung: Von der Kindertageseinrichtung in die Grundschule* (S. 107-120). Weinheim: Juventa.

Griebel, W. & Niesel, R. (2004). Transitionen. Fähigkeit von Kindern in Tageseinrichtungen fördern, Veränderungen erfolgreich zu bewältigen. Weinheim: Beltz.

Quasthoff, U., Fried, L., Katz-Bernstein, N., Lengning, A., Schröder, A. & Stude, J. (2011). (Vor)Schulkinder erzählen im Gespräch. Kompetenzunterschiede systematisch erkennen und fördern. *Baltmannsweiler: Schneider Verlag Hohengehren*.

Schneider, I. K. (2001). Kinder kommen in die Schule. Schulanfang aus biographischer Perspektive. In I. Behnken & J. Zinnecker (Hrsg.), *Kinder. Kindheit. Lebensgeschichte. Ein Handbuch* (S. 458 – 472). Seelze-Velber: Kallmeyer.





## **Kognitive Aktivierung und Konstruktive Unterstützung als Prozessqualitäten des Lehrens und Lernens**

### Chair(s): Gerhard Minnameier (Goethe-Universität Frankfurt)

Theoretischer Hintergrund: „Kognitive Aktivierung“ und „Konstruktive Unterstützung“ gelten als zentrale Kriterien für Unterrichtsqualität. Beide Variablen beziehen sich auf den Prozess des Unterrichts bzw. der Interaktion der Lernenden mit Lehrpersonen, Lernmaterial und Mitlernenden. Gleichwohl werden sie bislang typischerweise nicht im Prozess erhoben (z.B. in Form von Videoanalysen), sondern anhand von Unterrichtsmaterialanalysen und Befragungen (vgl. z.B. Kunter & Voss, 2011). Zwar lassen sich auf Basis hoch-inferenter Ratings Qualitätsdimensionen und Interaktionen abbilden und vergleichend evaluieren, dies bislang aber nur in der Gesamtschau der beobachteten Unterrichtsstunden (vgl. z.B. Clausen, Klieme & Reusser, 2003; Alonzo, Kobarg & Seidel, 2012).

Eine detailliertere Erfassung wäre aus verschiedenen Gründen sinnvoll. Die Unterrichtsforschung hat gezeigt, dass erstens der Lernerfolg der SchülerInnen wesentlich von Tiefenstrukturen des Unterrichts abhängt (vgl. Kunter & Voss, 2011; Pauli et al., 2008; Rakoczy et al., 2010). Dies erscheint zweitens umso wichtiger, als beim heutigen Stand der Forschung davon ausgegangen werden kann, dass „Kognitive Aktivierung“ und „Konstruktive Unterstützung“ nicht von allgemeinen Instruktionsmerkmalen (fragend-entwickelnder Unterricht, Frontalunterricht, Formen konstruktivistischen Unterrichts) abhängen (vgl. Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Pauli, Drollinger-Vetter, Hugener & Lipowsky, 2008), sondern von der Qualität der inhaltlichen Auseinandersetzung im Rahmen der jeweiligen Instruktionsmodi. Drittens wird heute durchgängig vom so genannten Angebots-Nutzungsmodell (vgl. Fend, 1998, 2002; Helmke, 2012; vgl. auch Pauli & Reusser, 2006; Kunter & Voss, 2011) ausgegangen, was bedeutet, dass gerade der Zusammenhang von Angebot und Nutzung von Potentialen für „Kognitive Aktivierung“ und „Konstruktive Unterstützung“ zu betrachten ist, und damit das Lehr-Lern-Prozessgeschehen.

Ziele, Forschungsfrage: Bisherige Versuche, Unterrichtsqualität empirisch zu erheben, basieren auf v.a. hoch-inferenten Ratings (vgl. Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Kunter & Voss, 2011), und die zentrale Herausforderung besteht darin, bzgl. der unterrichtlichen Tiefenstrukturen niedriger inferente Ratings zu realisieren bzw. einzelne Lehr-Lernprozesse innerhalb ganzer Unterrichtsstunden detaillierter im Hinblick auf die hier fokussierten Qualitätsaspekte einschätzen zu können. Dazu ist es erforderlich, den Unterricht mit videographischen Verfahren so zu erfassen, dass Informationen über Tiefenstruktur und stattfindende Lehr-Lernprozesse in den Videodaten abgebildet werden. Die anschließenden Analysen setzen an diesen Lehr-Lernprozessen und damit an der Inhaltlichkeit des Unterrichts an.

Methode: 4 x 90 Minuten Rechnungswesenunterricht wurden nach zwei verschiedenen fachdidaktischen Ansätzen vollständig (mit Unterrichtsskript und Arbeitsmaterialien) ausgearbeitet und in Teilschritte zerlegt, und zwar im Sinne von Teilprozessen des Problemlösens. Hier lehnen wir uns an eine Kategorisierung aus der PERLE-Studie an (vgl. Lauterbach, Gabriel & Lipowsky, 2013), die wir allerdings um der lernprozessbezogenen Systematik und der inhaltlichen Trennschärfe der Kategorien willen modifiziert haben. Wir verstehen dabei „Kognitive Aktivierung“ als Prozess der Induktion eines Problems beim Lernenden, die einen Problemlöse-Prozess, vermittelt über die Teilschritte (a) der Generierung eines Lösungsvorschlags, (b) der Umsetzung und (c) der Beurteilung dieses Lösungsvorschlags in Gang setzen soll. „Konstruktive Unterstützung“ bezieht sich dabei auf die Ermöglichung und Begleitung dieser drei Teilprozesse (a)-(c). Noch etwas detaillierter lassen sich diese Teilprozesse der Probleminduktion und Problemlösung anhand einer inferenziellen Lehr-Lerntheorie untersuchen, die diese Prozesse als Formen logischen Schließens (Abduktion, Deduktion, Induktion) rekonstruiert, was insbesondere für die Lokalisation von Fehlern und den Umgang mit ihnen von Bedeutung ist. Die Kodierung erfolgt dabei zunächst, indem die Teilschritte kodiert werden und im zweiten Schritt deren Qualitäten. Im dritten Schritt sollen Fehler und der Umgang mit Fehlern explorativ analysiert werden. Die Kodierungen werden direkt am Videomaterial vorgenommen. Grundlage für die Analysen sind Videoaufzeichnungen von Unterrichtsstunden, in denen SchülerInnen in Kleingruppen arbeiten sowie die zeitsynchron aufgezeichneten Desktop-Daten von Tablet-PCs, auf denen die Lernenden Arbeitsblätter bearbeiten.

Ergebnisse: Erste Videodaten aus Schulklassen liegen vor und werden derzeit analysiert (8. und 9. Klasse Gymnasium sowie berufliche Schulen). Dargestellt werden soll das Kodierverfahren und die ersten Ergebnisse qualitativer Art sowie Reliabilitätsanalysen.

Bedeutsam erscheint zunächst die exakte Phasenkodierung des Unterrichts, die detaillierte Analysen zur Unterrichtsqualität erst ermöglicht.

Alonzo, A., Kobarg, M., & Seidel, T. (2012). Pedagogical content knowledge as reflected in teacher-student interactions: analysis of two video cases. *Journal of Research in Science Teaching*, 49(10), 1211–1239.

Clausen, M., Klieme, E., & Reusser, K. (2003). Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilungen. Ein Vergleich zwischen Deutschland und der deutschsprachigen Schweiz. *Unterrichtswissenschaft*, 31, 122-141.

Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54, 222-237.

Kunter, M., & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 85-113). Münster: Waxmann.

Lauterbach, C., Gabriel, K., & Lipowsky, F. (2013). Hoch inferentes Rating: Kognitive Aktivierung im Mathematikunterricht. In M. Lotz, F. Lipowsky & G. Faust (Hrsg.), *Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE)*. 3. Technischer Bericht zu den PERLE-Videoanalysen (S.405-421). Materialien zur Bildungsforschung. Band 23/3. Frankfurt am Main: GEBF.

Pauli, C., Drollinger-Vetter, B., Hugener, I., Lipowsky, F. (2008). Kognitive Aktivierung im Mathematikunterricht. *Zeitschrift für Pädagogische Psychologie*, 22, 127-133.

Rakoczy, K., Klieme, E., Lipowsky, F., & Drollinger-Vetter, B. (2010). Strukturierung, kognitive Aktivität und Leistungsentwicklung im Mathematikunterricht. *Unterrichtswissenschaft*, 38, 229-246.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G150



## Kompetenzdiagnostik beim Durchführen von Experimenten im naturwissenschaftlichen Unterricht

Theoretischer Hintergrund und Fragestellung: Innerhalb der Bildungsstandards nimmt das Experimentieren in den naturwissenschaftlichen Fächern als Teil der Erkenntnisgewinnung einen zentralen Stellenwert ein. Über welche experimentellen Kompetenzen Schüler am Ende der Mittelstufe tatsächlich verfügen und wie sich konkret die Schülerleistungen empirisch erfassen lassen, ist sowohl für die Bildungsforschung als auch für die Schulpraxis von großer Relevanz. Als Problemlöseprozess setzt das Experimentieren eine Reihe von Kompetenzen voraus (vgl. u. a. Klahr & Dunbar, 1988; Mayer, 2007). Möller, Grube und Mayer (2008) konnten mit Hilfe schriftlicher Testverfahren empirisch nachweisen, dass der Prozess vier Teilkompetenzen impliziert: „naturwissenschaftliche Fragen formulieren“, „Hypothesen generieren“, „Untersuchungen planen“ sowie „Daten analysieren und Schlussfolgerungen“ ziehen. Bislang weitgehend unberücksichtigt blieb die Modellierung der Teilkompetenz „Untersuchungen durchführen“. Ein Messen kognitiver und manueller Fähigkeiten beim Durchführen von Experimenten setzt jedoch den Einsatz von *hands-on-Aufgaben* voraus. Generell werden zwei Arten von sogenannten *Performance Assessments* unterschieden, die sich hinsichtlich ihrer Auswertungsverfahren unterscheiden. Bei prozessbezogenen Experimentaltests erfolgt die Auswertung auf der Basis von Handlungsanalysen, die meist aus Beobachtungen mit Hilfe von Videoanalysen abgeleitet werden (z. B. Schauble et al., 1992; Baxter & Shavelson, 1994; von Aufschnaiter & Rogge, 2010). Bei produktbezogenen Tests erfolgt die Auswertung anhand von Schülerergebnissen, in der Regel mit Hilfe von Ergebnisprotokollen (z. B. Harmon, 1997; HarmoS, 2008). Ziel dieser Studie war es, ein Testinstrument zu entwickeln, welches u. a. den folgenden Anforderungen genügt: 1) Individuelle Erfassung der kognitiven und manuellen Schülerfähigkeiten bei der Durchführung von Experimenten, 2) Dokumentation prozeduraler Fähigkeiten ohne Videoanalysen, 3) Authentische Experimentiersituation, 4) Berücksichtigung eines biologischen Kontextes, 5) Balance zwischen offenen Testitems und leichter Codierbarkeit, 6) Bearbeitbarkeit ohne inhaltliches Vorwissen, 7) hohe Praktikabilität (Zeit, Kosten, Durchführbarkeit). Die sich daraus ableitenden Forschungsfragen lauten wie folgt: 1) Inwieweit lassen sich experimentelle Kompetenzen von Schülern bei der Durchführung biologischer Experimente mit Hilfe des entwickelten praktischen Experimentiertests messen? 2) Lassen sich die vier a priori angenommenen Subkompetenzen für die Durchführung eines Experiments reliabel erfassen? 3) Über welche experimentellen Kompetenzen bei der Durchführung von biologischen Experimenten verfügen Schüler der Sekundarstufe 1?

Untersuchungsdesign und Forschungsmethodik: Der im Rahmen dieser Studie entwickelte Experimentiertest berücksichtigt vier a priori definierte Subkompetenzen: 1) Experiment aufbauen, 2) Daten dokumentieren, 3) Beobachtungen und Messungen durchführen und 4) Fehler bei der Durchführung erkennen. 330 Gymnasialschüler (M Alter = 13,81, davon 56,4% Mädchen) führten je 45 Min. lang zwei zuvor unbekannte Experimente mit biologischem Kontext durch. In einem Testheft mit offenen, halboffenen und geschlossenen Aufgabenformaten dokumentierten die Schüler, begleitend zu ihrem Handlungsprozess, ihr experimentelles Vorgehen. Mit Hilfe eines Fotos wurde die Subkompetenz „Experiment aufbauen“ erfasst. Insgesamt wurden 52 kognitive und manuelle Fähigkeitsitems codiert (Intercoder-Reliabilität: Cohens Kappa = 0,91-1,00). Die Auswertung der Daten erfolgte mittels einer Rasch-Analyse mit Hilfe des Programms WINSTEPS, Korrelationsanalysen und Signifikanztests mittels nachgeschalteter Analysen in SPSS.

Ergebnisse und Diskussion: (1) Das Messinstrument weist eine hohe Testgüte auf. Analysiert wurden die MNSQ- und ZSTD-Werte der Personen sowie der Items. Die Item-Reliabilität beträgt .99, die Personen-Reliabilität .67. Zudem zeigen die Aufgaben eine angemessene Schwierigkeit ( $M\theta = .24$ ). (2) Die Subkompetenzen „Experiment aufbauen“ und „Daten dokumentieren“ lassen sich bereits reliabel erfassen. Die Subkompetenzen „Beobachtungen und Messungen durchführen“ sowie „Fehler bei der Durchführung erkennen“ bedürfen einer höheren Anzahl von Test-Items. (3) Insgesamt schneiden die Mädchen im Experimentiertest besser ab als die Jungen ( $p < .05$ , t-test). Es lassen sich sehr geringe Korrelationen zwischen Experimentierkompetenz und Lesekompetenz ( $r = .22$ ), mathematischen Fähigkeiten ( $r = .25$ ) und biologischem Fachwissen ( $r = .19$ ) erkennen, was darauf hindeutet, dass diese Fähigkeiten die Durchführung des Experimentiertests kaum beeinflussen. Detaillierte Analysen der einzelnen Items, der experimentellen Kompetenzen der Schüler in den Subkompetenzen sowie mögliche Implikationen für die Schulpraxis werden auf der Tagung ausführlich diskutiert.

Aufschnaiter, C. v. & Rogge, C. (2010). Wie lassen sich Verläufe der Entwicklung von Kompetenzen modellieren? Zeitschrift für Didaktik der Naturwissenschaften, 16, 95-114. Baxter, G. P. &

Shavelson, R. J. (1994). Science Performance Assessment. Benchmarks and Surrogates. International Journal of Educational Research, 21 (3), 279-298.

Grube, C., Möller, A. & Mayer, J. (2007). Dimensionierung wissenschaftsmethodischer Kompetenz. In Bayrhuber, H. et al. (Hrsg.): Ausbildung und Professionalisierung von Lehrkräften. (S. 31-34). Kassel: IPN.

Harmon, M., Smith, T., Martin, M., Kelly, D., Beaton, A., Mullis, I., et al. (1997). Performance Assessment in IEA's Third International Mathematics and Science Study (TIMSS). Chestnut Hill, MA: Boston College.

HarmoS: Konsortium HarmoS Naturwissenschaften (2008). HarmoS Naturwissenschaften+: Kompetenzmodell und Vorschläge für Bildungsstandards. Konsortium HarmoS Naturwissenschaften.

Klahr, D., & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. Cognitive Science, 12, 1-48.

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.). Theorien in der biologiedidaktischen Forschung (S. 177-186). Berlin, Heidelberg: Springer.

Schauble, L., Glaser, R., Raghavan, K. & Reiner, M. (1992). The Integration of Knowledge and Experimentation Strategies in Understanding a Physical System. Applied Cognitive Psychology, 6, 321-343.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G150



Theoretischer Hintergrund: Während Schüler/innen am Ende der Grundschulzeit naturwissenschaftlichen Fragestellungen gegenüber noch aufgeschlossen sind (Wittwer, Saß & Prenzel, 2008), zeigt sich im Verlauf der Schulzeit ein abnehmendes Interesse der Schüler/innen an den Naturwissenschaften (Baumert & Köller, 1998; Baumert, Bos & Lehmann, 2000; Hidi, 2000). Ein Faktor, der dieser Abnahme entgegenwirkt, ist ein anwendungsorientierter und kontextbezogener Unterricht (Hoffmann, Häußler & Lehrke, 1998; Parchmann & Gräsel, 2004), in dem die Schüler/innen Fragen aus ihrem Lebenskontext nachgehen und eigene Erfahrungen beim Experimentieren sammeln können (Duit & Mikelskis-Seifert, 2007). Trotz dieser Ergebnisse zeigen die PISA-Befunde, dass Schüler/innen im naturwissenschaftlichen Unterricht selten Gelegenheit haben, eigenständig zu forschen und Experimente zu machen (PISA-Konsortium, 2006). Nicht zuletzt aufgrund dieser Befunde hat sich in den vergangenen Jahren eine Vielzahl an mobilen Lernlaboren etabliert, um Schüler/innen Erfahrungen mit Experimenten zu ermöglichen und das Interesse der Schüler/innen zu steigern (Brandt, Möller, & Kohse-Höinghaus, 2008; Guderian, 2007; Pawek, 2009). Die hier vorgestellte Evaluationsstudie untersucht das mobile Lernlabor MobiLab (<http://mobilab-nw.ch>), das Schweizer Primar-Schüler/innen (4.-6. Klasse) die Möglichkeit bietet, im Unterricht eigenständig zu experimentieren. Das Besondere an diesen Experimenten ist, dass diese mit gewöhnlichen Alltagsgegenständen durchführbar sind, wodurch ein enger Lebensweltbezug erreicht werden soll. Dies soll den Schüler/innen einerseits erleichtern, sich naturwissenschaftliche Phänomene selbst zu erschließen, andererseits aber auch das Interesse der Schüler/innen am Experimentieren fördern. Die Kinder sollen zudem angeregt werden, die Experimente zu Hause mit ihren Eltern nachzumachen und diese somit in die naturwissenschaftliche Ausbildung einzubeziehen.

Fragestellung: Ausgehend von den genannten Zielen des MobiLabs werden folgende Fragen beantwortet: Frage 1: Wie verändert sich das Interesse der Schüler/innen am Experimentieren und am naturwissenschaftlichen Unterricht nach dem MobiLab-Einsatz? Frage 2: Machen die Schüler/innen die Experimente mit ihren Eltern nach? Welche Faktoren begünstigen, dass Nachmachen der Experimente? Es wird angenommen, dass aufgrund des engen Lebensweltbezugs der Experimente das Interesse der Schüler/innen gesteigert wird (Hypothese 1). Das hohe Interesse drückt sich auch darin aus, dass die Kinder nach dem MobiLab-Einsatz die Experimente zu Hause mit ihren Eltern nachmachen (Hypothese 2a). Dies auch deswegen, da die Experimente nur mit Alltagsgegenständen durchgeführt werden und daher leicht nachgemacht werden können. Als Faktoren, die das Nachmachen begünstigen, werden das generelle Interesse am Experimentieren, die Absicht das Experiment nachzumachen und das Interesse der Eltern an Naturwissenschaften angenommen (Hypothese 2b).

Methode: Die Evaluation des MobiLabs startete im Herbst 2013 und setzt sich aus verschiedenen Erhebungsmethoden zusammen (Fragebogen, Videostudie und qualitative Interviews). Bis Ende des Schuljahres 2014/2015 werden etwa 600 Schüler/innen und ihre Lehrpersonen zu drei Messzeitpunkten (direkt vor, direkt nach sowie 2.5 Monate nach dem MobiLab-Unterricht) mittels standardisierter Fragebogen befragt. Der vorliegende Beitrag fokussiert auf die Fragebogen-Daten der Schüler/innen, die bis zur Konferenz vorliegen (vermutetes N = 200 Schüler/innen). Zur Beantwortung der Frage 1 wird mittels varianzanalytischer Verfahren die Veränderung des Interesses am Experimentieren (Skala angelehnt an Kunter et al., 2002) und am Unterricht (Skala entnommen aus Rakoczy, Buff & Lipowsky, 2005) über die drei Messzeitpunkte verglichen. Zur Beantwortung der Frage 2 werden die Daten zum Messzeitpunkt 3 ausgewertet, zu dem die Schüler/innen angeben, ob sie die Experimente zu Hause mit ihren Eltern nachgemacht haben. Abschließend werden mittels multipler, logistischer Regressionen Faktoren herausgearbeitet, die die Wahrscheinlichkeit begünstigen, dass die Schüler/innen eines der Experimente zu Hause nachmachen.

Ergebnisse: Derzeit können lediglich die Daten einer Klasse analysiert werden. Erste Analysen zeigen, dass das Interesse der Schüler/innen am Unterricht mit dem MobiLab hoch zu sein scheint ( $M=3.21$ ,  $SD=0.51$ ;  $N=17$ ) und dass ein Grossteil der Schüler/innen (70.6%) vorhat, eines der Experimente zu Hause nachzumachen. Ob die Schüler/innen dies auch tun, werden die Ergebnisse des dritten Messzeitpunkts zeigen, für den bis zum Kongress Daten vorliegen. Zudem werden am Kongress die genannten Fragen mit einer grösseren Stichprobe beantwortet.

Baumert, J., Bos, W. & Lehmann, R. (Hrsg.) (2000). TIMSS/III - Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band I: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen: Leske + Budrich.

Baumert, J. & Köller, O. (1998). Interest research concerning secondary level I: An overview. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), Interest and learning. Proceedings of the Seoon-Conference on interest and gender (S. 241–256). Kiel: IPN.

Brandt, A., Möller, J., & Kohse-Höinghaus, K. (2008). Was bewirken außerschulische Experimentierlabors? Ein Kontrollgruppenexperiment mit Follow up-Erhebung zu Effekten auf Selbstkonzept und Interesse. Zeitschrift für Pädagogische Psychologie, 22(1), 5–12.

Duit, R. & Mikelskis-Seifert, S. (2007). Kontextorientierter Unterricht. Wie man es einbettet, so wird es gelernt. Naturwissenschaften im Unterricht. Physik, 18 (98), 4–8.

Guderian, P. (2007). Wirksamkeitsanalyse außerschulischer Lernorte. Der Einfluss mehrmaliger Besuche eines Schülerlabors auf die Entwicklung des Interesses an Physik. Dissertation, Berlin 2007. Verfügbar unter: (<http://edoc.hu-berlin.de/dissertationen/guderian-pascal-2007-02-12/PDF/guderian.pdf> (24.10.2013)).

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Hrsg.), Intrinsic and extrinsic motivation (S. 309–339). San Diego: Academic Press.

Hoffmann, L., Häußler, P. & Lehrke, M. (1998). Die IPN-Interessenstudie Physik. Kiel: IPN.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E. & Stanat, P. (2002). PISA 2000. Dokumentation der Erhebungsinstrumente. Berlin: Max-Planck-Institut für Bildungsforschung.

Parchmann, I. & Gräsel, C. (2004). Chemie im Kontext – ein Weg zur Verbesserung von (Chemie)Unterricht? Pädforum: unterrichten – erziehen, 32 (5), 278–280.

Pawek, C. (2012). Schülerlabore als nachhaltig das Interesse fördernde außerschulische Lernumgebungen. In D. Brovelli, K. Fuchs, R. von Niederhäusern & A. Rempfler (Hrsg.), Kompetenzentwicklung an Ausserschulischen Lernorten. Tagungsband zur 2. Tagung Ausserschulische Lernorte der PHZ Luzern vom 24. September 2011 (S. 69–94). Münster, Wien, Zürich: LIT.

PISA-Konsortium. (2008). PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich. Münster, New York, NY, München, Berlin: Waxmann.

Rakoczy, K., Buff, A. & Lipowsky, F. (Hrsg.). (2005). Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie. "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis". 1. Befragungsinstrumente. Frankfurt, Main: GFPP u.a.

Wittwer, J., Saß, S. & Prenzel, M. (2008). Naturwissenschaftliche Kompetenz im internationalen Vergleich: Testkonzeption und Ergebnisse. In: W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther. (Hrsg.). TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenz von Grundschulkindern in Deutschland im internationalen Vergleich. (S. 87–124). Münster: Waxmann.

Montag,  
03.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G150



## **Lehrerfeedback zur Aufgabenbearbeitung: Einfluss von Schwierigkeitseinschätzung und Unterrichtserfahrung**

Theoretischer Hintergrund und Fragestellung: Während der Aufgabenbearbeitung richtet der Lerner seine Aufmerksamkeit auf Aspekte des Selbst und seiner Umwelt, die notwendig sind, um die einzelnen Aufgabenlösungsschritte zu vollziehen (Achtziger & Gollwitzer, 2006, 2009; Heckhausen & Gollwitzer, 1987). Entsprechend sollte geeignetes Lehrerfeedback kognitive und metakognitive Informationen enthalten, die dem Lerner das selbstständige Weiterkommen ermöglichen (Butler & Winne, 1995). Insbesondere bei schwierigen Aufgaben sollte dabei Feedback vermieden werden, das die externe kognitive Belastung der Lernenden erhöht ohne bearbeitungsrelevante Hinweise zu liefern (Sweller, 2005). Unterrichtsskripts von Lehrenden sollten folglich auch Wissen darüber enthalten, welche Rückmeldungen zur Unterstützung der Aufgabenbearbeitung geeignet beziehungsweise ungeeignet sind (Blömeke, Eichler, & Müller, 2003; Bromme, 1997). Vor diesem theoretischen Hintergrund stellt sich die Frage, mit welchen Feedbackinhalten Lehrende Schüler bei der Bearbeitung von Aufgaben unterstützen, die sie als schwierig einschätzen, und welche Rolle ihre Unterrichtserfahrung dabei spielt.

Methode: Zur Beantwortung dieser Frage wurden 60 Lehrenden (30 unterrichtserfahrene und 30 unterrichtsunerfahrene Teilnehmer) Videovignetten von Fünftklässlerinnen bei der selbstständigen Aufgabenbearbeitung gezeigt. Die Teilnehmer erhielten die Instruktion, die Filme immer dann zu stoppen, wenn sie der gezeigten Schülerin Rückmeldung geben wollten. Zusätzlich wurde erfasst, wie schwierig die Teilnehmer die bearbeiteten Aufgaben für Schüler ihrer fünften Klasse einschätzten (McElvany et al., 2009). Die Äußerungen der Lehrenden wurden mit einer Videokamera aufgezeichnet und mittels einer qualitativen Inhaltsanalyse ausgewertet (Mayring, 2010). Zur Unterscheidung zwischen geeignetem und ungeeignetem Feedback fand dazu, auf Basis des generierten Kodierleitfadens, die Analyse der Lernprozesse der gezeigten Schülerinnen statt. Die so gewonnenen qualitativen Daten wurden in Regressionsmodellen auf ihren Zusammenhang mit der Schwierigkeitseinschätzung und der Unterrichtserfahrung geprüft.

Ergebnisse: Die Ergebnisse zeigen, dass die unterrichtserfahrene Teilnehmer mehr ungeeignetes Feedback gaben ( $M = 12.37$ ,  $SD = 7.61$ ), als ihre unterrichtsunerfahrenen Kollegen ( $M = 6.80$ ,  $SD = 4.97$ ,  $t(49.93) = 3.35$ ,  $p = .002$ ,  $d = .87$ ). Zudem konnten beim Einbezug des Interaktionsterms Zugehörigkeit zur Gruppe (unterrichtserfahren / unterrichtsunerfahren) x Schwierigkeitseinschätzung ( $\beta = .26$ ,  $t(59) = 2.16$ ,  $p = .035$ ) rund 30 Prozent der Varianz innerhalb des ungeeigneten Feedbacks aufgeklärt werden ( $R^2 = .31$ ,  $\text{adj } R^2 = .28$ ,  $F(3, 56) = 8.56$ ,  $p < .001$ ). Entsprechend wurde deutlich, dass besonders in der Gruppe der unterrichtserfahrenen Teilnehmer die Einschätzung der Aufgabenschwierigkeit einen Einfluss auf das phasenungeeignete Feedback hatte, was in der Gruppe der unterrichtsunerfahrenen Teilnehmer nicht der Fall war. Vor allem bei unterrichtserfahrenen Lehrenden nahm das ungeeignete Feedback somit ab, wenn sie die Schwierigkeit der Aufgaben höher einschätzten. Die Ergebnisse lassen vermuten, dass insbesondere unterrichtserfahrene Lehrende den Bearbeitungsprozess adäquater unterstützen könnten, wenn sie in einem Training über geeignetes und ungeeignetes Feedback zur Aufgabenbearbeitung informiert würden.

- Achtziger, A., & Gollwitzer, P. M. (2009). Rubikonmodell der Handlungsphasen. In V. Brandstätter, & J. H. Otto (Hrsg.), *Handbuch der Allgemeinen Psychologie: Motivation und Emotion* (S. 150-156). Göttingen: Hogrefe.
- Achtziger, A., & Gollwitzer, P.M. (2006). Motivation und Volition im Handlungsverlauf. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3., überarbeitete und aktualisierte Auflage, S. 277-302). Heidelberg: Springer.
- Blömeke, S., Eichler, D., & Müller, C. (2003). Rekonstruktion kognitiver Strukturen von Lehrpersonen als Herausforderung für die empirische Unterrichtsforschung. *Theoretische und methodologische Überlegungen zu Chancen und Grenzen von Videostudien. Unterrichtswissenschaft*, 31 (2), 103-121.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Psychologie des Unterrichts und der Schule* (S. 177-212). Göttingen, Bern, Toronto, Seattle: Hogrefe Verlag für Psychologie.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 245-281.
- Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, 11(2), 101-120.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (11., aktualisierte und überarbeitete Auflage). Weinheim, Basel: Beltz.
- McElvany, N., Schroeder, S., Richter, T., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerfähigkeiten und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für Pädagogische Psychologie*, 3/4, 223-235.
- Sweller, J. (2005). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 19 -30), Cambridge University Press: New York.





## Was macht Schulen in ungünstig segregierter Lage erfolgreich?

### Eine fallkontrastive Analyse der Unterrichtsqualität an erfolgreichen und weniger erfolgreichen Schulen

Seit einigen Jahren sind Schulen in sozioökonomisch ungünstig segregierten Stadtteilen sowohl international als auch zunehmend in Deutschland in den Fokus der Aufmerksamkeit gerückt, u.a. im Zusammenhang mit der Diskussion um Chancengerechtigkeit und Steuerung im Schulsystem. Aufgrund ihrer Lage kumulieren an diesen Schulen Belastungen, die eine günstige Lern- und Leistungsentwicklung erschweren, sodass diese Schulen bei Vergleichen oftmals schlechter abschneiden als Schulen in günstigeren Lagen (vgl. z.B. Baumert et al., 2006; van Ackeren, 2008; Harris, 2010; Ainscow & Goldrick, 2012; Arnz, 2012). Gleichwohl zeigen angloamerikanische und britische Studien wie auch PISA 2009 (OECD, 2011), dass Schulen trotz der kumulativ auf die Bildungskarrieren wirkenden Risiken einen signifikanten Zugewinn auf den Ebenen des Lernens und der Leistung der Schüler/innen erreichen können. Dabei erweist sich u.a. eine Fokussierung auf das unterrichtliche Lehren und Lernen als zentral (vgl. z.B. Maden, 2001; Muijs et al., 2004). Um guten Unterricht zu beschreiben haben sich in Deutschland auf Grundlage empirischer Befunde zehn Unterrichtsqualitätsmerkmale von Helmke (2012) etabliert, international ist insbesondere auf Hatties Metaanalyse (2009) hinzuweisen. Offen ist jedoch, inwiefern diese Qualitätskriterien auch für erfolgreiche Schulen in deprivierten Kontexten entscheidend sind. Zwar sind in der internationalen Forschung Hinweise auf unterrichtsbezogene Merkmale erfolgreicher Schulen in besonders herausfordernder Lage benannt, jedoch bleibt u.a. aufgrund divergierender systemischer Rahmenvorgaben und kultureller Kontextbedingungen offen, inwiefern diese Befunde auf deutsche Schulen in ungünstig segregierten Quartieren übertragbar sind. Vor diesem Hintergrund verfolgt dieses Vorhaben das Ziel, Faktorenkonstellationen und Prozessmerkmale von Unterricht an Schulen in schwieriger Lage zu erfassen und hinsichtlich ihrer Adaptivität bezüglich der spezifischen Rahmenbedingungen zu analysieren. Es ergibt sich folgende übergeordnete Forschungsfrage: *Wie sind unterrichtliche Lehr-Lernprozesse gestaltet, um den kontextbedingt ungünstigen Lern- und Leistungsvoraussetzungen Schüler/innen adäquat zu begegnen?* Der Untersuchung liegt ein Fallstudienansatz zugrunde, bei dem ausgehend von einem theoretischen Sampling (vgl. Glaser & Strauss, 1967) ein fallkontrastiver Zugang gewählt wurde. Basierend auf den Ergebnissen in Lernstand 8 wurden acht Fallschulen unterschiedlicher Schulformen identifiziert, die alle mit ungünstigen Bedingungen starten, sich jedoch hinsichtlich ihrer Ergebnisse unterscheiden, so dass zwei Gruppen von Schulen gebildet wurden: (1) Vier Schulen bilden die Gruppe der erwartungsgemäßen (leistungsschwachen) Schulen (2) Vier Schulen bilden die Gruppe der erwartungswidrig guten Schulen. Die Datenerhebung folgte einem Mixed-Methods-Design: Es erfolgte eine schriftliche Befragung der Schülerschaft (N= 2232) und der Lehrkräfte (N= 196) zur Beurteilung des Unterrichts bzw. der eigenen Unterrichtspraxis sowie bei den Lehrkräften zudem zur subjektiven Bedeutung verschiedener Unterrichtsmerkmale. Ergänzend wurden In-vivo-Unterrichtsbeobachtungen (N=123) durchgeführt. Das Rating-Verfahren erfolgt über eine erweiterte Fassung des hoch-inferenten Unterrichtsbeurteilungsinstruments der Schulinspektion Hamburg. Zudem fanden vertiefende semi-strukturierte Interviews mit den verschiedenen Akteuren statt, die inhaltsanalytisch ausgewertet werden. Abschließend erfolgt eine Triangulation (Flick, 2011) sowie eine Gegenüberstellung der Ergebnisse mit publizierten Vergleichswerten. Die Ergebnisse der Schülerbefragung zeigen im Gruppenvergleich einen signifikanten Unterschied in den Bereichen *Störungsfreiheit und Disziplin* sowie *soziale- und kognitive Aktivierung*, zu Gunsten der erwartungswidrig guten Schulen. Insgesamt sind die Standardabweichungen des Gruppenvergleichs relativ hoch, auch bei einer Betrachtung auf Einzelschuleben, was auf einen hohen Einfluss der einzelnen Lehrkräfte hinweist. Dementsprechend zeigen die Ergebnisse der Lehrerbefragung eine hohe Varianz der Einschätzungen auch innerhalb der einzelnen Schulen, obgleich sich in erster Linie Schulformunterschiede offenbaren. Unterschiede zwischen den untersuchten Schulgruppen, lassen sich jedoch für die Bereiche *Klarheit und Strukturiertheit* finden. Hinsichtlich der Frage nach der Wichtigkeit der verschiedenen Kennzeichen ‚guten Unterrichts‘ für Schulen in segregierter Lage, ist festzuhalten, dass alle auf einem hohem Niveau angesiedelt werden. Zwischen den beiden untersuchten Gruppen sind kaum konsistente Unterschiede festzustellen. Die Auswertungen der Interviews und Unterrichtsbeobachtungen stehen noch aus. Letztgenannte Ergebnisse sollen auf der zweiten Tagung der GEBF vorgestellt und schließlich die verschiedenen Akteursperspektiven verbunden werden.

Ackeren, I. van (2008). Schulentwicklung in benachteiligten Regionen. Eine exemplarische Bestandsaufnahme von Forschungsbefunden und Steuerungsstrategien. In W. Lohfeld, & F. Hamburger (Hrsg.), Gute Schulen in schlechter Gesellschaft. Wiesbaden: Verlag für Sozialwissenschaften, S.47-58.

Ainscow, M. & Goldrick, S. (2010). Making sure every child matter: enhancing equity within education systems. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (Hrsg.), Second international handbook of educational change (pp. 869-882). London: Springer.

Arnz, S. (2012). Turnaround von Schulen in kritischer Lage. Erfolg ermöglichen und organisieren. In S.G. Huber (Hrsg.), Schulverwaltung spezial (2), S. 17,f.

Baumert, J., Stanat, P. & Watermann, R. (2006). Herkunftsbedingte Disparitäten im Bildungswesen. Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften.

Flick, U.(2011). Triangulation – Eine Einführung. 2. Aufl., Wiesbaden: Verlag für Sozialwissenschaften.

Glaser, B. G. & Strauss, A. L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, Aldine Publishing Company.

OECD (2011). PISA 2009 Ergebnisse: Potenziale nutzen und Chancengerechtigkeit sichern – Sozialer Hintergrund und Schülerleistungen (Band II). <http://dx.doi.org/10.1787/9789264095359-de> [Zugriff: 08.10.2013].

Harris, A. (2010). Improving Schools in Challenging Contexts. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Hrsg.), Second international handbook of educational change (pp. 693-706). London: Springer.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (4. überarbeitete Aufl., Schule weiterentwickeln – Unterricht verbessern. Orientierungsband). Seelze: Klett-Kallmeyer.

Hattie, J. A. C. (2009). Visible Learning. A synthesis of over 800 meta-analyses relating to achievement. London & New York: Routledge.

Maden, M. (Ed.). (2001). Success against the odds: Five years on. London: Routledge.

Muijs, D., Harris, A., Chapman, C., Stoll, L. & Russ, J. (2004). Improving Schools in Socio-economically Disadvantaged Areas. A Review of the Research Evidence. School Effectiveness and School Improvement, 15(2), 149-175.





Chair(s): Tina Hascher (Universität Bern), Alexander Gröschner (TU München), Karl-Heinz Arnold (Universität Hildesheim)

Diskutant/in: Hans Gruber (Universität Regensburg)

Die Bedeutung von Praktika für die Ausbildung von Lehrerinnen und Lehrern steht weitgehend außer Zweifel, obschon empirische Befunde, die ihre Wirksamkeit für den Erwerb professioneller Handlungskompetenz für den Lehrerberuf belegen, fehlen (z.B. Arnold, Gröschner & Hascher, in Druck). Insbesondere mangelt es an Studien, die versuchen, die Qualität in Praktika differenziert zu beleuchten. Das Symposium versammelt je zwei Beiträge aus zwei Ländern (Deutschland, Schweiz), die aus verschiedenen Perspektiven (Studierende, Mentor/inn/en) und Fachrichtungen (Erziehungswissenschaft, Psychologie) der Frage nach Qualitätskriterien unterschiedlicher Formen der schulpraktischen Ausbildung (eher kurze, längere Praktika; verschiedene Betreuungsmodelle) explizit nachgehen und damit einen Beitrag zur empirischen Untersuchung von Lernprozessen in praxisbezogenen Lerngelegenheiten (im Primar- und Sekundarschulbereich) leisten. Die Breite der Methoden reicht von quantitativen und qualitativen Zugängen hin zu Interventionsstudien.

Der erste Beitrag von Lea Beyeler und Tina Hascher beleuchtet die Sicht von Studierenden und Mentor/inn/en in Bezug auf die Kriterien eines guten Praktikums anhand von Tiefeninterviews. Die Ergebnisse verdeutlichen beispielsweise, dass Studierende eher enge Vorstellungen von einem guten Praktikum aufweisen, da sie seine Qualität tendenziell mit einer qualitativ hochwertigen Praktikumsbegleitung gleichsetzen. Die Lehrpersonen weisen ein deutlich elaborierteres Verständnis auf.

Im zweiten Beitrag berichten Myriam Schlag und Margarete Imhof von einer Studie, bei der zwei Formen der Praktikumsdokumentation miteinander verglichen wurden. Die Befunde zeigen, dass eine strukturierte Form der Praktikumsdokumentation durch ein Arbeitsheft zu einem signifikanten Anstieg der selbsteingeschätzten Kompetenzen führt. Über einen längeren Zeitraum werden insbesondere praktikumsbegleitende Aufgaben eher bearbeitet und akzeptiert als durch offene Formen der Praktikumsreflexion.

Alexander Gröschner und Marc Kleinknecht stellen im dritten Beitrag ein internetgestütztes Lerntool vor, das die individuelle Förderung der Selbstreflexion durch Videofeedback zum Ziel hat. Im Rahmen des Projekts „V-Reflect“ lernen Studierende, eigenen Unterricht zu reflektieren und gemeinsam mit Kommiliton/inn/en über Unterricht zu kommunizieren. Der Beitrag stellt Befunde zur Zufriedenheit und Motivation der Studierende in der Nutzung dieser Lerngelegenheit vor.

Der vierte Beitrag von Annelies Kreis thematisiert die Frage, welche Aspekte im Praktikum sich durch das fachspezifische Coaching verbessern lassen. Die Ergebnisse verdeutlichen einerseits Qualitätssteigerungen hinsichtlich der Form der Lernunterstützung durch die Mentor/inn/en und des Lernertrags auf Seiten der Praktikant/inn/en, wohingegen die Überzeugungen der Studierenden (differenziert in transmissions- versus kognitiv-konstruktivistisch orientierte Überzeugungen) stabil bleiben.

Montag,

03.03.2014,

11:15 - 13:00,

IG-Farben-Haus

Raum 0.251



## **Gute Praktika aus der Sicht von Studierenden und Praxislehrpersonen**

Praktika kommt in der Ausbildung von Lehrerinnen und Lehrern ein grosser Stellenwert zu. Studierende, aber auch Lehrpersonen beurteilen Praktika fast uneingeschränkt positiv und messen ihnen eine sehr wichtige Rolle für die Entwicklung von Lehrkompetenzen bei. Dem allgemeinen Ruf nach mehr Praktika kann zwar entgegengestellt werden, dass nicht die Quantität, sondern die Qualität der Praktika entscheidend ist (z.B. Gröschner & Seidel, 2012; Hascher, 2012; Zeichner, 1990). Bisher jedoch fehlen Studien, die empirisch belegen, welche Qualitätskriterien erreicht werden sollten. Eine Ausnahme stellt die Studie von Beck und Kosnik (2002) dar. Sie identifizierten sieben Hauptkriterien für ein gutes Praktikum aus der Sicht von Studierenden: emotionale Unterstützung durch die Lehrperson, eine kollegiale Beziehung mit der Praxislehrperson, eine kollaborative Zusammenarbeit mit der Praxislehrperson, die Gewährung von Freiräumen und Flexibilität, ein qualitativ hochwertiges, konstruktives und regelmässiges Feedback, eine hohe professionelle Kompetenz der Praxislehrperson sowie eine intensive, aber nicht übertriebene Arbeitsbelastung. In Bezug auf diese allgemeinen Qualitätskriterien gilt es, spezifischere Realisierungsformen zu identifizieren: Wie kommen diese Qualitätskriterien zum Ausdruck und auf welche Bereiche wirken sie? Die von Beck und Kosnik (2002) befragten Studierenden erläuterten beispielsweise, dass Freundlichkeit und eine warme Atmosphäre emotional unterstützend sind und entsprechend wichtige Komponenten eines guten Praktikums darstellen. Aufgrund der Heterogenität der schulpraktischen Ausbildungen muss gefragt werden, inwiefern sich die Kriterien von Beck und Kosnik (2002) bestätigen lassen. Welche Aspekte gelten als Qualitätskriterien für Studierende und Praxislehrpersonen, und inwieweit stimmen die beiden Perspektiven überein? Dazu bedarf es einer vertiefenden, detaillierten Sicht auf die Prozesse in den Praktika und deren Qualitätskriterien (z. B. Allen & Wright, 2013). Für eine erste Annäherung an diese Fragen wurden 19 leitfadengestützte Tiefeninterviews durchgeführt. Die elf Studierenden im letzten Ausbildungsjahr sowie die acht Praxislehrpersonen (insgesamt fünf Tandems) wurden zu ihren Erfahrungen im kurz zuvor stattgefundenen Praktikum befragt. Es wurde erhoben, welche Kriterien ein gutes Praktikum, eine gute Praxislehrperson sowie eine gute Praktikantin/einen guten Praktikanten auszeichnen. Die Ergebnisse zeigen, dass Studierende ein gutes Praktikum tendenziell mit einer qualitativ hochwertigen Praktikumsbegleitung gleichsetzen. Darunter verstehen die Studierenden vor allem eine konstruktive Nachbesprechung, aber auch eine offene Haltung der Praxislehrperson gegenüber Ideen von Seiten der Studierenden sowie die Möglichkeit, sich mit Fragen an die Praxislehrperson wenden zu können. Der eigene Beitrag zu einem guten Praktikum (wie z.B. eine sorgfältige Vorbereitung auf den Unterricht oder eine motivierte Haltung) wird indes meistens erst bei der konkreten Nachfrage nach einem guten Praktikanten/einer guten Praktikantin erwähnt. Lehrpersonen zeigen neben der Nennung von Kriterien auf Studierendenseite (wie z.B. Zuverlässigkeit und eine lernoffene Haltung) vielseitigere Punkte auf wie die Möglichkeit, weitere Erfahrungen im schulischen Umfeld sammeln zu können oder die Einbindung ins Lehrerkollegium. Zudem nehmen Praxislehrpersonen ihre eigene Rolle für ein gutes Praktikum eher in den Fokus als die Studierenden. Die Kriterien von Beck und Kosnik (2002) werden in dieser Befragung prinzipiell bestätigt. Allerdings lassen sich die sieben Kriterien für ein gutes Praktikum nicht trennscharf voneinander abgrenzen und es erscheint notwendig, sie stärker in Bezug zueinander zu setzen. Zum Beispiel wird deutlich, dass aus der Sicht der Studierenden die Ausgewogenheit zwischen den beiden Kriterien gewährte Freiräume und Unterstützung durch die Praxislehrperson sehr wichtig ist. Zu grosse Freiheit, aber auch zu viel Unterstützung werden jeweils negativ beurteilt. Zu beachten gilt des Weiteren, dass sich die Umsetzung der Qualitätskriterien im Verlauf des Praktikums verändert, beispielsweise indem die Unterstützung im Verlauf des Praktikums je nach Voraussetzungen und Entwicklung der Studierenden zunimmt oder verringert wird, oder verschiedene Formen der Kollaboration praktiziert werden wie Hospitation oder Teamteaching. Diese Ergebnisse werden in Hinblick auf weitere Studien zur Qualitätsmessung von Praktika diskutiert.

Allen, J. M. & Wright, S. E. (2013). Integrating theory and practice in the pre-service teacher education practicum. *Teachers and Teaching*, (ahead-of-print), 1-16.

Beck, C. & Kosnik, C. (2002). Components of a Good Practicum Placement: Student Teacher Perceptions. *Teacher Education Quarterly*, 81-98.

Gröschner, A. & Seidel, T. (2012). Lernbegleitung im Praktikum – Befunde und Innovationen im Kontext der Reform der Lehrerbildung. In W. Schubarth, K. Speck, A. Seidel, C. Gottmann, C. Kamm & M. Krohn (Hrsg.), *Studium nach Bologna: Praxisbezüge stärken?! Praktika als Brücke zwischen Hochschule und Arbeitsmarkt* (S. 170-183). Wiesbaden: Springer.

Hascher, T. (2012). Lernfeld Praktikum – Evidenzbasierte Entwicklungen in der Lehrer/innenbildung. *Zeitschrift für Bildungsforschung* 2, 109-129.

Zeichner, K. M. (1990). Changing Directions in the Practicum: looking ahead to the 1990s. *British Journal of Teacher Education*, 16(2), 105-132.

## Kompetenzentwicklung im Schulpraktikum durch verschiedene Formen von Praktikumsdokumentationen

Forschungsergebnisse zur Kompetenzentwicklung während der schulpraktischen Phasen haben gezeigt, dass die selbsteingeschätzten Kompetenzen der Lehramtsstudierenden in diesen Abschnitten der Ausbildung meist ansteigen (Bach, 2013; Schubarth et al., 2012). Als wichtige Einflussfaktoren wurden dazu bisher die Betreuung durch die Mentor/inn/en in der Schule und das Stresserleben identifiziert (Bach, 2013). Die Effekte von unterschiedlichen Formen der Praktikumsdokumentationen in der gleichen Art von Schulpraktikum sind jedoch bis jetzt noch nicht in den Blick genommen worden. Ein zentrales Unterscheidungskriterium von Dokumentationsformen ist die offene im Vergleich zur gelenkten Beobachtung und Reflexion. Während bei offenen Aufgaben- und Dokumentationsformen nur sehr allgemeine Fragen gestellt werden, werden bei gelenkter Beobachtung sehr spezifische Situationen angesprochen. Die Führung der Studierenden in den Praktikumsphasen könnte den Vorteil haben, den Blick für die Vorgänge im Unterricht zu fokussieren (vgl. Seidel, Blomberg, & Stürmer, 2010). Damit stellt sich die Frage, inwiefern gelenkte im Vergleich zu offenen Formen der Praktikumsdokumentation zu unterschiedlichen Veränderungen in der selbsteingeschätzten Kompetenz der Studierenden führen. Um dieser Frage nachzugehen werden in dieser Studie zwei Formen von Praktikumsdokumentationen verglichen: offene Reflexionsaufgaben und ein strukturiertes Portfolio (Arbeitsheft). Die offenen Reflexionsaufgaben sind zu 2 Seiten zu bearbeiten und stellen dabei wenig Verbindung zu den Inhalten der universitären Lehre dar. Das Arbeitsheft bezieht sich auf universitäre Lehrinhalte und beinhaltet sowohl konkrete Beobachtungsaufträge, Materialsammlungen als auch Vorüberlegungen und Reflexionen zum Praktikum. An der Untersuchung nahmen  $n = 66$  Lehramtsstudierende in der Experimentalgruppe und  $n = 97$  in der Referenzgruppe der Universität Mainz teil. Die Studierenden der Experimentalgruppe füllten den Online-Fragebogen zu 3 Messzeitpunkten (vor, nach und 1,5 Monate nach dem Praktikum) aus, die Referenzgruppen 1,5 Monate nach dem Praktikum. Das Praktikum war ein zweiwöchiges Hospitationspraktikum in der vorlesungsfreien Zeit und wurde von den Studierenden in der Regel nach dem zweiten Semester belegt. Die Studierenden der Experimentalgruppe bearbeiteten das Arbeitsheft, die Referenzgruppe die offenen Reflexionsaufgaben. Die Onlinefragebögen beinhalteten neben den Rahmenbedingungen zum Schulpraktikum die Skalen zur Erfassung von Kompetenzen in der Lehrerbildung (Gröschner & Schmitt, 2012), zur Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999) und zur Akzeptanz der Praktikumsdokumentation. Die Akzeptanz wurde dabei mit einer eigenen Skala erhoben, welche aus sechs Items besteht. Die Ergebnisse zeigen einen signifikanten Anstieg der selbsteingeschätzten Kompetenz über die 3 Messzeitpunkte in der Experimentalgruppe ( $F(1.66, 108.19) = 20.26, p < .05$ ). Im Follow-up 1,5 Monate nach dem Schulpraktikum unterschieden sich die Experimental und die Referenzgruppe in der selbsteingeschätzten Kompetenz jedoch nicht signifikant voneinander. Auch die Selbstwirksamkeitserwartung stieg in der Experimentalgruppe signifikant über die 3 Messzeitpunkte an ( $F(1.98, 128.40) = 11.90, p < .05$ ), unterschied sich zum letzten Messzeitpunkt jedoch nicht signifikant von der Referenzgruppe. Im Gegensatz dazu war die Akzeptanz der Praktikumsdokumentation für die Form des Arbeitshefts signifikant höher ( $t(161) = 11.86, p < .05$ ). Dabei zeigten die Studierenden der Experimentalgruppe, die mit dem Arbeitsheft arbeiteten, eine deutlich höhere Akzeptanz für die Aufgaben als die Referenzgruppe, die offene Reflexionsaufgaben bearbeiteten. Die bisherigen Ergebnisse des Forschungsprojekts werden durch strukturierte Interviews ergänzt. Dazu werden die Daten von  $n = 10$  Studierenden der Experimentalgruppe und  $n = 10$  Lehramtsstudierende der Kontrollgruppe, welche der Referenzgruppe 2 entspricht, erhoben. Ziel der Datenerhebung ist es, den Erstellungs- und Bearbeitungsprozess sowie den Inhalt der Praktikumsdokumentationen genauer zu analysieren. Außerdem werden die Praktikumsbetreuung durch Mentor/inn/en und die selbsteingeschätzte Nützlichkeit der Praktikumsaufgaben durch die Studierenden erfasst.

Bach, A. (2013). Kompetenzentwicklung im Schulpraktikum. Ausmaß und zeitliche Stabilität von Lerneffekten hochschulischer Praxisphasen. Münster: Waxmann.

Gröschner, A. & Schmitt, C. (2012). Kompetenzentwicklung im Praktikum? Entwicklung eines Instruments zur Erfassung von Kompetenzeinschätzungen und Ergebnisse einer Befragung von Lehramtsstudierenden im betreuten Blockpraktikum. *Lehrerbildung auf dem Prüfstand*, 5 (2), 112–128.

Schubarth, W., Speck, K., Seidel, A., Gottmann, C., Kamm, C., Kleinfeld, M. et al. (2012). Kompetenzentwicklung im Praxissemester: Ergebnisse einer Längsschnittanalyse zum "Potsdamer Modell der Lehrerbildung". In T. Hascher & G. H. Neuweg (Hrsg.), *Forschung zur (Wirksamkeit der) Lehrerinnenbildung* (Bd. 8, 1. Aufl., S. 201–220). Münster, Westf: LIT.

Schwarzer, R. & Jerusalem, M. (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. *Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*, Freie Universität Berlin. Zugriff am 16.10.2012. Verfügbar unter <http://userpage.fu-berlin.de/health/germscal.htm>

Seidel, T., Blomberg, G. & Stürmer, K. (2010). Observer - Validierung eines Videobasierten Instruments zur Erfassung der professionellen Wahrnehmung im Unterricht. *Zeitschrift für Pädagogik*, 56 (Beiheft 56), 296–306.



## ***Selbstreflexion durch Videofeedback im Lehramtsstudium – Eine Intervention im Rahmen des studienbegleitenden Praktikums***

Praktika nehmen in der Lehrerbildung eine bedeutsame Rolle ein und dienen insbesondere der Überprüfung der eigenen professionellen Eignung (u.a. Hascher, 2012). Zur Unterstützung dieses Ziels werden zunehmend Bemühungen um eine qualitativ hochwertige universitäre Lernbegleitung unternommen, die sich als relevant für den wahrgenommenen Lernertrag erwiesen hat (Gröschner, Schmitt, & Seidel, 2013). Einen zentralen Aspekt, zu dem Studierende im Praktikum reflektieren sollten, stellen Erfahrungen eigenen unterrichtlichen Handelns dar. Im Projekt „V-Reflect“ werden mit Video aufgezeichnete, von den Studierenden ausgewählte eigene Unterrichtssequenzen anhand einer internetbasierten Lernplattform im Rahmen eines praktikumsbegleitenden Seminars diskutiert und reflektiert. Die Form des „Videofeedbacks“ wird für die Förderung von Fähigkeiten zur Selbstreflexion als wirksames Mittel betrachtet (Fukkink, Trienekens, & Kramer, 2011). Allerdings ist bisher unklar, welchen Mehrwert die internetgestützte, videobasierte Selbstreflexion gegenüber anderen Formen der fallbasierten Reflexion hat (z.B. mittels verschriftlichter Erfahrungsberichten, die in Präsenzveranstaltungen ausgetauscht werden). Im Vortrag werden empirische Befunde zur Zufriedenheit und Motivation der neu eingeführten Möglichkeit des Videofeedback bei einer Gruppe von Studierenden, die am Projekt teilnahmen (IG), im Vergleich zu einer Gruppe, die ein herkömmliches Reflexionsseminar besuchte (KG), vorgestellt.

Den folgenden Fragen wird im Vortrag nachgegangen:

1. Inwieweit schätzen die Studierenden der IG die Zufriedenheit mit dem Seminar und den subjektiven Lernertrag höher ein im Vergleich zur KG?
2. Inwieweit wird die Lernumgebung gemäß der Selbstbestimmungstheorie als positiver erlebt in der IG im Vergleich zur KG und wie verändert sich die Wahrnehmung im Verlauf des Seminars?

Für die Umsetzung des Projekts „V-Reflect“ wurden zwei Seminare zum semesterbegleitenden Praktikum im Masterstudium angeboten, die den gleichen Inhalt (Adaptivität und innere Differenzierung im Unterricht) hatten. Die Studierenden der IG (n = 12) nahmen sich einmal im Unterricht mit Video auf und stellten eine Sequenz zur Diskussion auf der Internetplattform zur Verfügung. Die Studierenden der KG (n = 30) reflektierten ihre Unterrichtserfahrungen im Praktikum mittels schriftlicher Berichte und tauschten sich im Seminar aus. In der Semesterhälfte (t 1) und am Semesterende (t 2) erhielten beide Gruppen einen Fragebogen zur Zufriedenheit mit den reflexionsbezogenen Lernanlässen, dem subjektiv erlebten Lernertrag sowie der Motivation. Darin wurde der subjektive Lernertrag erfasst (4 Items, vierstufig), die Zufriedenheit mit dem Seminarverlauf (11 Items, vierstufig) sowie die Wahrnehmung der Lernumgebung auf der Basis der Selbstbestimmungstheorie nach Deci und Ryan (2002) in den Dimensionen Autonomieunterstützung (5 Items, vierstufig), Kompetenzunterstützung (5 Items, vierstufig) und soziale Eingebundenheit (5 Items, vierstufig). Die Analysen und Gruppenvergleiche erfolgten mittels nichtparametrischer Tests mit SPSS bzw. mit R.

Die Ergebnisse verdeutlichen zunächst, dass beide Gruppen (zu t 1) insgesamt zufrieden mit dem jeweiligen Seminar sind. Im Verlauf der Seminare nehmen die Zufriedenheit und die Wahrnehmung der Lernumgebung insbesondere in der IG zu. Die Befunde werden kritisch vor dem Hintergrund der Anforderungen an Praktika in der Lehrerbildung diskutiert.

Fukkink, R. G., Trienekens, N. & Kramer, L. J. C. (2011). Video Feedback in Education and Training: Putting Learning in the Picture. *Educational Psychology Review*, 23, 45-63. doi: 10.1007/s10648-010-9144-5

Gröschner, A., Schmitt, C. & Seidel, T. (2013). Veränderung subjektiver Kompetenzeinschätzungen von Lehramtsstudierenden im Praxissemester. *Zeitschrift für Pädagogische Psychologie*, 27(1-2), 77-86. doi: 10.1024/1010-0652/a000090

Hascher, T. (2012). Lernfeld Praktikum – Evidenzbasierte Entwicklungen in der Lehrer/innenbildung. *Zeitschrift für Bildungsforschung*, 2(2), 109-129.



## **Überzeugungen zum Unterrichten von Mathematik: Effekte von Fachspezifischem Unterrichtscoaching mit Lehrstudierenden im Praktikum**

Mit dem Fachspezifischen Unterrichtscoaching (Staub, 2004; Staub & Kreis, 2013), das auf einem kognitiv-konstruktivistischen Verständnis von Lehr-Lernprozessen (Reusser, 2001) basiert, sollen unterrichtsbezogene Lernprozesse von Lehrpersonen unterstützt werden. In dialogischen Vor- und Nachbesprechungen von Unterricht erweitern Lehrpersonen respektive Lehrerstudierende situations- und fachspezifisch ihre Überzeugungen, ihr Wissen und ihre Unterrichtskompetenz. Eine Coachingsequenz besteht dabei aus einer gemeinsamen Unterrichtsvorbereitung, der Unterrichtsdurchführung in geteilter Verantwortung und abgesprochener Kooperation sowie einer Nachbesprechung. Besonders betont werden hierbei der Fokus auf fachspezifische und fachdidaktische Wissensbestände sowie eine hohe Dialogizität der Interaktionen. Eine wichtige Rolle spielen Kernperspektiven der Planung und Reflexion von Unterricht, die fachübergreifend eingesetzt werden, jedoch durch den Fokus auf die zu lehrenden Fachinhalte das Gespräch auf fachdidaktische Aspekte lenken. In der Interventionsstudie „Unterrichtsentwicklung durch fachspezifisches Coaching“ (Kreis & Staub, 2011) wurden Praxislehrpersonen (PL) zu Fachspezifischen Unterrichtscoaches für Mathematik (Primarstufe) ausgebildet. Die fachdidaktischen Ausbildungselemente waren auf eine kognitiv-konstruktivistische Gestaltung von Lernumgebungen ausgerichtet. Die PL der Interventionsgruppe coachten sodann während eines siebenwöchigen Praktikums Lehrstudierende für die IG. Studierende der Kontrollgruppe (KG) wurden während des gleichen Praktikums nach den herkömmlichen Vorgaben der Ausbildungsinstitution betreut. Die Studierenden wurden zufällig einer Praxislehrperson der IG oder KG zugeteilt. In einem multi-methodischen Untersuchungsdesign konnte mittels verschiedener Indikatoren dargelegt werden, dass zwischen den beiden Gruppen Unterschiede bezüglich der Form der Lernunterstützung durch die PL sowie des Lernertrags für die Lehrstudierenden bestehen (Kreis & Staub, 2011). Dies deutet auf die Wirksamkeit sowohl der Intervention mit den PL als auch des durch die PL mit Studierenden durchgeführten Fachspezifischen Unterrichtscoachings. Ein weiterer relevanter Indikator für Lernen von Lehrstudierenden in Zusammenhang mit berufspraktischen Lerngelegenheiten sind Veränderungen in deren fachdidaktischen Überzeugungen. So konnten Staub und Stern (2002) zeigen, dass eine kognitiv-konstruktivistische Ausrichtung von Lehrpersonen mit höherem Leistungszuwachs der Lernenden im Problemlösen einhergeht (vgl. auch Hill, Rowan & Ball, 2005). In einer Teilstudie der Interventionsstudie wurden deshalb Studierende der IG (N = 55) und KG (N = 189) vor und im Anschluss an das Praktikum mit einem Fragebogen hinsichtlich von Überzeugungen zum Unterrichten von Mathematik gefragt. Es wurden 6 Skalen eines in PISA 2003 eingesetzten Fragebogens für Lehrpersonen (Prenzel et al., 2006) verwendet, die sich am Instrument von Staub und Stern (2002) orientieren. Drei Skalen (vierstufige Zustimmung) umfassen Aussagen zu transmissionsorientierten Überzeugungen über Mathematikunterricht (Selbstständiges und verständnisvolles diskursives Lernen, 12 Items,  $\alpha = .78$ ; Vertrauen auf mathematische Selbstständigkeit der Schüler/-innen, 5 Items,  $\alpha = .63$ ; in PISA .84; Kognitiv aktivierende Aufgaben, 8 Items,  $\alpha = .72$ ). Drei Skalen erfassen kognitiv-konstruktivistisch orientierte Überzeugungen (Rezeptives Lernen durch Beispiele und Vormachen, 12 Items,  $\alpha = .80$ ; Einschleifen von technischem Wissen, 4 Items,  $\alpha = .53$ , in PISA .66); Eindeutigkeit des Lösungswegs, 2 Items,  $\alpha = .71$ ). Untersucht wird die Frage nach Ausprägungen in den fachdidaktischen Überzeugungen der Studierenden sowie nach Unterschieden zwischen IG und KG sowie den beiden Erhebungszeitpunkten. Die Skalen zu kognitiv-konstruktivistischen Überzeugungen erfahren zu beiden Erhebungszeitpunkten eine höhere Zustimmung durch die Lehrstudierenden beider Gruppen (Mittelwerten um 3.5) als die Skalen zu transmissionsorientierten Überzeugungen (Mittelwerte um 2). Zum ersten Erhebungszeitpunkt finden sich keine signifikanten Unterschiede zwischen den beiden Gruppen. Bei einer Skala („Eindeutigkeit des Lösungswegs“) besteht in der IG ein signifikant grösserer Zuwachs der Zustimmung zu t2, was der intendierten Richtung entgegenläuft. In der KG ist die Zustimmung zur Skala „Kognitiv aktivierende Aufgaben“ nach dem Praktikum (t2) signifikant geringer als bei t1. Zwischen den Gruppen und den Erhebungszeitpunkten vor und nach dem Praktikum bestehen somit kaum Unterschiede, was an einem Deckeneffekt liegen könnte. Effekte der Intervention können somit anhand des Indikators „Überzeugungen“ – im Gegensatz zu den Indikatoren „Qualität des erteilten Unterrichts“, „selbstberichtete Lernereignisse“ und „Lernereignisse während Unterrichtsbesprechungen“ (Kreis & Staub, 2011) – kaum nachgewiesen werden.

Brown, J. S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.

Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.

Kreis, A. & Staub, F. C. (2011). Fachspezifisches Unterrichtscoaching im Praktikum – eine quasi-experimentelle Interventionsstudie. *Zeitschrift für Erziehungswissenschaft*, 14(1), 61-83.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J., Schiefele, U., Rolff, H.-G., Ramm, G. (2006). PISA 2003: Dokumentation der Erhebungsinstrumente. Münster: Waxmann.



## ***Kann Bildung die deprivierte sozioökonomische Lage von Geringgebildeten verbessern? Eine Analyse der Wirkung von Bildungsreformen in Großbritannien in den 1980er Jahren***

Chair(s): Martin Neugebauer (Universität Mannheim)

Bildung ist in der Wissensgesellschaft zentral für eine gleichberechtigte Teilhabe in allen Lebensbereichen, wobei der Teilhabe am Arbeitsmarkt aufgrund ihrer weitreichenden Konsequenzen eine besondere Bedeutung zugemessen wird. Empirisch ist zu beobachten, dass geringe Bildung häufig mit ungünstigen Arbeitsmarktpositionen bzw. Ausschluss vom Arbeitsmarkt verbunden ist. Ausgangspunkt dieses Beitrages ist die Frage nach den kausalen Wirkungsmechanismen, welche niedrige Bildung und Arbeitsmarktprobleme miteinander verbinden. Nach dem bildungsökonomischen Erklärungsansatz (z.B. Becker, 1993) führt geringe Bildung zu mangelnder Produktivität und somit zu niedrigen Einkommen und hoher Arbeitslosigkeitgefährdung. Produktivitätssteigerungen in Form von Humankapitalinvestitionen sind somit das primäre Mittel, um die negativen sozioökonomischen Folgen von geringer Bildung zu bekämpfen. Einen gegenläufigen Erklärungsansatz liefern Konflikttheorien, welche davon ausgehen, dass die angestrebten Bildungsrenditen nicht durch Produktivität, sondern durch die kollektive Aushebelung von Marktmechanismen erzielt werden (z.B. Sørensen, 2000). Während soziale Schließung die Bildungsrenditen der höheren Klasse schützt, müssen weniger Privilegierte zum niedrigeren Marktniveau arbeiten. Die sozioökonomische Lage Geringqualifizierter kann daher eher durch Umverteilung der Marktrenditen (z.B. über Mindestlöhne) und Marktzugänge (z.B. Steigerung der Durchlässigkeit des Bildungssystems) als durch reine Humankapitalinvestitionen verbessert werden. Bisher gibt es jedoch nur wenige Studien, die beide Erklärungsansätze vergleichend analysieren. Dieser Beitrag versucht daher die sozioökonomischen Folgen einer exogenen Variation zwischen Herkunft und Bildung mit Hilfe einer quasi-experimentellen Analyse zu bestimmen. Betrachtet werden hierbei Veränderungen in Großbritannien seit den späten 1970er Jahren, wobei die Tatsache genutzt wird, dass Bildungsreformen in Schottland und England/Wales unterschiedlich verlaufen sind. Der herkunftsspezifische Bildungserfolg wird dabei indirekt über die Folgen der Education Acts 1980, 1981 und 1983 (England und Wales) geschätzt, welche die Gesetzesgrundlagen der ersten Vermarktlichungsreformen des englisch-walisischen Bildungssystems bilden. Im Gegensatz implementierte Schottland Bildungsreformen, welche stärker auf Chancengleichheit ausgerichtet waren. Sozial schwächere Kinder dürften in Schottland somit ceteris paribus mehr Bildung erworben haben, wobei die hieraus generierten sozioökonomischen Renditen den Untersuchungsgegenstand bilden. Als Datengrundlage dient die British Cohort Study, eine Vollerhebung aller in einer Woche 1970 in Großbritannien geborenen Personen. Seit der Geburt wurden zunächst die Eltern und anschließend die Kohortenmitglieder in regelmäßigen Abständen befragt, sowie eine Reihe non-kognitiver Tests durchgeführt. Personen, die über den gesamten Untersuchungszeitraum in England und Wales gelebt haben, bilden die kontrafaktische Kontrollgruppe. Personen, die ausschließlich in Schottland lebten, die Experimentalgruppe. Das Quasi-Experiment ist in drei Schritten aufgebaut. Allen Schritten gemein ist die Paarbildung zwischen Individuen der Experimental- und Kontrollgruppe mittels eines kernrichtebasierten Propensity-Score-Matching Verfahrens, welches anhand statistischer Ähnlichkeit in sieben Dimensionen Paare generiert um etwaige Endogenitätsverzerrungen auszubalancieren. Obwohl diese Matching-Indikatoren im frühen Kindheitsalter (vor den Reformen) erhoben wurden, korrelieren sie dennoch deutlich mit dem individuellen Bildungs- und Arbeitsmarkterfolg im Erwachsenenalter. In den ersten beiden Schritten wurden zudem Herkunftsindikatoren aufgenommen, welche im dritten Schritt durch Humankapitalinformationen aus dem Jugend- und Erwachsenenalter ersetzt wurden.

Wie erwartet zeigen sich im ersten Analyseschritt, dass Schotten mit niedriger sozialer Herkunft signifikant mehr Humankapital erwerben als die Kontrollgruppe. Im zweiten Schritt wurde anschließend die Arbeitsmarktposition der Experimental- und Kontrollgruppe verglichen. Hierbei hatten Schotten trotz höherer Bildung weder bessere Berufspositionen noch ein höheres Einkommen als die Kontrollgruppe. Im dritten Schritt zeigte sich, dass Schotten nach der Paarbildung auf Grundlage der vollen Humankapitalausstattung aus einem niedrigeren Elternhaus stammten als die gleichgebildete Kontrollgruppe, was damit einherging, dass sie ein geringeres Einkommen erzielten und in einer niedrigeren Berufsposition arbeiteten. Zusammengefasst widerlegen diese Ergebnisse die humankapitaltheoretische Annahme, dass die negativen sozioökonomischen Konsequenzen niedriger Bildung monokausal auf den Mangel an Bildung und Produktivität zurückzuführen sind. Vielmehr deuten die Analysen daraufhin, dass die angestrebte Bildungsrendite aufgrund sozialer Schließung und Ressourcenhortung nicht am unteren Ende der Bildungsverteilung erzielt werden können. Bildungsinvestitionen erzeugen demnach nur den Anschein die negativen sozioökonomischen Folgen von geringer Bildung bekämpfen zu können, wozu diese allerdings ohne gleichzeitige Umverteilungsmaßnahmen nicht in der Lage sind.

Becker, Gary S. 1993: Human Capital. Chicago: The University of Chicago Press.

Sørensen, Aage B. 2000: Toward a Sounder Basis for Class Analysis, American Journal of Sociology, 105 (6), 1523-1558.



## ***The Causal Effect of Shortening School Instruction Time on Grade Repetitions and Graduation Rates***

Background: When earning a university entrance degree, German students have traditionally spent more time in school than those in other countries. It has been suggested to shorten the length of the high school track (*Gymnasium*) from nine to eight years, leaving the curriculum unchanged. It is suggested that this reform could increase Germany's labour market competitiveness by allowing students to enter the labour market at younger ages. By increasing the labour force by one birth cohort, the reform could reduce skilled worker shortage while reducing the per capita education costs. Eventually, it would help ease the public pension scheme. Since 2001, most German federal states have introduced the suggested school reform. Since the curriculum and the requirements for the *Abitur* were left unchanged, the reform increased the learning intensity for students. Our research investigates whether this so called G8-reform affected grade repetition and *Abitur* graduation rates. Understanding the educational consequences is not only important for individuals' education history, but also for society as a whole as this might inflict unintended costs.

Literature & Research Hypotheses: The move to reduce the time spent at *Gymnasium* is at the centre of debate and there are few conclusive studies on the subject. Previous research analysing the effects of the G8-reform are limited in that they concentrate on single states (e.g. Büttner et al., 2010; Büttner et al., 2011; Vieluf et al., 2012; Meyer et al., 2012). The reform effect cannot reliably be distinguished from time trends or other policy changes implemented at the same time. Additionally, these studies rely on self-collected survey data that may lack representativeness. There are at least two reasons why shortening school time increases repetition rates. First, increased learning intensity imposes higher requirements on students to follow class material. Lower-ability students may need extra time to fulfil the requirements for the university entrance degree and would rather repeat a year. Examining the effects of shortened school instruction time induced by a 1966/67 reform, Pischke (2007) identifies an increase in grade repetition and a decrease in the number of students moving to the academic school track. Second, students might be less reluctant to repeat a class because for long time it has been expected to earn the degree within nine years. Additionally, the increased learning intensity could reduce graduation rates. Jacob und Lefgren (2009) suggest that grade repetition increases the odds of school drop-outs.

Data and Research Design: We employ administrative data from the Federal Statistical Office, Destatis. We use data covering 1990 through 2012 from the "Fachserie 11, Reihe 1 – Allgemein bildende Schulen", which contains the relevant information aggregated by class, gender, school type and federal state. The data cover the universe of all students in Germany in primary and secondary education. Treating the education reform as a quasi-experiment, we exploit the regional and temporal variations in the implementation of this education reform with a flexible difference-in-difference approach (Borjas, 2007). We examine in depth the common trend assumption through graphical inspections of pre-treatment trends in key variables. In addition, we conduct extensive research for potential co-treatments that were implemented at the same time and could moderate our findings.

Results: We find that the shortening of the high school track length increased grade repetition rates by about 2 percentage points, equalling an increase by 50-100 per cent, for students qualifying for university entrance qualification (zero to two years to graduation). This estimated effect also persists when grade level effects, time effects, state effects and further school reforms are taken into account. We also find a small decrease in overall *Abitur* graduation rates.

Borjas, G. J. (2007), *Labour Economics*. Boston: Mcgraw-Hill Higher Education

Büttner, B. & Thomsen, S.L., (2010). Are we spending too many years in school? Causal evidence of the impact of shortening secondary school duration. ZEW Discussion Papers, (10-011).

Büttner, B., Thiel, H. and Thomsen, S.L., (2011). Variation of learning intensity in late adolescence and the impact on noncognitive skills. ZEW Discussion Papers, (11-007).

Jacob, B. A. & Lefgren, L. (2009). The Effect of Grade Retention on High School Completion. *American Economic Journal: Applied Economics* 1(3), 33–58.

Meyer, T., Thiel, H., & Thomsen, S. L. (2012). Effekte des Turbo-Abiturs: Leistungen in Mathematik schlechter, weniger Einschreibungen in den Naturwissenschaften. *niw-info spezial* No. 2/2012.

Pischke, J.-S. (2007). The impact of length of the school year on student performance and earnings: Evidence from the German short school years. *Economic Journal*, 117(523):1216-1242.

Vieluf, U., Ivanov, S., & Nikolova, R. (2011), HANSE - Hamburger Schriften zur Qualität im Bildungswesen: Vol. 10. KESS 10/11: Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. Münster, München, Berlin: Waxmann.



## ***Return and Risk of Human Capital Contracts***

Despite political will to invest more in higher education and to increase the number of graduates, the public financing of tertiary education has come under pressure. In effect, the growing demand for tertiary education is accompanied by increased other social expenditures related to public pensions, health care or early childhood education as well as the financial burden of the recent financial crisis. Increasing private contributions to tertiary education may be one solution to help to alleviate limited public budgets.

By issuing human capital contracts (HCC), private investors gain the right to obtain a share of student's future earnings. They allow students a yearly or monthly amount of money for the duration of their studies. After graduation, students pay a fixed share of their net income to the fund during a fixed amount of time. Depending on the achieved income, students may thus pay back more or less than the amount invested by the fund. The financial return to a HCC depends on the time of reimbursement as well as on the share of net income that is to be paid into the fund. In this paper, we evaluate the return and risk properties that can be expected from human capital contracts.

We start by replicating and complementing the analysis of Palacios-Huerta (2003), who assesses the value-added of including human capital into a financial portfolio using US data. Based on traditional mean-variance spanning tests (Huberman & Kandel, 1987; Bekaert & Urias, 1996), he finds that human capital should be a highly valuable complement to a financial portfolio. While our methodology to evaluate return and risk properties of human capital contracts is consistent with mean-variance spanning tests, we base our analysis on cash-flows investors can realistically receive from human capital contracts.

We use the Mikrocensus data for the years 1995-2009, a yearly representative survey dataset including 1% of all German households. This represents about 380,000 households with 820,000 individuals. In order to assess the financial performance of human capital contracts, we compute the internal rate of return (IRR), abnormal return (alpha), and the risk exposures to the stock and bond markets (betas) of human capital contracts. Our estimates are based on a novel GMM estimator by (Driessen et al., 2012) which allows calculating return and risk properties of non-traded assets using their cash-flows. We simulate the defined human capital contracts for funds covering 50 randomly chosen males and 50 randomly chosen females.

The results of our simulations show that funds of human capital contracts provide low risk exposures to stocks and bonds. As a result, risk-adjusted returns of funds of human capital contracts are significantly positive. The performance of HCC depends on the time of reimbursement and share of net income to be paid into the fund. Even under less favorable conditions of reimbursement from the investor's perspective and taking into account default risk, the performance of HCC is reasonable. Thus, human capital contracts potentially offer large diversification benefits for investors and might be a way to improve the state's educational budget. Besides attracting private investors into the educational sector, issuing human capital contracts may also be an interesting additional form of student support for the state that could improve public educational budgets through its high returns.

Bekaert, G., & Urias, M. S. (1996). Diversification, Integration and Emerging Market Closed-End Funds, *Journal of Finance*, 51, 835-869.

Driessen, J., T.-C. Lin, & Phalippou, L. (2012). A New Method to Estimate Risk and Return of Nontraded Assets from Cash Flows: The Case of Private Equity Funds, *Journal of Financial and Quantitative Analysis*, 47, 511-535.

Huberman, G. & Kandel, S. (1987). Mean-Variance Spanning, *Journal of Finance*, 42, 873-888.

Palacios-Huerta, I. (2003). An Empirical Analysis of the Risk Properties of Human Capital Returns, *American Economic Review*, 93, 948-964.



## *Der Lehrerarbeitsmarkt und die Wahl des Lehrerberufs im zeitlichen Wandel*

Wie haben sich die Arbeitsmarktaussichten von westdeutschen Lehrerinnen und Lehrern in den letzten 30 Jahren gewandelt? Haben sich die Eingangsmerkmale der Lehramtsanwärterkohorten über diesen Zeitraum geändert, möglicherweise in Reaktion auf veränderte Arbeitsmarktaussichten? Diese Fragen, die an die Debatte um die Rekrutierung „geeigneter“ Kandidaten in das Lehramt anknüpfen, stehen im Zentrum des Vortrages. Auf Basis der Konstanzer Studierendensurveys von 1982 bis 2009 wird nachgezeichnet, inwiefern sich Abiturleistungen, ebenso wie fachliche und sicherheitsbezogene Studienwahlmotive im Laufe der letzten 30 Jahre gewandelt haben. Den Lehramtsstudierenden wird eine Vergleichsgruppe von anderen Studierenden gegenübergestellt, um die charakteristischen Merkmale der Lehramtsanwärter aufzuzeigen. Darüber hinaus beschreibt die Studie auf Basis der Mikrozensus von 1980 bis 2009 die Arbeitsmarktaussichten (Einkommen und Arbeitslosigkeitsrisiken) von Lehrern im Vergleich zu anderen akademischen Berufen. Es wird herausgearbeitet, wie ein Wandel in den Arbeitsmarktaussichten Personen mit unterschiedlichen Leistungsvoraussetzungen und Motivlagen anziehen kann. Führen, wie einige Ökonomen annehmen, höhere Lehrereinkommen dazu, dass Personen mit besseren Abiturleistungen ein Lehramt anstreben? Führt die Aussicht auf einen sicheren Arbeitsplatz dazu, dass Personen mit einem ausgeprägten Sicherheitsbedürfnis in das Lehramt drängen? Die Ergebnisse lassen sich wie folgt zusammenfassen: Die Lehrereinkommen sind in Deutschland, im Gegensatz zu den USA oder Großbritannien, nicht gefallen – weder absolut noch relativ zu den Einkommen anderer Akademiker. Entsprechend ist auch kein Abfall in den Abiturleistungen von Lehramtsanwärtern zu verzeichnen. Die Gesamtgruppe der Lehramtsanwärter ist über die Zeit hinweg zwischen 0.17 und 0.27 Standardabweichungen schlechter als andere Studierende an westdeutschen Universitäten. Hinsichtlich der Arbeitslosigkeitsrisiken zeigen sich deutliche Schwankungen auf dem Teilarbeitsmarkt Schule. Für Berufsanfänger im Lehramt steigt das Arbeitslosigkeitsrisiko bis Mitte der 1980er stark an, bevor es ab Mitte der 1990er deutlich unter das Niveau anderer Akademiker zurückfällt. Die Studienwahlmotive korrespondieren mit diesen Schwankungen. Die Wahrscheinlichkeit das Lehramt zu wählen, ist für Personen mit einer hohen beruflichen Sicherheitsmotivation sehr gering in Zeiten eines hohen Arbeitslosigkeitsrisikos. Sie steigt aber deutlich ab Mitte der 1990er, als die Arbeitsmarktaussichten deutlich vielversprechender sind. Im Gegensatz dazu war der Anteil an Personen mit einer hohen fachlichen Motivation bis Mitte der 1990er im Lehramt höher als in anderen Studienfächern, vermutlich weil Personen mit niedriger Fachmotivation zu dieser Zeit in andere Studienfächer ausgewichen sind. Der Beitrag zeigt am Beispiel des Lehramtes, wie Arbeitsmarktaussichten die Studienfachwahl beeinflussen können. Dadurch ergibt sich für den Lehrerberuf je nach Arbeitsmarktlage ein unterschiedlicher Pool an Bewerbern.

Marion Spengler (Universität Luxemburg), Martin Brunner (Freie Universität Berlin), Oliver Lüdtke (Humboldt-Universität zu Berlin), Rodica Damian (University of Illinois), Romain Martin (Universität Luxemburg), Brent Roberts (University of Illinois)

## ***Pessimismus oder Fleiß – was ist ausschlaggebend für einen erfolgreichen Bildungsweg bzw. beruflichen Erfolg?***

**Chair(s):** Marion Spengler (Universität Luxemburg, Freie Universität Berlin)

Erfolgreiche Bildungswege stellen eine der wichtigsten Stellschrauben für eine erfolgreiche berufliche Karriere dar. Der Grundstein für einen erfolgreichen Lebensweg wird bereits in den frühen Phasen einer schulischen Laufbahn gelegt (Almlund, Duckworth, Heckman & Kautz, 2011; Blossfeld, Roßbach & von Maurice, 2011). Studien im schulischen Kontext belegen hierbei die wichtige Rolle kognitiver und nicht-kognitiver Prädiktoren - wie beispielsweise IQ, Motivation und auch Persönlichkeitseigenschaften - für die Vorhersage von akademischem Erfolg (Gottfredson, 2002; Gustafsson & Undheim, 1996; Hailikari, Nevgi & Komulainen, 2007; Kuncel, Hezlett & Ones, 2004; Spengler, Lüdtke, Martin & Brunner, 2013; Steinmayr & Spinath, 2009). Vor allem die Rolle von Persönlichkeitseigenschaften ist in den letzten Jahren in den Mittelpunkt vieler Untersuchungen gerückt. Es zeigte sich, dass Jugendliche, die fleißiger und zielstrebigere waren als ihre Mitschüler, erfolgreicher in der Schule waren (Spengler et al., 2013; Poropat, 2009). Neben diesen förderlichen Eigenschaften weisen Eigenschaften wie Angst und Pessimismus einen hemmenden Einfluss auf Erfolg im schulischen Bereich auf (Judge & Bono, 2002; Yates, 2002).

Im Erwachsenenalter spielen Persönlichkeitseigenschaften ebenfalls eine große Rolle für die Vorhersage von beruflichem Erfolg. Vor allem die Faktoren Gewissenhaftigkeit und Neurotizismus leisten hierbei einen signifikanten Beitrag (Judge et al., 2002). Weniger ist jedoch darüber bekannt, welche Persönlichkeitseigenschaften, die ihren Ursprung schon im Kindes- oder Jugendalter haben (wie etwa Fleiß und Pessimismus), wichtig für eine erfolgreiche Karriere sind. Wir haben daher untersucht, ob (1) Bildungserfolg durch die Persönlichkeitseigenschaften Fleiß und Pessimismus vorhersagbar ist, (2) beruflicher Erfolg über die Lebensspanne hinweg durch diese Faktoren vorhersagbar ist, (3) dieser Effekt auf Berufserfolg über Bildung mediiert ist, (4) ob der Effekt robust gegenüber empirisch gut belegten Prädiktoren wie IQ und SES (sozioökonomischer Status) ist. Den Analysen lag eine repräsentative Stichprobe aus Luxemburg zugrunde, die über eine Spanne von 40 Jahren zu zwei Messzeitpunkten untersucht wurde. Zum ersten Messzeitpunkt im Jahr 1968 wurden Intelligenz, Persönlichkeitsmaße und der familiäre Hintergrund der Kinder (SES) in einer Stichprobe 12-jähriger Schüler erfasst, die etwa die Hälfte der damaligen luxemburgischen Schülerpopulation in Klassenstufe 6 darstellte (N = 2.824, M=11.9 Jahre, SD=0.6, 49,9% weiblich). Eine randomisiert stratifizierte und repräsentative Stichprobe wurde im Jahr 2008 ein zweites Mal getestet (N=745; M=51.7 Jahre, SD=0.6; 53,3% weiblich). Für diesen Messzeitpunkt liegen Daten zur Bildungs- und beruflichen Karriere der Teilnehmer vor. In verschiedenen längsschnittlichen Pfadmodellen wurde der Einfluss von Persönlichkeit (Fleiß, Pessimismus) im Jugendalter auf verschiedene Maße des subjektiven und objektiven Berufserfolgs (berufliche Stellung, Einkommen, bereichsspezifische Lebenszufriedenheit) 40 Jahre später untersucht. Hierbei wurde um wichtige weitere Einflussfaktoren wie Intelligenz, SES der Eltern und Bildung der Kinder kontrolliert. Die Befunde zeigen, dass

- (1) Fleiß einen positiven und Pessimismus einen negativen direkten Effekt auf die Bildung (operationalisiert durch Jahre an Bildung) haben;
- (2) Fleiß einen positiven und Pessimismus einen negativen direkten Effekt auf verschiedene Maße des Berufserfolgs haben (berufliche Stellung, Einkommen);
- (3) Bildung einen direkten Effekt auf Berufserfolg hat;
- (4) die Effekte von Persönlichkeit über die Bildung mediiert werden;
- (5) die Effekte größtenteils robust auch nach Kontrolle von Intelligenz und sozioökonomischem Status zu T1 waren.

Die Befunde legen nahe, dass Persönlichkeitseigenschaften, die sich bereits im Jugendalter manifestieren eine weitreichende Rolle für Bildungs- und Berufserfolg über die Lebensspanne hinweg spielen.

- Almlund, M., Duckworth, A.L., Heckman, J.J. & Kautz, T. (2011). Personality Psychology and Economics, In E. Hanushek, S. Machin, & L. Woessman (Eds.) Handbook of the Economics of Education, Volume 4, (pp. 1-181). Amsterdam: Elsevier.
- Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (2011). Editorial. Zeitschrift für Erziehungswissenschaft, 14, 1-4.
- Gottfredson, L.S. (2002). Where and why g matters: Not a mystery. Human performance, 15, 25-46.
- Gustafsson, J.-E. & Undheim, J.O. (1996). Individual differences in cognitive functions. In D.C. Berliner & R.C. Calfee (Eds.), Handbook of educational psychology (pp. 186-242). New York: Prentice Hall International.
- Hailikari, T., Nevgi, A. & Komulainen, E. (2007). Academic self-beliefs and prior knowledge as predictors of student achievement in mathematics: A structural model. Educational Psychology, 28, 59-71.
- Judge, T. A. & Bono, J. E. (2002). A rose by any other name: Are self-esteem, generalized self-efficacy, neuroticism, and locus of control indicators of a common construct? In B. W. Roberts & R. T. Hogan (Eds.), Personality psychology in the workplace (pp. 93-118). Washington, DC: American Psychological Association.
- Kuncel, N.R., Hezlett, S.A. & Ones, D.S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? Journal of Personality and Social Psychology, 86, 148-161.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. Psychological Bulletin, 135, 322.
- Spengler, M., Lüdtke, O., Martin, O. & Brunner, M. (2013). Personality is Related to Educational Outcomes in Late Adolescence: Evidence From Two Large-Scale Achievement Studies. Journal of Research in Personality, 47, 613-625.
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school Achievement. Learning and Individual Differences, 19, 80-90. DOI:10.1016/j.lindif.2008.05.004
- Yates, S.M. (2002). The influence of optimism and pessimism on student achievement in mathematics. Mathematics Education Research Journal, 14, 4-15.





## **Mathematische Förderung am Übergang von der Schule in den Beruf: Effekte einer Sommerakademie für Hauptschülerinnen und Hauptschüler**

Insbesondere Hauptschülerinnen und Hauptschülern gelingt ein direkter Übergang von der Schule in den Beruf deutlich seltener (Protsch & Dieckhoff, 2011). Dabei hat sich allgemein gezeigt, dass sich das Fehlen eines Hauptschulabschlusses, niedrige schulische Leistungen und ein Migrationshintergrund negativ auf den Übergang von der Schule in den Beruf auswirken (Kohlrusch & Solga, 2012). Gleichzeitig erhöhen gute Leistungen, insbesondere in Mathematik, die Chance eines erfolgreichen Übergangs (List et al., 2010). Sommerakademien können dabei eine Möglichkeit darstellen, die Leistungen von Hauptschülerinnen und Hauptschülern bereits vor Schulabschluss zu erhöhen und den Übergang von der Schule in den Beruf erfolgreicher zu gestalten. Eine Metaanalyse von Cooper et al. (2000) hat gezeigt, dass sich Sommerakademien positiv auf den Lernzuwachs der Teilnehmerinnen und Teilnehmer auswirken ( $d = 0.26$ ). Dabei profitieren leistungsschwächere Schülerinnen und Schüler ( $d=0.20$ ) jedoch in einem leicht schwächeren Maße, wobei die Zuwächse signifikant bleiben. Auch für Deutschland konnten bereits positive Effekte für Sommerakademien beobachtet werden (vgl. u.a. Stanat et al., 2012; Odau & Ehmke, eingereicht).

Dennoch liegen für Deutschland bisher erst sehr wenige Studien vor und insbesondere ein langzeitlicher Effekt der Maßnahmen ist noch weitestgehend ungeklärt. Daher soll untersucht werden: 1. Welche Effekte zeigen sich für die dreiwöchige Förderung von Hauptschülerinnen und Hauptschüler in Mathematik? 2. Wie stabil bleiben die beobachteten Effekte über die Zeit? Grundlage der Untersuchung ist die Leuphana Sommerakademie (Czerwenka, 2008), die zum Ziel hat, Hauptschülerinnen und Hauptschüler vor Verlassen der Schule zu fördern, um sie auf einen erfolgreichen Übergang von der Schule in den Beruf vorzubereiten. Hierzu kommen die Jugendlichen zwischen der achten und neunten Klasse in einem dreiwöchigen Sommercamp zusammen und werden im Anschluss während der gesamten neunten Klasse betreut. Neben der Förderung in Mathematik erhalten die Jugendlichen in weiteren Bereichen (z.B. Bewerbungstraining) Unterstützung. In 2012 nahmen 243 Schülerinnen und Schüler an sieben Standorten an der Leuphana Sommerakademie teil. Zusätzlich wurden 86 Schülerinnen und Schüler einer Kontrollgruppe befragt, so dass insgesamt 329 Jugendliche an der Untersuchung beteiligt waren. Die Daten wurden in einem Prä-Post Design erhoben, so dass die Jugendlichen zu Beginn und zum Ende des Sommercamps sowie ein Dreiviertel Jahr später Fragebögen bekamen. Der eingesetzte Mathematiktest setzt sich insgesamt aus 55 Items aus fünf verschiedenen Bereichen aufgeteilt auf zwei verschiedene Testhefte mit zehn Ankeritems zusammen. Zur Berechnung der Zuwächse wurden die Daten rasch-skaliert und z-standardisiert. Die Berechnung der Mittelwerte, Standardfehler und Standardabweichungen erfolgte mit Hilfe von MPlus, um das Mehrebenen-Design zu berücksichtigen. Für die einzelnen Testhefte liegt insgesamt eine gute interne Konsistenz vor. Die Ergebnisse zeigen, dass die Teilnehmerinnen und Teilnehmer innerhalb der drei Wochen deutlich von den Maßnahmen profitieren können. So werden teilweise Zuwächse von über einer Standardabweichung für einige Camps deutlich. Im Gegensatz hierzu kann für die Kontrollgruppe keine signifikante Veränderung beobachtet werden. Bei der langzeitlichen Betrachtung der Ergebnisse fällt allerdings auf, dass die beobachteten Zuwächse nicht stabil bleiben. So erzielen die Teilnehmerinnen und Teilnehmer der Leuphana Sommerakademie ein Dreiviertel Jahr später signifikant niedrigere Leistungen als direkt im Anschluss an die Camps. Gleichzeitig kann die Kontrollgruppe die Lücke schließen. Somit können bei den Teilnehmerinnen und Teilnehmern die hohen Effekte im Anschluss an die Camps nicht stabilisiert werden. Insgesamt wird in den Ergebnissen der Leuphana Sommerakademie deutlich, dass die Jugendlichen im Anschluss an die Camps deutlich von den Maßnahmen profitieren können, wohingegen die Kontrollgruppe auf ihrem Niveau bleibt. Dabei liegen die beobachteten Effekte sogar deutlich über den von Cooper et al. (2000) berichteten Effektstärken. Dennoch gelingt es trotz einer Nachbetreuung nicht, die hohen Zuwächse zu stabilisieren. Zukünftige Analysen müssen daher klären, welchen Einfluss die Nachbetreuung auf die Stabilisation der Leistungszuwächse hat und wie die Nachbetreuung gestaltet sein muss, um langfristige Effekte in der mathematischen Leistung zu erreichen.

Cooper, H.; Charlton, K.; Valentine, J.C. & Muhlenbruck, L. (2000). Making the Most of Summer School. A Meta-Analytic and Narrative Review. Monographs of the Society for Research in Child Development. Malden, USA: Blackwell Publishers.

Czerwenka, K. (2008): Sommerakademie. Fit für die Lehrstelle. Weinheim & Basel: Beltz Verlag.

Kohlrusch, B. & Solga, H. (2012). Übergänge in die Ausbildung: Welche Rolle spielt die Ausbildungsreife? Zeitschrift für Erziehungswissenschaft, 15 (4), 753–773. doi: 10.1007/s11618-012-0332-6

List, J., Schnabel, C. & Klaus, A. (2010). Erfolg und Misserfolg bei der Ausbildungsplatzsuche: Eine empirische Untersuchung von Hauptschülern im Landkreis Nürnberger Land. Erlangen-Nürnberg: Univ. Lehrstuhl für Arbeitsmarkt- und Regionalpolitik.

Odau, S. & Ehmke, T. (eingereicht). Förderung von Hauptschülerinnen und Hauptschülern über die Sommerferien am Übergang von der Schule in den Beruf.

Protsch, P. & Dieckhoff, M. (2011). Noten, kognitive Fähigkeiten oder Persönlichkeit: Was bei der Suche nach einem Ausbildungsplatz zählt.

Stanat, P.; Becker, M.; Baumert, J.; Lüdtke, O. & Eckhardt, A.G. (2012): Improving second language skills of immigrant students: A field trial study evaluating the effects of a summer learning program. In: Learning and Instruction. Amsterdam: Elsevier.

## **Intrapreneurship-Kompetenz von Auszubildenden – Befunde aus Reports zur mündlichen Abschlussprüfung von Industriekaufleuten**

Theoretischer Hintergrund: Schnelle Veränderungen im sozialen und ökonomischen Leben verlangen nach veränderten Fähigkeiten, die unter dem Begriff Intrapreneurship (IP) zusammengefasst werden können (u. a. Pinchot, 1987; Korunka et al., 2009; Wunderer & Bruch, 2000). Empirische Evidenzen dafür liefert die vierte europäische Erhebung über die Arbeitsbedingungen (Parent-Thirion et al., 2008). Anforderungen an ein IP-orientiertes Verhalten zeigen sich auch in den Ordnungsmitteln der beruflichen Erstausbildung IKL: IP ist seit 2002 in der betrieblichen Ausbildungsordnung sowie im Lehrplan (Lernfeld 12: Unternehmensstrategien/ -projekte umsetzen, 80h) verankert und findet seinen Niederschlag in den Reports zur mündlichen Abschlussprüfung (KMK, 2002).

Pinchot (1987) sequenziert IP in (A) die „Entwicklung einer (IP)-Projektidee“ und (B) die „Umsetzung einer (IP)-Projektidee“. Weber et al. (in Vorbereitung) identifizierten 14 Intrapreneurship-Fähigkeiten, die sich wie folgt den beiden Phasen zuordnen lassen:

1. Wahrnehmen von Problemen (A&B), 2. Analysieren von Situationen (A&B), 3. Kreieren einer Idee (A), 4. Kreativitätstechniken anwenden (A), 5. Aspekte sequenzieren und Ressourcen abschätzen (B), 6. Informationen recherchieren, selektieren, bewerten (A&B), 7. Fachbegriffe und ökonomische Theorien anwenden (B), 8. Tools anwenden (A&B), 9. Begründete Entscheidungsfindung (A&B), 10. Risikoanalyse (B), 11. Teamarbeit (B), 12. Umgang mit Störungen (B), 13. Reflektion (B), 14. Projekte verkaufen (A&B)

Fragestellungen: Betrachtet man den Report als eine Aufgabe, in der zahlreiche IP-Fähigkeiten zur Lösung angewendet werden müssen, dann stellen sich folgende Fragen: (1.) Zeigen Auszubildende (IKL) in ihren Reports zur mündlichen Abschlussprüfung IP-Fähigkeiten?, (2.) Weist das IP-Kompetenzstrukturmodell entsprechend der theoretischen Überlegungen zwei Dimensionen auf?, (3.) Lassen sich Gruppen identifizieren, die auf unterschiedliche Fähigkeitsprofile schließen lassen?

Methode: Zur Beantwortung der Forschungsfragen wurden 820 Reports aus den Abschlussprüfungen IKL (Winter 2011/2012, Sommer 2012, Winter 2012/2013) der IHK für München und Oberbayern hinsichtlich der theoretisch formulierten IP-Fähigkeiten ex-post analysiert (deduktive Inhaltsanalyse). Dabei gingen die Reports aller Zeitpunkte in eine Gesamtstichprobe ein (One-Shot-Design). Die IP-Fähigkeiten werden als Teilaufgaben (=Items) angesehen, die es im Rahmen der Reporterstellung zu lösen galt. Die Fähigkeiten 8, 11 und 12 sind dabei nicht prüfbar, da bspw. eine Teamarbeit als Prüfungsleistung ausgeschlossen ist. Anhand einer zufällig gezogenen Teilstichprobe von 25 Reports wurde das Kodierschema entwickelt; weitere 200 Reports doppelt kodiert ( $\kappa=0,88$ ).

Die Analysen erfolgen bezogen auf die FF(1) primär mittels der deskriptiven Statistik. Zur Bestimmung der Kompetenzstruktur (Modellgüte) FF(2) wurde das eindimensionale Raschmodell mit dem zweidimensionalen Random-Coefficients-Multinomial-Logit-Modell (MRCML-model: within-item-multidimensionality) verglichen. Zur Identifizierung möglicher latenter Leistungsgruppen (FF(3)) wurde eine Mixed-Rasch-Analyse nach Rost (1990) durchgeführt. Zur Prüfung von Mastery-Einflüssen auf die Daten werden zusätzlich diagnostische Messmodelle, wie das Fusion-Modell (Hartz & Roussos, 2005) angewendet. Eine Q-Matrix mit den zwei Fähigkeits-Komponenten (A) und (B) bildet die Grundlage für diese weiterführenden Analysen.

Ergebnisse: FF(1): Die Ergebnisse der deskriptiven Statistik zeigen, dass IP-Fähigkeiten eingesetzt wurden. Die Range der relativen Lösungshäufigkeiten liegt zwischen 0,5% und 98,5%. Für die weiteren Analysen werden Items mit Lösungshäufigkeiten unter 5% und über 95% eliminiert (4, 5, 6, 7).

FF(2): Bezogen auf die Kompetenzstruktur fittet im Ergebnis das zweidimensionale within-Modell die Daten besser (BIC = 5.700,8; AIC = 5.653,7; EAP/PV-Reliabilität Dimension (A)/ (B): 0.68/ 0.49). Die zugehörigen Itemgüten weisen Werte von  $0.75 < wMNSQ < 1.34$  auf.

FF(3): Die Mixed-Rasch-Analyse (MRM) (Rost, 1990) zur Prüfung einer latenten Gruppenstruktur schlug bei Verwendung des BIC drei Gruppen mit unterschiedlichen Itemschwierigkeiten vor (BICk=2 = 5218.8; BICk=3 = 5158.8; BICk=4 = 5254.0). Auch der AIC bestätigt diese Zahl. Alternativ sollen Gruppen mittels diagnostischer Modelle (Fusion-Modell; Hartz & Roussos, 2005) identifiziert und mit den Gruppen der MRM-Analysen verglichen werden.

Zusammenfassend zeigt sich, dass IP-relevante Fähigkeiten in unterschiedlicher Intensität bei der Erstellung der Reports zum Einsatz kommen. Außerdem gibt es erste Hinweise darauf, dass die theoretische Kompetenzstruktur gilt und unterschiedliche Fähigkeitsprofile existieren. Letzteres wird noch vertiefend (insbesondere im Hinblick auf Mastery-Einflüsse) analysiert.

Hartz, S. M., & Roussos, L. A. (2005). The Fusion Model for skills diagnosis: Blending theory with practice, ETS Research Report. Princeton, NJ: Educational Testing Service.

KMK (The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany) (2002). Rahmenlehrplan für den Ausbildungsberuf Industriekaufmann/Industriekauffrau. Bonn.

Korunka, C., Frank, H., Lueger, M. & Ebner, M. (2009). Entwicklung und Prüfung eines Modells zur Förderung von Intrapreneurship in der dualen Berufsausbildung. Zeitschrift für Personalpsychologie, 8(3), 129-146.

Parent-Thirion, A., Fernández Macías, E., Hurley, J. & Vermeylen, G. (2008). Vierte Europäische Erhebung über die Arbeitsbedingungen. Retrieved on 04. 07. 2012 from <http://www.eurofound.europa.eu/pubdocs/2006/98/de/1/ef0698de.pdf>

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. Applied Psychological Measurement, 14, 271-282.

Weber, S., Trost, S., Wiethe-Körprich, M., Weiß, C. & Achtenhagen, F. (in Vorbereitung). Intrapreneur: an entrepreneur within a firm – an approach on modeling and measuring intrapreneurship competence. In: S. Weber, F. Oser, F. Achtenhagen, M. Fretschner & S. Trost (Eds). Becoming an entrepreneur. Rotterdam, Taipei: Sense.

Wunderer, R. & Bruch, H. (2000). Umsetzungskompetenz. Diagnose und Förderung in Theorie und Unternehmenspraxis. München: Vahlen.



## **Probleme und Herausforderungen der Messung von Kompetenzen bei kaufmännischen Auszubildenden für ein nachhaltiges Wirtschaften**

Der wissenschaftliche Diskurs über Bildung für eine nachhaltige Entwicklung (BNE) und den zu fördernden Kompetenzen hat sich in den vergangenen Jahren intensiviert (Barth & Michelsen, 2013). Kompetenzmodelle wurden insbesondere für den allgemeinbildenden Bereich entwickelt wie beispielsweise System-, Bewertungs- sowie Gestaltungskompetenz (Rost, Lauströer & Raack, 2003). Erste Befunde zur Modellierung und Validierungen liegen zum Teil vor (für die Gestaltungskompetenz: Rost, Lauströer & Raack, 2003; Rost, 2008; für die Bewertungskompetenz: Eggert, 2008). Daneben sind Ergebnisse aus internationalen Studien aus dem Bereich „environmental literacy“ zu nennen, bei denen positiv korrelierte Zusammenhänge zwischen Wissen, Einstellungen, Skills, und Partizipation zu nachhaltigkeitsrelevanten Themen gemessen wurden (Marcinkowski et al., 2013). Im Bereich der kaufmännischen Berufsausbildung, bestehen bisher keine Kompetenzmodelle, mit Ausnahme des Referenzmodells von Schlömer (2009), welches nicht empirisch-quantitativ, sondern qualitativ-explorativ geprüft wurde. Insgesamt ist festzuhalten, dass bisher nur wenige empirische Zugänge im Bereich der Messung von Kompetenzen in der BNE vorliegen. Schwierigkeiten in der Kompetenzforschung im Bereich von BNE sind vor allem auf die inhaltliche Komplexität des Nachhaltigkeitsgedanken zurückzuführen, insbesondere wegen des Vernetzungsgrads ökonomischer, ökologischer sowie sozialer Nachhaltigkeitsdimensionen (Gräsel et al., 2012, S. 13). Das Ziel dieses Beitrags besteht in der Beschreibung von Problemen und Herausforderungen der empirischen Validierung eines Kompetenzmodells für ein nachhaltiges Handeln kaufmännischer Auszubildender. Exemplarisch erfolgt dies am Beispiel der Ausbildungsberufe Verkäufer/-in (N=105), Kaufmann/-frau im Einzelhandel (N=449) sowie Bürokaufmann/-frau (N=392). Das Kompetenzmodell wurde in Anlehnung an die Theorie des geplanten Verhaltens (Ajzen, 1985) entwickelt, bei der die Intention, nachhaltig zu wirtschaften, die wichtigste bestimmende Variable für das später gezeigte Verhalten ist. Diese Intention wird von verschiedenen Aspekten beeinflusst: Einstellungen, der subjektiven Norm sowie der wahrgenommenen Verhaltenskontrolle zum nachhaltigen Wirtschaften. Dem Stand der gegenwärtigen Kompetenzdiskussion entsprechend wurden Handlungsmotivation und Wissen über Nachhaltigkeit in das Modell integriert. Anforderungen zur Messung der Intention wurden als Situational-Judgement-Testitems (SJT) gestaltet, bei dem die Auszubildenden eine Handlungsempfehlung im Rahmen einer Situation geben müssen. Die Items des SJT werden über ein Partial-Credit-Modell bewertet. Der Wissenstest ist als Multiple-Choice-Test konzipiert und beinhaltet Fakten- sowie Konzeptwissen. Die Auswertung des SJT und Wissenstests erfolgt über das 1PL-Rasch-Modell. Zur Messung von Normen, Einstellungen sowie der Verhaltenskontrolle wurden eigene Tests entwickelt und explorativ (Cronbachs Alpha: .625 bis .848) sowie konfirmatorisch (CFI: .97 bis 1.000; RMSEA: .090 bis .000) ausgewertet. Für die Motivation wurden die Skalen von Prenzel et al. (1996) adaptiert. Bei allen Tests außer dem Wissenstest wird bezüglich des Handlungskontexts eine betriebliche und private Dimension unterschieden. Die Items des Wissenstests sind sowohl für den betrieblichen als auch für den privaten Kontext relevant. Die vorliegenden Daten zeigen signifikant positive Zusammenhänge zwischen den Ajzen-Konstrukten. Den größten Einfluss auf die Handlungsintention weist das Wissen über Nachhaltigkeit auf. Allerdings erweist sich die Validierung des SJT als große Herausforderung. Das geschlossene Design mit der Vorgabe potenzieller Handlungsalternativen begünstigt sozial erwünschtes Antwortverhalten. Weiterhin können niemals alle potenziellen Handlungsintentionen für eine bestimmte Situation abgebildet werden. Daneben führten heterogen gewählte Handlungsintentionen zu einer bisher noch nicht zufriedenstellenden Reliabilität. Aus diesen Gründen erfolgte eine Neukonzipierung dieses Testteils mit offenerem Antwortformat. Die ersten Auswertungen für den Bereich Spedition und Logistik (N=354) werden vorgenommen und sollen im Beitrag ebenfalls vorgestellt werden. Das zugrunde gelegte Kompetenzmodell wird anhand von Strukturgleichungsanalysen diskutiert.

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl, & J. Beckmann (Hrsg.), *Action Control* (S. 11-39). Berlin: Springer.

Barth & Michelsen (2013). Learning for change: an educational contribution to sustainability science. *Sustainability Science*, 8, S. 103-119.

Eggert, S. (2008). Bewertungskompetenz für den Biologieunterricht - Vom Modell zur empirischen Überprüfung. Abgerufen am 15.10.2013 von: <http://hdl.handle.net/11858/00-1735-0000-0006-AD10-8>

Gräsel, C., Bormann, I., Schütte, K., Trempler, K., Fischbach, R. & Asseburg, R. (2012). Perspektiven der Forschung im Bereich Bildung für nachhaltige Entwicklung. In Bundesministerium für Bildung und Forschung (Hrsg.), *Bildungsforschungsband 19: Bildung für nachhaltige Entwicklung - Beiträge der Bildungsforschung*. Berlin: BMBF.

Marcintowski, T., Shin, D., Noh, K.-I., Negev, M., Sagy, G., Garb, Y., McBeth, B., Hungerford, H., Volk, T., Meyers, R. & Erdogan, M. (2013). National Assessments of Environmental Literacy. In R. B. Stevenson, M. Brody, J. Dillon & A. E. J. Wals (Hrsg.), *International Handbook of Research on Environmental Education* (S. 310-330). New York: Routledge.

Prenzel, M., Kristen, A., Dengler, P., Eitle, R. & Beer, T. (1996). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, Supplement 13, S. 108-127.

Rost, J., Lauströer, A. & Raack, N. (2003). Kompetenzmodelle einer Bildung für Nachhaltigkeit. *Praxis der Naturwissenschaften*. In: *Chemie in der Schule*, 8, S. 10-15.

Rost, J. (2008). Zur Messung von Kompetenzen einer Bildung für nachhaltige Entwicklung. In I. Bormann & G. de Haan (Hrsg.), *Kompetenzen der Bildung für nachhaltige Entwicklung: Operationalisierung, Messung, Rahmenbedingungen, Befunde* (S. 7-12). Wiesbaden: VS Verlag für Sozialwissenschaften.

Schlömer, T. (2009). Berufliches Handeln und Kompetenzen für nachhaltiges Wirtschaften: Ein Referenzmodell auf der Grundlage theoretischer und empirischer Explorationen. In K. Rebman (Hrsg.), *Schriften zur Berufs- und Wirtschaftspädagogik*, Band 6. München: Rainer Hampp.



## Schülerfeedback an beruflichen Gymnasien – zur Problematik des fairen Vergleichs

Schülerratings werden sowohl in der Unterrichtsforschung als auch zu schulpraktischen Zwecken, d. h. zur Evaluation der Lehre und des Unterrichts genutzt. Damit stellt sich das Problem, wann die Schülerrückmeldungen als eine eher positive und wann als eine eher negative Rückmeldung zu interpretieren sind. In gängigen Verfahren der Schul- und Unterrichtsevaluation ist es üblich, den Lehrerinnen und Lehrern das Feedback der Klasse im Vergleich mit dem „aggregierten Bewertungsprofil“ der eigenen Schule oder auch einer schulübergreifenden Stichprobe zurückzumelden und das Schülerfeedback im Vergleich zum jeweiligen Stichprobenmittelwert zu beurteilen (vgl. Wagner, 2010). Angesichts des Forschungsstandes zur Aussagefähigkeit des Schülerfeedbacks ist jedoch zweifelhaft, ob ein solcher bildungsgang- und fächerübergreifender sozialer Vergleich als fair gelten kann. Vor allem das Fachinteresse der Schülerinnen und Schüler ist ein gut belegter und relativ stark wirksamer Einflussfaktor auf die Schülerrückmeldungen, dessen Beeinflussbarkeit durch die Lehrperson zumindest fraglich ist (vgl. u. a. Greimel-Fuhrmann, 2003; Greimel-Fuhrmann & Geyer, 2005). Schülerinteressen differenzieren sich mit steigender Schulbesuchsdauer zunehmend domänenspezifisch aus (vgl. Baumert & Kunter, 2006; Daniels, 2008) und insbesondere im Vergleich zwischen beruflichen Gymnasien verschiedenen Typs ist von Unterschieden in der „dominanten Interessenorientierung der Schülerschaft“ auszugehen (vgl. Lüdtke & Trautwein, 2004). Im Rahmen einer Vorstudie zu der aktuellen Untersuchung haben sich varianzanalytisch signifikante Mittelwertunterschiede im Interesse der Schüler am Unterrichtsfach in Abhängigkeit vom unterrichteten Fach, aber nicht in Abhängigkeit von der Fachrichtung des beruflichen Gymnasiums gezeigt. Überdies konnte mit Hilfe von Strukturgleichungsmodellen gezeigt werden, dass sich das Fachinteresse der Schüler zwar stets signifikant, aber im Fächervergleich nicht gleichförmig auf die wahrgenommene Unterrichtsqualität auswirkt. Es wird erwartet, dass sich diese Befunde anhand der Stichprobe der Hauptuntersuchung reproduzieren lassen. In der Konsequenz wäre – so die dem Vortrag zugrundeliegende These – ein fairer Vergleich zwar gymnasialtypenübergreifend, aber nicht fächerübergreifend möglich. Der Vortrag geht mit anderen Worten der Frage nach, inwieweit die Gefahr besteht, dass die Lehrenden an beruflichen Gymnasien durch einen fach- und bildungsgangübergreifenden Vergleich der Rohwerte des Schülerfeedbacks zur Qualität ihres Fachunterrichts eine positiv bzw. eine eher negativ verzerrte Rückmeldung erhalten. Zu diesem Zweck werden die Daten mehrebenenanalytisch ausgewertet und in Analogie zur Debatte um die „fairen Vergleiche von Schülerleistungen“ (vgl. Arnold, 2002; Helmke, 2008) auf der Grundlage der Befunde multipler Regressionsanalysen „adjustierte Mittelwerte“ berechnet und mit den Rohwerten verglichen. Als Datenbasis der im Rahmen der GEBF-Tagung zu präsentierenden Analysen dienen die im Sommer 2013 im Zuge eines laufenden, von der DFG geförderten Forschungsprojektes (Rahn, Gruehn & Böttcher, 2012), an 39 kaufmännischen, technischen und erziehungswissenschaftlichen beruflichen Gymnasien erhobenen Schülerrückmeldungen zur Unterrichtsqualität. Befragt wurden insgesamt 1532 Schülerinnen und Schülern in 88 Lerngruppen zur Qualität des Unterrichts in Mathematik, Deutsch und dem jeweiligen profildbildenden Fach (BWL/Rechnungswesen, Elektrotechnik und Erziehungswissenschaft). Die Unterrichtsqualität bei der Entwicklung der Erhebungsinstrumente wurde in Anlehnung an das Angebots-Nutzungsmodell von Helmke modelliert. Bei der Operationalisierung der Prozessmerkmale des Unterrichts und des Fachinteresses der Schülerinnen und Schüler wurde überwiegend auf eingeführte Skalen und Items der empirischen Unterrichts- und Bildungsforschung zurückgegriffen, die gekürzt und ggf. geringfügig sprachlich angepasst wurden. Die theoretisch erwartete Skalenstruktur der Erhebungsinstrumente konnte in der Voruntersuchung faktorenanalytisch bestätigt werden. Die Reliabilitäten der Skalen sind mit Alphawerten zwischen .67 und .90 als akzeptabel bis (sehr) gut zu bezeichnen. Erste deskriptive Analysen der Daten zeigen die zu erwartenden Mittelwertunterschiede. Ebenfalls divergiert das Interesse am Unterrichtsinhalt zwischen den Fächern. In beiden Fällen fallen die Unterschiede zu Gunsten des profildbildenden Fachs aus.

Arnold, K.-H. (2002): Schulentwicklung durch Rückmeldung der Lernwirksamkeit an die Einzelschule: Möglichkeiten und Grenzen der Schuleffizienzforschung. In: Zeitschrift für Pädagogik. 48. Jg./Heft 5, S. 741-764.

Baumert, J. & Kunter, M. (2006): Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft. 9. Jg./Heft 4, S.469-520.

Daniels, Z. (2008): Entwicklung schulischer Interessen im Jugendalter. Münster: Waxmann.

Greimel-Fuhrmann, B. (2003): Evaluation von Lehrerinnen und Lehrern. Einflussgrößen auf das Gesamturteil von Lernenden. Innsbruck/Wien/München: Studienverlag.

Greimel-Fuhrmann, B. & Geyer, A. (2005): Die Wirkung von Interesse und Sympathie auf die Gesamtbeurteilung in der Lehrevaluation. Direkte und indirekte Effekte unter Berücksichtigung des Lehrverhaltens. In: Empirische Pädagogik. 19. Jg./Heft 2, S.103-120.

Gruehn, S. (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster: Waxmann.

Helmke, A. (2008): Unterrichtsqualität. Erfassen, Bewerten, Verbessern. Seelze: Kallmeyer.

Lüdtke, O. & Trautwein, U. (2004): Die gymnasiale Oberstufe und psychische Ressourcen. Gewissenhaftigkeit, intellektuelle Offenheit und die Entwicklung von Berufsinteressen. In: Wege zur Hochschulreife in Baden-

Württemberg: TOSCA - eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen: Leske und Budrich, S.367-401.

Wagner, C. (2010): Unterrichtsentwicklung durch Evaluation am Beispiel des „Netzwerk Schülerbefragung“. In: Die berufsbildende Schule. 62. Jg./Heft 10, S.291-296.

Wagner W. (2008): Methodenprobleme bei der Analyse der Unterrichtswahrnehmung aus Schülersicht. Am Beispiel der Studie DESI (Deutsch Englisch Schülerleistungen International) der Kultusministerkonferenz. Koblenz-Landau: Univ., FB Psychologie.

Montag,  
03.03.2014,  
11:15 - 13:00,  
IG-Farben-Haus  
Raum 1.411



## ***Ins Studium navigieren – den Überblick nicht verlieren!***

### ***Zur Ausdifferenzierung der RIASEC-Interessensdimensionen im STUDIEN-NAVI***

Chair(s): Anne Milatz (Universität Wien)

Die Wahl des passenden Studiums stellt jede Schulabschlusskohorte und ihre Familien vor eine herausfordernde Entscheidung mit dem Ziel eine optimale Passung von Interessen, Eignung und Studienfach zu erzielen. Auch aus bildungspolitischer und sozioökonomischer Perspektive ist es ein zentrales Anliegen, diese Passung zu optimieren. Um angehende Studierende in ihrer Studienwahlentscheidung zu unterstützen, werden u.a. Instrumente zur Erfassung von Studieninteresse eingesetzt, die zumeist auf der etablierten Interessentheorie - dem sechs-dimensionalen RIASEC-Modell - von J. Holland (1997) basieren. Mit dem Ziel, die Beratungspräzision durch (1) einen evidenzbasierten Zugang und (2) eine Ausdifferenzierung der sechs Interessensdimensionen zu erhöhen, wurde ein neuer Interessenstest STUDIEN-NAVI (Gittler, 2012) entwickelt. Er umfasst 120 einfache Aussagen und dauert im Durchschnitt 10.6 Minuten. Insgesamt wurde der Online-Fragebogen im Rahmen einer Normdatenerhebung N = 13.818 Studierenden (M = 23;4 Jahre, SD = 3;11; 59 % weiblich) aus 125 verschiedenen Studienrichtungen in Österreich vorgegeben. Die daraus empirisch gewonnenen Interessensprofile der 125 Studienrichtungen stellen die Grundlage für die im Ergebnisbericht enthaltenen Studienrichtungsvorschläge dar.

Die vorliegende Studie will einerseits der Frage nachgehen, ob sich die sechs RIASEC-Dimensionen replizieren bzw. in weitere Subdimensionen ausdifferenzieren lassen. Andererseits soll geklärt werden, ob das erweiterte Interessensmodell zur Beratungspräzision beiträgt.

Anhand der Normdaten (N = 13.818 Studierende) konnte mittels konfirmatorischer Faktorenanalyse ein 15-dimensionales Interessensmodell identifiziert werden, in welchem die sechs Hauptdimensionen nach Holland in zwei bis drei weitere Subdimensionen ausdifferenziert wurden. Zudem konnte das 15-dimensionale Interessensmodell anhand einer weiteren Stichprobe von N = 1312 MaturantInnen (M = 18;11 Jahre, SD = 1;1; 52 % weiblich) kreuzvalidiert werden. Hinsichtlich der Kriteriumsvalidität zeigte sich, dass die Beratungspräzision anhand der 15 Interessensdimensionen im Vergleich zu den herkömmlichen 6 Dimensionen erhöht wird. Darüber hinaus zeigten sich für das STUDIEN-NAVI zufriedenstellende interne Konsistenzen und Retest-Reliabilitäten. Die Relevanz der Ergebnisse sollen in Bezug auf die aktuelle Interessensforschung sowie Studienberatungspraxis diskutiert werden.



## ***Psychologie Lernen im Hochschulkontext – Die Bedeutsamkeit von Vorwissenstests und wie sich das Wissen von Studierenden mit wenigen vs. vielen Misconceptions entwickelt***

Theoretischer Hintergrund: Vorwissen ist ein wichtiger Prädiktor zukünftiger Leistung; dabei ist die Art des Vorwissens entscheidend, da sich falsches Vorwissen auch negativ auswirken kann (Thompson & Zamboanga, 2004). Dies ist besonders für psychologische Erkenntnisse relevant, da diese auch oft in den Medien diskutiert, aber z. B. verkürzt dargestellt werden (Thompson & Nelson, 2001). Somit stehen Hochschullehrende im Fach Psychologie vor der Herausforderung, falsches Vorwissen der Studierenden, sog. Misconceptions (Vaughan, 1977), durch korrektes Wissen zu ersetzen. Hierzu muss das falsche Vorwissen zunächst identifiziert werden.

Vorwissenstests sind jedoch in universitären Veranstaltungen nicht weit verbreitet, trotz diverser Vorteile (vgl. Thompson & Zamboanga, 2003): Lehrende können das vorhandene (richtige oder falsche) Vorwissen der Studierenden einschätzen, die Vermittlung der Lerninhalte danach ausrichten, den Einfluss ihrer Lehre evaluieren und überprüfen, ob der Lernerfolg von den Vorkenntnissen der Studierenden beeinflusst wird. Außerdem erhalten die Lernenden einen Einblick in die Themen der Lehrveranstaltung und ggfs. das Format der Prüfung. Um eine Vorlesung für Lehramtsstudierende in Pädagogischer Psychologie im Sinne des Forschenden Lehrens (Spinath & Seifried, 2012) kontinuierlich zu verbessern, wird das Wissen der Studierenden über das Semester hinweg beobachtet. Hierbei wird nicht nur die Korrektheit des Wissens, sondern auch die Sicherheit dieses Wissens erfasst. Somit können sowohl die Misconceptions der Studierenden erfasst als auch die Entwicklung des Wissens untersucht werden.

Fragestellung: Der vorliegenden Studie liegen die folgenden Forschungsfragen zugrunde:

- 1) Geht mit einem Lernzuwachs in psychologischen Lehrveranstaltungen auch eine positive Entwicklung der Sicherheit des Wissens einher?
- 2) Wie viele Misconceptions sind bei Studierenden über psychologische Inhalte vorhanden und wie entwickelt sich das Wissen bei Studierenden mit wenigen gegenüber Studierenden mit vielen Misconceptions?

Methode: Zu drei Messzeitpunkten (erste Vorlesungssitzung, Mitte des Semesters und letzte Vorlesungssitzung) bearbeiteten N = 198 Studierende einen Wissenstest. Hierbei wurde das Antwortformat von Dutke und Barenberg (2009) eingesetzt: Die Studierenden gaben an, ob sie vorgegebenen Aussagen zustimmen, und mit welcher Sicherheit sie dies jeweils tun. Die 15 Items des Vorwissenstests wurden in gleicher Form in die Abschlussklausur integriert. Dadurch können u.a. Entwicklungseffekte des Wissens überprüft werden. Basierend auf den Ergebnissen des Vorwissenstests (15 Items) wurden die Studierenden nachträglich in zwei Gruppen unterteilt: 1) Studierende, die keine oder nur wenige Misconceptions hatten (n = 111), und 2) Studierende, die einige bis viele Misconceptions hatten (n = 87). Misconceptions wurden definiert als das Wissen, das die Studierenden als sicher einstufen, das jedoch falsch ist. Die Daten wurden mithilfe varianzanalytischer Verfahren ausgewertet.

Ergebnisse: Bezüglich der ersten Fragestellung zeigte sich zunächst allgemein ein signifikanter Wissenszuwachs über die drei Messzeitpunkte. Dies ging damit einher, dass der prozentuale Anteil an sicherem falschem Wissen abnahm und der prozentuale Anteil an sicherem richtigem Wissen zunahm.

Bezüglich der zweiten Fragestellung ergab sich, dass die Studierenden bei den 15 Items des Vorwissenstests zwischen 0 und 9 Misconceptions (M = 2.52, SD = 1.95) hatten. Dies entspricht einem prozentualen Anteil von 0% bis 60% an Misconceptions im Vorwissen der Studierenden. Um zu überprüfen, wie sich das Wissen von Studierenden mit wenigen gegenüber Studierenden mit vielen Misconceptions entwickelt, wurden lediglich die 15 Items betrachtet, die sowohl im Vorwissenstest als auch in der Abschlussklausur in gleicher Form vorhanden waren. Innerhalb der allgemeinen Wissenszunahme zeigte sich hierbei zu beiden Messzeitpunkten ein Vorteil der Studierenden mit wenigen Misconceptions. Hinsichtlich der Anzahl an Misconceptions zeigte sich eine signifikante Interaktion zwischen der Gruppenzugehörigkeit und der Zeit: Die Anzahl an Misconceptions beider Studierendengruppen näherte sich im Verlauf des Semesters auf einem niedrigen Niveau an. Aufgrund des Designs der vorliegenden Studie können keine kausalen Schlussfolgerungen gezogen werden. Dennoch vermag die Studie auf die mögliche Bedeutsamkeit von Vorwissenstests und damit identifizierbaren Misconceptions im Hochschulkontext hinzuweisen.

Dutke, S. & Barenberg, J. (2009). Ökonomisch anwendbare Wissenstests für teilnehmerreiche Psychologieveranstaltungen. In M. Krämer, S. Preiser & K. Brusdeylins (Hrsg.), Psychologiedidaktik und Evaluation VII (S. 367-376). Aachen: Shaker.

Spinath, B. & Seifried, E. (2012). Forschendes Lehren: Kontinuierliche Verbesserung einer Vorlesung. In M. Krämer, S. Dutke & J. Barenberg (Hrsg.), Psychologiedidaktik und Evaluation IX (S. 171-180). Aachen: Shaker.

Thompson, R. A., & Nelson, C. A. (2001). Developmental science and the media: Early brain development. *American Psychologist*, 56, 5–15.

Thompson, R. A., & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in Introduction to Psychology. *Teaching of Psychology*, 30, 96–101.

Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, 96, 778-784.

Vaughan, E.D. (1977). Misconceptions about psychology among introductory psychology students. *Teaching of Psychology*, 4, 138–141.



## **Veränderungen absoluter und relativistischer epistemologischer Überzeugungen über den Verlauf des Psychologiestudiums**

Theoretischer Hintergrund: Epistemologische Überzeugungen (EB) werden als Auffassungen einer Person über die Beschaffenheit und den Prozess des Wissens definiert (vgl. Hofer & Pintrich, 1997). Ausgehend von Perry (1970) lassen sich vier aufeinanderfolgende Phasen der Entwicklung von EB über den Verlauf der schulischen und universitären Ausbildung hinweg beschreiben: Dualismus, Multiplizität, Relativismus und Postrelativismus; dabei werden die beiden erstgenannten Phasen häufig verallgemeinernd als absolute Überzeugungen zusammengefasst. Insbesondere für die Entwicklung in späteren Phasen gilt höhere Bildung als entscheidende Voraussetzung (Hofer, 2001). EB werden oftmals mittels multidimensionaler, disziplinspezifischer Fragebogen-Verfahren gemessen, deren Items aus epistemologischen Aussagen bestehen, die auf Likert-Skalen hinsichtlich ihres Zutreffens zu beurteilen sind (z.B. Schommer, 1990). Die meisten dieser Fragebögen enthalten ausschließlich oder überwiegend Items mit Aussagen, welche absolute oder postrelativistische Überzeugungen widerspiegeln (z.B. Bendixen, Schraw, & Dunkle, 1998). Häufig wird eine zunehmende Ablehnung absoluter Aussagen oder zunehmende Zustimmung zu postrelativistischen Aussagen als Hinweis auf eine günstige Entwicklung von EB interpretiert. Mithilfe dieser Vorgehensweise scheint es allerdings kaum möglich, den Relativismus von EB angemessen abzubilden, da relativistische Überzeugungen zu ähnlichen Antwortmustern führen würden wie postrelativistische. Beispielsweise würden sowohl Personen mit relativistischen als auch solche mit postrelativistischen Überzeugungen eine absolute Aussage (z.B. „Die Wahrheit ist in dieser Disziplin unveränderlich.“; Hofer, 2000) ablehnen und einer postrelativistischen Aussage (z.B. „Auffassungen in der Wissenschaft verändern sich manchmal.“; Conley, Pintrich, Vekiri, & Harrison, 2002) zustimmen. Demgegenüber könnte mithilfe relativistischer Aussagen (z.B. „Das was Wissenschaftler als richtig erachten ist im ständigen Wandel“) zwischen relativistischen und postrelativistischen Überzeugungen unterschieden werden, da Personen mit postrelativistischen Überzeugungen nicht nur absolute, sondern auch relativistische Aussagen ablehnen sollten.

Fragestellung: Die Studie zielte darauf ab, die Entwicklung von EB über den Verlauf des Psychologiestudiums im Rahmen eines querschnittlichen Untersuchungsdesigns zu untersuchen. Es wurde angenommen, dass (1) absolute und relativistische Aussagen relativ unabhängige, allenfalls marginal korrelierte Dimensionen von EB konstituieren und (2) die Zustimmung zu absoluten und relativistischen Aussagen bei steigendem Studienfortschritt abnimmt.

\*Methode: Es wurde ein Item-Inventar mit  $k = 35$  Aussagen erstellt, die auf Likert-Skalen zu beurteilen waren. Jedes der Items ließ sich einer von acht Kategorien zuordnen, die sich aus der Kreuzung der Entwicklungsphase von EB (absolut vs. relativistisch) mit vier inhaltlichen Facetten von EB (Stabilität, Komplexität, Quelle und Rechtfertigung; Hofer, 2000). Das Inventar wurde  $N = 278$  Studierenden der Psychologie (davon 85% weiblich; mittleres Alter  $M = 22.98$  Jahre) zur Bearbeitung vorgelegt. Für die Prüfung von Hypothese 2 wurde die Stichprobe entsprechend dem Studienfortschritt in vier Gruppen unterteilt: 1. Studienjahr Bachelor ( $n = 84$ ), 2. Studienjahr Bachelor ( $n = 57$ ), Bachelor ab dem 3. Studienjahr ( $n = 77$ ); Graduierte / Master-Studierende ( $n = 60$ ).

Ergebnisse: Zunächst wurde für jede der vier inhaltlichen Facetten eine exploratorische Faktorenanalyse durchgeführt. Chi-Quadrat Vergleichstests legten bei jeder Faktorenanalyse eine Zwei-Faktorenlösung nahe (alle  $p < .01$ ), wobei absolute und relativistische Items auf jeweils unterschiedlichen Faktoren luden.

Im Vergleich verschiedener getesteter Strukturgleichungsmodelle wies ein Modell die höchste Anpassungsgüte auf, das zum einen absolute und relativistische Überzeugungen als zwei orthogonale latente Variablen (LV) zweiter Ordnung und zum anderen acht LV erster Ordnung (entsprechend den acht Item-Kategorien) enthielt. Zudem wurden Zusammenhänge zwischen den Residuen der Items gleicher inhaltlicher Facetten angenommen. Sämtliche Goodness-of-Fit Indizes zeigten zufriedenstellende Werte ( $\chi^2 = 305.81$ ,  $df = 228$ ,  $\chi^2/df = 1.34$ , CFI = .93, RMSEA = .04, SRMR = .06). Untersuchungen der Zusammenhänge zwischen Studienfortschritt und den beiden LV zweiter Ordnung zeigten keinen signifikanten Zusammenhang zu absoluten Überzeugungen und einen signifikant negativen Zusammenhang zu relativistischen Überzeugungen ( $\beta = -.25$ ,  $SE = 0.08$ ,  $p < .01$ ). Mit zunehmendem Studienfortschritt tendierten die Studierenden weniger dazu, relativistischen Aussagen zuzustimmen. Dieser Befund legt nahe, dass während des Studiums der Psychologie der Übergang von relativistischen zu postrelativistischen Überzeugungen wesentlich ist.

Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic beliefs and moral reasoning. *Journal of Psychology: Interdisciplinary and Applied*, 132(2), 187–200. doi:10.1080/00223989809599158

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186–204. doi:10.1016/j.cedpsych.2004.01.004

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25(4), 378–405. doi:10.1006/ceps.1999.1026

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, 13(4), 353–383. doi:10.1023/A:1011965830686

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. doi:10.2307/1170620

Perry, W. G., Jr. (1970). *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. Rinehart and Winston, Inc: Holt.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498–504. doi:10.1037/0022-0663.82.3.498



## ***Kann ein Studienabbruch Perspektiven bieten? Empirische Befunde auf Basis quantitativer Daten***

Die hohe Relevanz des Themas lässt sich u.a. an der steigenden Anzahl von Fachpublikationen abzulesen (s. dazu beispielsweise Schröder-Gronostay 1999, Pohlenz et al., 2007). Knapp ein Drittel aller Bachelorstudierenden an deutschen Hochschulen verlässt über alle Studienbereiche hinweg das begonnene Studium ohne Abschluss (Heublein et al., 2012, S. 12). Für diese Entscheidung zum Abbruch sind mehrere Gründe verantwortlich, die für die betroffenen Studierenden unterschiedliches Gewicht haben. Der Entschluss das begonnene Studium abzubrechen, ist also das Ergebnis eines komplexen Entscheidungsgefüge, der von unterschiedlichen Einflussfaktoren geprägt ist. Der Studienabbruch ist neben hohen gesellschaftlichen Kosten oft auch mit individuellen Kosten verbunden. Jedoch kann so ein Lebenschnitt mitunter unterschiedliche Chancen bieten, die erst durch perspektivistische Betrachtungsweise als solche erkennbar werden. Daher ist die Erforschung der Studienabbruchursachen eine wichtige Aufgabe der Bildungsforschung. Trotz der Relevanz und Aktualität des Themas muss allerdings konstatiert werden, dass der Studienabbruch noch nicht ausreichend untersucht ist (Köster, 2002). Insbesondere wurde die Entwicklung integrativer und empirisch ausreichend gestützter Modelle vernachlässigt. Aufgabe des Beitrags ist somit u.a., dieses Forschungsdesiderat zu erfüllen.

So steht im Zentrum des vorliegenden Beitrags die Frage nach den Motiven für das vorzeitige Verlassen einer Hochschule ohne einen qualifizierten Abschluss erlangt zu haben. Demzufolge sollen konkret diese Forschungsfragen beantwortet werden: Aus welchen Gründen verlassen Studierende die Hochschule ohne Abschluss? Gibt es signifikante Unterschiede hinsichtlich des Studienabbruchs nach Studienabschluss, Geschlecht, Alter und sozialer Herkunft? Lassen sich im Prozess des Studienabbruchs bei den Gruppen der Hochschulwechsler vs. Studienabbrecher Unterschiede feststellen?

Die Fragestellung wird mit einer empirisch-quantitativen und standardisierten Studie auf der Grundlage schriftlicher Befragungen an der Technischen Hochschule Mittelhessen (n=165) angegangen. Auf dieser Datenbasis werden vorgestellte Fragestellungen mithilfe multivariater Datenanalysen untersucht. Im Ergebnis stellt sich der vollzogene Studienabbruch als mangelnde Verpflichtung hinsichtlich des Studiums bzw. des begonnenen Studienfaches dar. Extrinsische Motive als auch ein fehlendes fachliches Interesse erhöhen die Wahrscheinlichkeit auf einen Studienabbruch. Auch eine verstärkte Erwerbstätigkeit während des Studiums kann den Studienabbruch begünstigen. Auf der anderen Seite bietet ein solcher 'Bruch' in der Biographie unter Umständen Möglichkeiten, aus denen sich alternative Perspektiven entwickeln können. Einige Forschungsbefunde zeigen, dass Studienabbrecher/innen durchaus beruflich erfolgreicher, vor allem aber zufriedener als die Absolventen/innen sein können (Lewin, 1996).

Heublein, Ulrich/Richter, Johanna/Schmelzer, Robert/Sommer, Dieter (2012). Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen. Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010, HIS: Forum Hochschule.

Köster, Fritz (2002). Studienabbruch. Perspektiven und Chancen Bund-Verlag, Frankfurt am Main.

Lewin, Karl (1999). Studienabbruch in Deutschland. In: M. Schröder-Gronostay und H.-D. Daniel (Hg.): Studienerfolg und Studienabbruch. Beiträge aus Forschung und Praxis. Neuwied: Luchterhand Verlag.

Pohlenz, Philipp; Seyfried, Markus & Tinsner, Karen (2007). Studienabbruch. Ursachen, Probleme, Begründungen. Saarbrücken: VDM, Müller.

Schröder-Gronostay, Manuela & Daniel, Hans-Dieter (1999). Studienerfolg und Studienabbruch. Beiträge aus Forschung und Praxis. Neuwied: Luchterhand Verlag.

Chair(s): Mareike Kunter (Goethe-Universität Frankfurt)

Unterrichtsqualität ist eines der zentralen Themen der aktuellen empirischen Lehr-/Lernforschung. Im Vortrag sollen am Beispiel des Fachs Mathematik ausgewählte Aspekte von „gutem Fachunterricht“ beleuchtet werden: Wie kann selbständiges Lernen gefördert werden? Wie können Diagnose, Feedback und Unterstützung aufeinander abgestimmt werden? Welcher Lehrerqualifikationen bedarf es hierzu? Und welche Rolle spielt in diesem Kontext die Fachdidaktik? Exemplarisch wird dabei auf einige interdisziplinäre empirische Studien zum Lehren und Lernen von Mathematik Bezug genommen: auf DISUM zum selbständigen Lernen, auf Co<sup>2</sup>CA zu Diagnose und Feedback sowie auf COACTIV zu Lehrerqualifikationen.

Montag,  
03.03.2014,  
14:00 - 15:00,  
Casino-Gebäude  
Raum 823



Ilonca Hardy (Goethe-Universität Frankfurt) und Melihan Cinar (Goethe-Universität Frankfurt)

## ***Die Bedeutung des Wortschatzes für Sprachentwicklung und Fördermaßnahmen bei Kindern mit Deutsch als Zweitsprache***

Chair(s): Ilonca Hardy (Goethe-Universität Frankfurt), Melihan Cinar (Goethe-Universität Frankfurt)

Diskutant/in: N. N.

Kinder mit Deutsch als Zweitsprache liegen nach verschiedenen Studien schon im Kindergarten in unterschiedlichen Bereichen der sprachlichen Entwicklung unter dem Niveau von einsprachig deutschen Kindern. In der Forschung zur Sprachentwicklung hat sich insbesondere der Bereich des Wortschatzes als bedeutsam erwiesen, da dieser auf vielfältige Weise mit der Entwicklung von anschlussfähigen Fähigkeiten im bildungssprachlichen Bereich verbunden ist. So ist die Wortschatzentwicklung beispielsweise prädiktiv für die Ausbildung schriftsprachlicher Kompetenzen und der Lesekompetenz. Gleichzeitig steht der Wortschatz als Teil der Bildungssprache selbst im Fokus von Förderansätzen im Elementar- und Primarbereich. Allerdings gibt es theoretisch wie empirisch Klärungsbedarf hinsichtlich der Zusammenhänge zwischen der Wortschatzentwicklung und spezifischen Komponenten der (schrift-) sprachlichen Fähigkeiten insbesondere bei der Gruppe von Kindern mit Deutsch als Zweitsprache. Gleichmaßen gibt es trotz vielfältiger Ansätze der Sprachförderung in der Unterrichtspraxis wenige empirisch validierte Befunde zu den Gestaltungsprinzipien von Wortschatztrainings, den Möglichkeiten einer Kombination unterschiedlicher Sprachförderdomänen und dem Wissen über Wortschatz bei Sprachförderkräften. Dieses Symposium beleuchtet die Wortschatzentwicklung und –förderung bei Kindern mit Deutsch als Zweitsprache aus den unterschiedlichen Perspektiven der Linguistik, Deutschdidaktik, Erziehungswissenschaft und Psychologie und stellt deren Bedeutung für ausgewählte Bereiche der schriftsprachlichen Entwicklung im Grundschulbereich dar. Während im ersten Teil des Symposiums die sprachlichen Voraussetzungen von Kindern mit Deutsch als Zweitsprache aus Sicht der Linguistik und der Fachdidaktik thematisiert werden, stehen im zweiten Teil Aspekte der Förderung im Primarbereich anhand von spezifischen Interventionen im Fokus, welche dem Bereich der Lehr-Lernforschung entstammen.

Barbara Geist und Anja Müller stellen zwei Studien vor, in denen der Zusammenhang zwischen Wortschatz und Grammatik untersucht wird. Dabei stellt sich der Wortschatz für Kinder mit Deutsch als Zweitsprache als bedeutsam für die syntaktische Entwicklung heraus; bei Sprachförderkräften zeigt sich jedoch nur ein geringes Wissen hinsichtlich dieser Sprachdomänen. Der Beitrag von Magdalena Spaude und Ulrich Mehlem thematisiert den Zusammenhang zwischen dem Wortschatz und der Entwicklung von Rechtschreibkompetenz in einem längsschnittlichen Design. Auf Grundlage von Sprachstanderhebungen in der Schuleingangsphase wird insbesondere untersucht, inwieweit Wortschatz und Morphologie prädiktiv für die Rechtschreibkompetenz sind. Der Beitrag von Annika Ohle, Nele McElvany et al. fokussiert die Wortschatzförderung bei Kindern mit Deutsch als Zweitsprache durch Lesetexte unter Nutzung der Muttersprache. In einer experimentellen Trainingsstudie konnte gezeigt werden, dass sich der Ansatz des Kontextlernens für Kinder mit Deutsch als Zweitsprache als produktiv erwies. Im Beitrag von Sabrina Förster und Frank Hellmich wird die Wortschatzarbeit schließlich als Teil eines Lesestrategietrainings in Grundschulklassen herangezogen, wobei sich signifikante Effekte für die Lesekompetenzentwicklung der Kinder zeigten.

**GEBF** TAGUNG  
2014

**B02**

**Symposium**

Montag,

03.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 311

2



## **Wortschatz und Grammatik am Übergang zwischen Kita und Grundschule: Evidenzen aus der Spracherwerbs- und Kompetenzforschung**

Die Frage, ob lexikalische Fähigkeiten ein Prädiktor für die Entwicklung grammatischer Fähigkeiten sind, ist für die Spracherwerbsforschung zentral (für einen Überblick Marhman & Thal, 2005). Ergebnisse der Erstspracherwerbsforschung belegen, dass der Lexikonerwerb und der Ausbau grammatischer Fähigkeiten miteinander einhergehen (Moyle et al., 2007). Studie mit Kindern mit Deutsch als Zweitsprache (DaZ) zeigen, dass diese die gleichen Erwerbsschritte in Bezug auf den Erwerb von Haupt- und Nebensatzstrukturen durchlaufen, wie Kinder mit Deutsch als Muttersprache (DaM) (u.a. Thoma & Tracy, 2009). Auch der Wortschatzerwerb bei DaZ-Kindern ist in Bezug auf den Verlauf und die Komposition vergleichbar mit dem Erwerb bei DaM-Kindern (Jeuk, 2003). Auch wenn ein qualitativ ähnlicher Erwerbverlauf in beiden Bereichen zu verzeichnen ist, unterscheiden sich die sprachlichen Fähigkeiten der DaZ-Kinder in diesen Bereichen quantitativ von denen der DaM-Kinder (u.a. Duboway et al., 2008). Im Gegensatz zum Erstspracherwerb ist in der Zweitspracherwerbsforschung der Zusammenhang zwischen Wortschatz und Grammatik bisher wenig erforscht (jedoch Fallstudie Rothweiler, 2009). In der ersten Studie wird daher die folgende Frage untersucht: Ist der Umfang des Wortschatzes ein Prädiktor für die Entwicklung syntaktischer Fähigkeiten von Kindern mit DaZ?

In der Sprachdiagnostik und Sprachförderung nehmen die Bereiche Grammatik und Wortschatz eine zentrale Rolle ein (Geist, 2013). Für das Gelingen von Sprachförderung spielt die Sprachförderkompetenz der pädagogischen Fachkraft eine zentrale Rolle (Müller im Druck). Jedoch ist bislang kaum untersucht, über welches Fachwissen Sprachförderkräfte verfügen (u.a. Thoma et al., 2011). Untersuchungen zum Wissen über die Bereiche Wortschatz und Grammatik gibt es bislang nicht. Daher wird in der zweiten Studie die folgende Frage fokussiert: Über welches Fachwissen im Bereich Wortschatz und Grammatik verfügen Sprachförderkräfte? In Studie (1) wurden die lexikalischen und syntaktischen Fähigkeiten von 37 6-jährigen DaZ-Kindern zu zwei Testzeitpunkten erfasst. Zu T1 wurde der expressive Wortschatz mittels AWST-R (Kiese-Himmel, 2005) erhoben. Die syntaktischen Fähigkeiten wurden zu T1 und T2 (im Abstand von vier Monaten) mittels LiSe-DaZ (Schulz & Tracy, 2011) erfasst. Die Daten des Wortschatztests wurden getrennt für Nomen und Verben ausgewertet. Für die syntaktischen Fähigkeiten wurde die Produktion der Verbzweitstellung in Hauptsätzen und der Verbendstellung in Nebensätzen ausgewertet. Der Vergleich zwischen dem Verbwortschatz zu T1 und den syntaktischen Fähigkeiten zu T2 zeigt eine signifikante Korrelation ( $r=.441$ ,  $p=.006$ ). Der Nomenwortschatz zu T1 korreliert nicht mit den syntaktischen Fähigkeiten zu T2 ( $r=.224$ ,  $p=.183$ ). In Studie (2) wurde das Fachwissen von 31 Sprachförderkräften mit einem standardisierten Test (SprachKoPF, Tracy & Thoma, 2012) erfasst. Der Test umfasst 55 Multiple-Choice-Fragen zum Basiswissen über Sprache, bestehend aus Fragen zu den einzelnen sprachlichen Ebenen (N=22) und zum praxisrelevantem Wissen über Spracherwerb, Sprachdiagnostik und Sprachförderung (N=33). Die Sprachförderkräfte erreichten einen Durchschnittswert von .28 (SD=.19; Min=.00; Max=.62). In den Fragen zum praxisrelevantem Wissen zeigten die Sprachförderkräfte signifikant bessere Leistungen als in den Fragen zum linguistischen Basiswissen ( $t(df=30)=2.34$ ;  $p<.05$ ). Werden innerhalb des Bereichs Sprache die Fragen nach den sprachlichen Ebenen differenziert wird deutlich, dass in den Bereichen Lexikon (mean=.18; SD=.27) und Syntax (mean=.19; SD=.37) die niedrigsten Punktwerte innerhalb des Bereichs Sprache erreicht wurden (Phonologie: mean=.22; SD=.31; Morphologie: mean=.36; SD=.30; Semantik/Pragmatik: mean=.42; SD=.32). Die Ergebnisse der ersten Studie zeigen, dass im frühen Zweitspracherwerb der Umfang des Verbwortschatzes, jedoch nicht der Umfang des Nomenwortschatzes, prädiktiv für die Entwicklung syntaktischer Fähigkeiten ist. Die Ergebnisse der zweiten Studie deuten darauf hin, dass zentrale Aspekte der Bereiche Lexikon und Grammatik nicht im Fachwissen von Sprachförderkräfte verankert sind. Zusammenfassend lassen sich aus diesen Ergebnissen zwei Schlussfolgerungen ableiten: Zum einen sollte in der Sprachförderung die Erweiterung des Verbwortschatzes stärker fokussiert werden. Zum anderen sollten Sprachförderkräfte vor allem in den Bereich Lexikon und Grammatik sensibilisiert werden.

- Bates, E. & Goodman, J.C. (2001). On the inseparability of grammar and the lexicon: Evidence from acquisition. In M. Tomasello and E. Bates (Eds.), *Language Development: The Essential Readings*. Malden, MA: Blackwell Publishers Inc. /
- Duboway, M.; Ebert, S.; von Maurice, J. & Weinert, S. (2008): Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten: Ein Vergleich von Kindern mit und ohne Migrationshintergrund. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40 (3), 124–134.
- Geist, B. (2013). Sprachdiagnostische Kompetenz von Sprachförderkräften. Berlin: de Gruyter.
- Kiese-Himmel, C. (2005). *Aktiver-Wortschatztest (AWST-R)*. Göttingen: Hogrefe.
- Moyle, M.J.; Weismer, S.E.; Evans, J.L. & Lindstrom, M.J. (2007). Longitudinal Relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research*, Vol. 50, 508-528.
- Müller, A. (im Druck): Profession und Sprache: Die Sicht der (Zweit)Spracherwerbs-forschung. In Tanja Betz & Peter Cloos (Hrsg.): *Kindheitspädagogische Beiträge*. Beltz Juventa.
- Müller, A.; Geist, B. & Schulz, P. (2013): Wissen und Handeln von Sprachförderkräften im Elementar- und Primarbereich. *Diskurs Kindheits- und Jugendforschung*.
- Pinker, S. (1999). *Words and rules*. New York: basic Books.
- Ricart Brede, J. (2011). Videobasierte Qualitätsanalyse vorschulischer Sprachfördersituationen. Freiburg: Fillibach.
- Rösch, H. (2011): *Deutsch als Zweit- und Fremdsprache*. Berlin: Akademie Verlag.
- Rothweiler, M. (2009). Über den Zusammenhang von Lexikon, Grammatik und Mehrsprachigkeit. *Sprachheilarbeit* 6, 246-254.
- Schulz, P. & Tracy, R. (2011). *Linguistische Sprachstandserhebung – Deutsch als Zweitsprache (LiSe-DaZ)*. Göttingen: Hogrefe.
- Thoma, D. & Tracy, R. (2012). *SprachKoPF Wissenstest*. Mannheim: Universität Mannheim.
- Tracy, R. (2007): *Wie Kinder sprechen lernen*. Tübingen: Narr.

## **Wortschatz als Prädiktor von Rechtschreibleistungen mehrsprachiger Kinder in der Schuleingangsphase**

Die Rechtschreibkompetenz hängt u.a. von einem individuellen Faktorenkomplex ab, der zwecks Vorhersagepotenzial und damit Interventionsmöglichkeiten intensiv beforscht wird. Solche Faktoren, die man bereits vor oder kurz nach dem Schuleintritt messen kann, werden auch Vorläuferfertigkeiten des Schriftspracherwerbs bezeichnet und können anhand der Dimensionen Schriftsprachspezifität und Verortung mit je zwei Ausprägungen beschrieben werden. Für den Vortrag ist die Sprachentwicklung zentral, die einen spezifischen internalen (in der Person und nicht der Umwelt verorteten) Faktor darstellt, was ebenfalls auf die in diesem Zusammenhang häufig genannte phonologische Bewusstheit zutrifft. Im Fokus stehen wird die Entwicklung in der L2 Deutsch unter besonderer Berücksichtigung des Wortschatzes und der Grammatik als Prädiktor für die Rechtschreibleistung in der Schuleingangsphase. Sowohl für DaM- als auch DaZ-Kinder konnte die Bedeutung des Sprachstands neben oder vor anderen häufig genannten Prädiktoren, wie phonologische Bewusstheit oder kognitive Leistungsfähigkeit, empirisch nachgewiesen werden, so zum Beispiel von Goldammer, Mähler, Bockmann, & Hasselhorn (2010), Knievel, Daseking, & Petermann (2010), Schröder-Lenzen & Merkens (2006), Mücke & Schröder-Lenzen (2010) und Weber, Marx, & Schneider (2007). In der letzteren Untersuchung wurde herausgefunden, dass DaZ-Kinder zwar im Rechtschreiben von Realwörtern schlechter abschnitten (mehr als eine halbe Standardabweichung), allerdings nicht beim Schreiben von Kunstwörtern. Daraus kann abgeleitet werden, dass die Ergebnisse des Rechtschreibtests der DaZ-Kinder niedriger ausfallen aufgrund geringerer Erfahrung mit dem Realwortschatz. Da aber für das System der deutschen Wortschreibung neben dem phonographisch-silbischen und dem syntaktischen Prinzip auch das morphologische gilt (vgl. Bredel, Furhop, & Noack, 2011, Blatt, Voss, Kowalski, & Jarsinski, 2011), wird die Hypothese aufgestellt, dass sowohl der Umfang des Wortschatzes als auch die Beherrschung der Morphologie ein relevanter Prädiktor für die Rechtschreibleistung ist. Dieser Frage wird im Vortrag mithilfe von Regressions- und Mediatoranalyse nachgegangen. Die der Untersuchung zugrundeliegenden Daten stammen aus dem Forschungsprojekt Literalität und Interaktion in der Sprachförderung (LISFör), das in einer Längsschnittstudie die Wirksamkeit von Sprachfördermaßnahmen in Bielefeld für mehrsprachige Kinder in der Schuleingangsphase evaluiert hat. Im Rahmen der Evaluation wurden im Winter 2010/11, im Frühjahr/Sommer 2011 und 2012 Sprachdaten von über 100 mehrsprachigen Schülerinnen und Schülern erhoben, die zum ersten Messzeitpunkt die 1. Klasse besuchten. Zu allen drei Zeitpunkten wurde u.a. der standardisierte Sprachstandstest SET 5-10 (Petermann, Metz, & Fröhlich 2010) sowie bei den beiden späteren Messungen der Rechtschreibtest Hamburger Schreib-Probe HSP (May 2012) erhoben. Die Untertests des SET 5-10 vom Anfang der 1. Klasse stellen die Prädiktorvariablen dar, während der Zuwachs der Rechtschreibleistung zwischen der HSP 1 und HSP 2 (Ende der 2. Klasse) als abhängige Variable dient. Dabei beträgt die Stichprobengröße 99 mehrsprachige Kinder. Die bivariaten Korrelationen zwischen den Rohwerten des Lexiktests, des Pluralbildungstests und dem Graphemtrefferzuwachs (unstandardisierte Residuen) fallen mittel bis hoch aus:  $r_{LG} = 0,33$ ;  $r_{PG} = 0,43$ ;  $r_{LP} = 0,60$ ;  $p \leq 0,001$ . Die Ergebnisse der Mediatoranalyse zeigen, dass es einen signifikanten indirekten Effekt von Lexik über den Mediator Pluralbildung auf den Zuwachs von Graphem-treffern (Rechtschreibung) gibt:  $b = 0,35$ ; 95% Bootstrap CI95 [0,13; 0,59]. Der Effekt ist dabei als mittel bis groß zu bezeichnen:  $\kappa^2 = 0,19$ ; Bootstrap CI95 [0,08; 0,28]. Den signifikanten Effekt bestätigt ebenfalls der Sobel-Test,  $Z = 2,90$ ;  $p = 0,004$ . Im weiteren Vorgehen sollen diese Ergebnisse anhand von Strukturgleichungsmodellen auf latenter Ebene geprüft werden. Aus spracherwerbstheoretischer Sicht kann diese Mediation damit erklärt werden, dass eine bestimmte Grundmenge an Wortschatz vorhanden sein muss, damit der Grammatikerwerb in Gang kommen kann (vgl. Szagun, 2006: 128). In jedem Fall stellt die Pluralbildung als Teil der Morphologie neben der Lexik einen starken Prädiktor für die Rechtschreibleistung von mehrsprachigen Kindern in der Schuleingangsphase dar.

Bredel, Ursula; Fuhrhop, Nanna; Noack, Christina (2011): Wie Kinder lesen und schreiben lernen. 1. Aufl. Tübingen: Francke.

Blatt, Inge; Voss, Andreas; Kowalski, Kerstin; Jarsinski, Stephan (2011): Messung von Rechtschreibleistung und empirische Kompetenzmodellierung. In: Ursula Bredel und Winfried Ulrich (Hg.): Weiterführender Orthographieerwerb. Baltmannsweiler: Schneider Hohengehren (Deutschunterricht in Theorie und Praxis, 5), S. 226–256.

Knievel, Julia; Daseking, Monika; Petermann, Franz (2010): Kognitive Basiskompetenzen und ihr Einfluss auf die Rechtschreib- und Leseleistung. In: Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie (42), S. 15–25.

Marx, Peter (2007): Lese- und Rechtschreiberwerb. Paderborn [u.a.]: Schöningh.

May, Peter (2012): Hamburger Schreib-Probe. Neustandardisierung 2012, 1. Aufl. Stuttgart: Verlag für pädagogische Medien im Ernst Klett Verlag.

Mücke, Stephan; Schröder-Lenzen, Agi (2010): Schulleistungen von Schülern im Anfangsunterricht - Welche Rolle spielt der Migrationshintergrund. In: Karl-Heinz Arnold, Katrin Hauenschild, Britta Schmidt und Birgit Ziegenmeyer (Hg.): Zwischen Fachdidaktik und Stufendidaktik. Perspektiven für die Grundschul-pädagogik. 1. Aufl. Wiesbaden: VS Verl. für Sozialwiss (Jahrbuch Grundschulforschung, 14), S. 81–84.

Petermann, Franz; Metz, Dorothee; Fröhlich, Linda Paulina (2010): SET 5-10. Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren. Göttingen: Hogrefe.

Schröder-Lenzen, Agi; Merkens, Hans (2006): Differenzen schriftsprachlicher Kompetenzentwicklung bei Kindern mit und ohne Migrationshintergrund. In: Agi Schröder-Lenzen (Hg.): Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang. 1. Aufl. Wiesbaden: VS Verl. für Sozialwiss, S. 15–44.

Szagun, Gisela (2006). Sprachentwicklung beim Kind: ein Lehrbuch. Weinheim: Beltz.

von Goldammer, Arianne; Mähler, Claudia; Bockmann, Ann-Katrin; Hasselhorn, Marcus (2010): Vorhersage früher Schriftsprachleistungen aus vorschulischen Kompetenzen der Sprache und der phonologischen Informationsverarbeitung. In: Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie (42), S. 48–56.

Weber, Jutta; Marx, Peter; Schneider, Wolfgang (2007): Die Prävention von Lese-Rechtschreibschwierigkeiten bei Kindern mit nichtdeutscher Herkunftssprache durch ein Training der phonologischen Bewusstheit. In: Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie (21), S. 65–75.



Annika Ohle (TU Dortmund), Nele McElvany (TU Dortmund), Judith Razakowski (Universität Duisburg-Essen),  
Ilonca Hardy (Goethe-Universität Frankfurt), Melihan Cinar (Goethe-Universität Frankfurt)

## **Förderung des Bildungswortschatzes im Deutschen unter Einbezug der Familiensprache. Eine Interventionsstudie in der Primarstufe**

Rationale: Die Schülerschaft in Deutschland ist durch eine hohe Heterogenität geprägt. Viele Schülerinnen und Schüler (SuS) haben einen Migrationshintergrund und sprechen eine andere Familiensprache als Deutsch (OECD, 2010). SuS mit türkischem Migrationshintergrund bilden die größte Gruppe und zeigen im Vergleich zu SuS ohne Migrationshintergrund deutlich geringere Leistungen in Schulleistungstests (z.B. Stanat et al., 2010). Die Förderung von bildungssprachlichen Kompetenzen bei SuS mit Migrationshintergrund steht zunehmend im Fokus, um bestehende Disparitäten in Schulleistungen abzubauen. Der Wortschatz ist substantiell für die Entwicklung von (bildungs-)sprachlichen Kompetenzen und bietet einen Ansatzpunkt zu deren gezielter Förderung. Die Theory of Learning from Context (Sternberg & Powell, 1983) kann als Grundlage für einen impliziten Förderansatz dienen, bei dem die Bedeutung von Wörtern aus dem Kontext – beispielsweise einem Text – entnommen wird. Einige Studien belegen die Effektivität dieser Methode (z.B. Fukkink, 2005). Allerdings gibt es bislang wenige Studien zur Nutzbarmachung des Kontextlernens für SuS mit Migrationshintergrund.

Fragestellung: Es ist anzunehmen, dass die Zurverfügungstellung zusätzlichen Textinputs in der Familiensprache die Bedeutungsentnahme für SuS mit Migrationshintergrund erleichtert. In dem Vortrag werden daher zwei zentrale Fragestellungen untersucht: 1. Lernen Kinder mit türkischer Familiensprache in verschiedenen Interventionsbedingungen mehr deutschen Wortschatz als vergleichbare Kinder, die keine Intervention erhalten? Es wird angenommen, dass Kinder in allen Interventionsbedingungen mehr deutschen Wortschatz lernen als in der Kontrollgruppe. 2. Ist bei Kindern mit türkischer Familiensprache der Wortschatzerwerb aus dem sprachlichen Kontext deutscher Texte effektiver, wenn das gleiche Fördermaterial auch in der Familiensprache zur Verfügung gestellt wird? Es wird angenommen, dass der Einbezug der Familiensprache das Wortschatzlernen positiv auswirkt.

Methode: Insgesamt nahmen an der Untersuchung N = 143 SuS mit türkischem Migrationshintergrund und grundlegendem türkischen Textverständnis teil. Das Durchschnittsalter lag bei Malter = 9,80 (SD = 0,58) Jahren und 49% waren weiblich. Im Rahmen der Intervention mit 15 Sitzungen lasen die SuS Texte, die jeweils 9 Zielwörter aus dem deutschen Bildungswortschatz enthielten. Diese Texte wurden vorher auf ihre geschlechtsneutrale, Interessantheit und Unbekanntheit der Zielwörter hin überprüft. Die SuS wurden 4 Interventionsgruppen zugeordnet. Die Textpräsentation in Interventionsgruppe A erfolgte zwei Mal auf Deutsch, während Interventionsgruppe B den Text erst auf Türkisch und dann auf Deutsch las. Interventionsgruppe C unterschied sich darin, dass sie den Text auf Deutsch mit türkischer Übersetzung der Zielwörter erhielt und anschließend ohne Übersetzung. Gruppe D wurde als reine Kontrollgruppe in die Untersuchung miteinbezogen und bekam das Interventionsmaterial ohne Zielwörter zu lesen. Die Zuordnung der SuS auf die Interventionsbedingungen erfolgte randomisiert; die SuS unterschieden sich nicht signifikant hinsichtlich ihrer kognitiven Grundfähigkeiten, ihres Alters und Geschlechts. Direkt vor und nach der Intervention wurde der Zielwortschatz im Deutschen mittels eines Synonymetests erhoben.

Ergebnisse: Zur Beantwortung der ersten Forschungsfrage nach der Effektivität der Intervention für Kinder mit türkischem Migrationshintergrund (Interventionsbedingungen A-C vs. D) zeigt eine einfaktorielle univariate Kovarianzanalyse für den Faktor Intervention nur einen Effekt für den Zielwortschatzzuwachs auf dem 10%-Signifikanzniveau ( $F[1, 137] = 3.07, p < .10, \eta^2 = .02$ ); die Kovariate Prätest zeigte einen statistisch signifikanten Effekt ( $F[1, 137] = 237.01, p < .001, \eta^2 = .64$ ). Die zweite Forschungsfrage zielt auf das Potenzial der Muttersprache ab (Gruppe B vs. D). Hier zeigte sich kein statistisch signifikanter Haupteffekt des Faktors Bedingung ( $F[1, 60] = 1.20, p > .05, \eta^2 = .02$ ). Die Kovariate Prätest war hingegen statistisch signifikant ( $F[1, 60] = 81.22, p < .001, \eta^2 = .58$ ).

Diskussion: Der Ansatz Theory of Learning from Context ist auch bei Kindern nichtdeutscher Familiensprache effektiv und stellt ein geeignetes Verfahren zur Wortschatzförderung dar. Allerdings hat sich in diesem Setting kein zusätzliches Potential durch den Einbezug der Familiensprache gezeigt. Weitere Forschung mit älteren und im Türkischen kompetenteren Kindern sollte dieses Ergebnis absichern.

Fukkink, R. G. (2005). Deriving word meanings from context: A process analysis. *Learning and Instruction*, 15, 23-43.

OECD. (2010). *OECD Reviews of Migrant Education - Closing the Gap for Immigrant Students: Policies, Practice and Performance*. PARIS: OECD.

Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrati-onshintergrund. In E. Klieme et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200-230). Münster: Waxmann Verlag.

Sternberg, R. & Powell, J. (1983). Comprehending verbal comprehension. *American Psychologist*, 39, 878-893.

GEBF TAGUNG  
2014

B02 c

Symposiums  
beitrag

Montag,  
03.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 311

2

Die Bedeutung des Wortschatzes für Sprachentwicklung und Fördermaßnahmen

## **Evidenzbasierte Förderung der Lesekompetenzen von Kindern mit Deutsch als Zweitsprache durch eine kombinierte Lerneinheit aus Wortschatzarbeit und Lesestrategietraining**

Als eine besondere Herausforderung gilt seit der Veröffentlichung der Befunde aus den Internationalen Grundschul-Lese-Untersuchungen (vgl. Bos et. al., 2012) die Förderung der Lesekompetenzen von Kindern, die Deutsch als Zweitsprache erwerben. Es zeigte sich hier, dass rund zwei Drittel aller Kinder mit Deutsch als Zweitsprache am Ende des vierten Schuljahres nicht über elementare Lesefähigkeiten und -fertigkeiten verfügen. Besondere Schwierigkeiten bereiten diesen Kindern das Dekodieren von Wörtern und Sätzen und die Identifikation explizit angegebener Einzelinformationen, ebenso wie das Auffinden und Verknüpfen relevanter Einzelheiten und Informationen in Erzähl- oder Sachtexten (vgl. Schwippert, Wendt, & Tarelli, 2012). Offenbar scheinen Kinder mit Deutsch als Zweitsprache nur wenig von unterrichtlichen Bemühungen im Leseunterricht der Grundschule zu profitieren.

Im Rahmen der Bildungsforschung stellt sich vor der dargestellten Befundlage die Frage, welche Maßnahmen notwendig sind, um Kindern mit Deutsch als Zweitsprache effiziente Lerngelegenheiten zu ermöglichen. Als eine Chance zur Verbesserung der Lesekompetenzen von Kindern mit Deutsch als Zweitsprache wird für den Leseunterricht in der Grundschule eine Kombination verschiedener Ebenen des Leseprozesses angesehen: Hierarchieniedrige Prozesse beziehen sich auf Verarbeitungsvorgänge im Bereich der Worterkennung (Dekodierung), hierarchiehohe Prozesse zielen auf die Erfassung globaler Zusammenhänge auf einer höheren Abstraktionsebene ab (Richter & Christmann, 2002). Programme, die explizit Wortschatztrainings (hierarchieniedrige Ebene) enthalten, existieren – mit wenigen Ausnahmen (vgl. El-Khechen, Gebauer, & McElvany, 2012) – derzeit nicht, obwohl grundsätzlich der kausale Einfluss des Wortschatzes auf Lesefähigkeiten empirisch nachgewiesen werden konnte (z.B. Beck, Perfetti & McKeown, 1982). Zwar liegen momentan zahlreiche Programme zur Förderung von Lesestrategien (hierarchiehohe Ebene) vor, allerdings sind diese nicht speziell auf die Eingangsvoraussetzungen von Kindern mit Deutsch als Zweitsprache ausgerichtet.

Im Rahmen unserer Studie wurde vor diesem Hintergrund eine Fördereinheit für Kinder mit Deutsch als Zweitsprache entwickelt, bei der hierarchieniedrige und -höhere Leseprozesse aufeinander bezogen sind. Bei dieser Fördereinheit werden auf der einen Seite Lesestrategien vermittelt und geübt (z.B. Verknüpfung relevanter Einzelheiten und Informationen in Texten, Erfassung globaler Zusammenhänge, Analyse und Verarbeitung größerer Textteile) und auf der anderen Seite Fördermaßnahmen in der Form von Wortschatzarbeit bereit gestellt, die im Wesentlichen das Dekodieren von Wörtern, das simultane Erfassen von bekannten Wörtern sowie – unter semantischem Gesichtspunkt – das Ableiten der sprachlichen Bedeutung von Wörtern betreffen. Die Wirksamkeit dieses wortschatzbasierten Lesestrategietrainings wurde im Rahmen eines Kontrollgruppen-Designs evaluiert. Insgesamt waren an der Untersuchung N=232 Schülerinnen und Schüler der dritten Jahrgangsstufe beteiligt, davon N=71 Kinder mit Deutsch als Zweitsprache. Die Hälfte der Kinder nahm über einen Zeitraum von zehn Unterrichtsstunden an dem wortschatzbasierten Lesestrategietraining teil, während die Schülerinnen und Schüler der Kontrollgruppe regulären Deutschunterricht (ohne explizite Förderung von Wortschatz oder Lesestrategien) erhielten. Prä- und postexperimentell bearbeiteten die Kinder Aufgaben zur Erfassung ihrer Lesefähigkeiten. Hierbei wurde u.a. der Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6; vgl. Schneider & Lenhard, 2006) eingesetzt. Zur Erfassung überfachlicher Effekte (Leseinteresse, Selbstkonzept im Lesen, Selbstwirksamkeitsüberzeugungen) der Implementation des wortschatzbasierten Lesestrategietrainings bearbeiteten die Schülerinnen und Schüler prä- und postexperimentell einen Fragebogen. Um den Sprachstand der an der Studie beteiligten Kinder kontrollieren zu können, wurde der Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren (SET 5-10; vgl. Petermann, Fröhlich, & Metz, 2010) präexperimentell eingesetzt. Die auf der Basis von Varianzanalysen ermittelten Befunde geben Hinweise darauf, dass die Kinder mit Deutsch als Zweitsprache aus der Experimentalgruppe von dem wortschatzbasierten Lesestrategietraining – im Vergleich zu den Kindern mit Deutsch als Zweitsprache aus der Kontrollgruppe – in Hinblick auf ihre Lesefähigkeiten signifikant profitiert haben. Durch das eingesetzte Training können bei den Kindern der Experimentalgruppe keine signifikanten Effekte in Hinblick auf einen überfachlichen Zugewinn im Bereich ihrer Selbstwirksamkeitserwartungen und ihres Leseinteresses gegenüber den Kindern der Kontrollgruppe nachgewiesen werden. Allerdings profitiert die Experimentalgruppe signifikant in Bezug auf ihre lesebezogenen Selbstkonzepte im Vergleich zu der Kontrollgruppe.

- Beck, I. L., Perfetti, C. A. & McKeown, M. G. (1982). Effects on long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, 74(4), 504-521.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.) (2012). IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.
- El-Khechen, W., Gebauer, M. M. & McElvany, N. (2012). Wortschatzförderung bei Grundschulkindern – Ein Vergleich von Kindern mit und ohne Migrationshintergrund. *Zeitschrift für Grundschulforschung*, 5(2), 48-64.
- Petermann, F., Fröhlich, L. P. & Metz, D. (2010). SET 5-10. Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren. Göttingen: Hogrefe.
- Richter, T. & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (S. 25-59). Weinheim: Beltz.
- Schneider, W. & Lenhard, W. (2006). ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.
- Schwippert, K., Wendt, H. & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 191-207). Münster: Waxmann.

## ***Self-Assessments als Instrumente der Studienberatung: Eine kritische Prüfung aktueller Verfahren***

Chair(s): Michaela Köller (Christian-Albrechts-Universität zu Kiel), Uta Klusmann (IPN)

Diskutant/in: Oliver Wilhelm (Universität Ulm)

Self-Assessments im Zusammenhang mit der Beratung angehender Studierender haben mittlerweile eine immense Verbreitung gefunden. Sie dienen der Selbsterkundung eigener Interessen, Fähigkeiten und Persönlichkeitsmerkmale via Internet im Hinblick auf die Studienfachwahl. Studieninteressierte können dabei zum einen fächerübergreifende Verfahren bearbeiten. Hierbei erhalten sie nach der Verfahrensbearbeitung konkrete Vorschläge für verschiedene Studienfächer, die mit ihren Interessen, Fähigkeiten und ggf. Persönlichkeitsmerkmalen übereinstimmen. Zum anderen können sie ihre Eignung oder ihr Interesse für ein bestimmtes angestrebtes Studium mithilfe eines studiengangspezifischen Verfahrens überprüfen. Die Selbsterkundungsverfahren haben dabei zwei Funktionen. So dienen sie der Informationsgewinnung, da sie mit den Anforderungen verschiedener Studiengänge vertraut machen und die Entscheidung für oder gegen die Aufnahme eines Studiums unterstützen. Darüber hinaus haben sie eine Selektionsfunktion, da das Ergebnis des Selbsterkundungsverfahrens (Passung vs. Nicht-Passung für ein bestimmtes Studium, Vorschlag verschiedener infrage kommender Studiengänge) einen Einfluss auf die Studienentscheidung hat. Das Ziel der Self-Assessments besteht darin, die Passung der Merkmale von Studierenden mit den Anforderungen des Studiums sicherzustellen, um die Studienzufriedenheit und den Studienerfolg zu erhöhen und damit Studienabbrüche zu vermeiden.

Das Ziel des Symposiums ist es, die Möglichkeiten und Grenzen der Unterstützung der Studienwahlentscheidung durch Self-Assessments kritisch-konstruktiv zu beleuchten. Dazu nehmen vier Autorengruppen sowohl fächerübergreifende als auch studiengangspezifische Verfahren in den Blick. Insbesondere Fragen zu Inhalten und Effekten der einzelnen Verfahren sollen beantwortet werden. Dazu werden Ergebnisse qualitativer Inhaltsanalysen, quantitativer Quer- und Längsschnitt- sowie quasiexperimenteller Studien vorgestellt.

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G168



## **Unterstützen Self-Assessments die Studienfachwahl? – Befunde und Erfahrungen aus dem Projekt Study-Finder der Universität des Saarlandes**

Die Studienfachwahl wird maßgeblich durch die spezifischen Interessen der Person und Ihre Fähigkeiten in bestimmten Bereichen beeinflusst (Schnabel, 2001). So versuchen Personen generell, solche Studiengänge zu wählen, in die Sie ihre Fähigkeiten und Neigungen einbringen können (Holland, 1997). Die Übereinstimmung zwischen individuellen Interessen und den Anforderungen im gewählten Studiengang wird in der Theorie beruflicher Interessen als Interessenskongruenz beschrieben (Holland, 1997). Vielfältige Untersuchungen liefern Belege dafür, dass die Interessenskongruenz positive Zusammenhänge mit der Zufriedenheit im Studium, der Studienleistung und dem Verbleib in Studium aufweist (Nagy, 2006; Rolfs & Schuler, 2002; Tsabari, Tziner, & Meir, 2005).

Demgegenüber zeigen andere Studien, dass die Entscheidung für oder gegen ein Studienfach auch durch vage Vorstellungen, Erwartungen und Absichten beeinflusst werden kann (Guggenberger & Kellermann, 1991; Heublein, Hutzsch, Schreiber, Sommer, & Besuch, 2010a). Falsche Erwartungen und mangelnde Informationen zu Studienbeginn weisen Zusammenhänge mit der späteren Studien(un)zufriedenheit auf (Hasenberg & Schmidt-Atzert, 2013) und können ein erhöhtes Risiko für späteren Studienabbruch darstellen (Heublein, Hutzsch, Schreiber, Sommer, & Besuch, 2010b; Schmidt-Atzert, 2005).

Ausgehend von diesen Befunden wurde an der Universität des Saarlandes ein zweistufiges Self-Assessment konzipiert, das die Aspekte Interessen und Erwartungen erfasst. Die erste Stufe beinhaltet einen fächerübergreifenden Interessentest, während die zweite Stufe studiengangspezifische Erwartungschecks für die grundständigen Studiengänge umfasst. Der Interessentest basiert auf der Theorie beruflicher Interessen (Holland, 1997) und dient als erste Orientierungshilfe bei der Suche nach einem passenden Studienfeld. Mit Hilfe des Interessentests können Schülerinnen und Schüler ihr individuelles Interessenprofil ermitteln. Durch einen Abgleich des individuellen Profils mit den Anforderungsprofilen sämtlicher grundständiger Studiengänge der Universität wird zusätzlich eine Empfehlung für passende Studiengänge angezeigt.

Die studiengangspezifischen Erwartungschecks ermöglichen es den Schülerinnen und Schüler, sich intensiver mit ihren Erwartungen an einen bestimmten Studiengang auseinanderzusetzen. Dabei sind die Aufgaben so gestaltet, dass falsche Erwartungen identifiziert und mit Hilfe ergänzender Informationstexte direkt während der Bearbeitung des Self-Assessments korrigiert werden können.

Berichtet werden Befunde zu Validität und Akzeptanz der Self-Assessment-Module. Dabei wird auf Daten aus mehreren unabhängigen längs- und querschnittlichen Studien zurückgegriffen. Betrachtet werden (1) Zusammenhänge zwischen der Empfehlung passender Studiengänge mittels des Interessentests und relevanten Merkmalen einer erfolgreichen Studienfachwahl wie z.B. Studienzufriedenheit, Tendenz zu Studienabbruch und die Einschätzung der eigenen Studienleistung; (2) Veränderungen der selbsteingeschätzten Informiertheit über den jeweiligen Studiengang durch die Bearbeitung eines Erwartungschecks sowie (3) Zusammenhänge zwischen der Korrektheit der Erwartungen zu Studienbeginn und Merkmalen einer erfolgreichen Studienfachwahl im Studienverlauf.

Guggenberger, H., & Kellermann, P. (1991). Hochschulzugang und Studienwahl: empirische und theoretische Ergebnisse (Bd. 24). Kärntner.

Hasenberg, S., & Schmidt-Atzert, L. (2013). Die Rolle von Erwartungen zu Studienbeginn: Wie bedeutsam sind realistische Erwartungen über Studieninhalte und Studienaufbau für die Studienzufriedenheit? Zeitschrift für Pädagogische Psychologie, 27(1), 87–93.

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010a). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. In HIS Forum Hochschule.

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010b). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres, 8.

Holland, J. L. (1997). Making vocational choices: A theory of vocational personalities and work environments (3. Aufl., Bd. 14). Odessa: Psychological Assessment Resources.

Nagy, G. (2006). Berufliche Interessen, kognitive und fachgebundene Kompetenzen (Dissertation). Freie Universität Berlin, Universitätsbibliothek, Berlin.



## Welche Effekte haben Self-Assessments: Eine quasi-experimentelle Studie mit Oberstufenschüler(inne)n

Hochschulen verpflichten angehende Studierende vor Aufnahme eines Studiums immer häufiger webbasierte Selbsterkundungsverfahren, sogenannte Self-Assessments, durchzuführen (vgl. Kubinger, Frebort, & Müller, 2013). Die Implementierung solcher Verfahren zielt darauf ab, die Passung zwischen den Eingangsvoraussetzungen der Studierenden und den Anforderungen des jeweiligen Studiums zu erhöhen und somit zu vermehrten Studienerfolgen und einer Verringerung von Studienabbrüchen und Studienfachwechseln beizutragen. Trotz des großen Angebots an diesen Verfahren und ihrer teils verpflichteten Bearbeitung steht eine Evaluation der Effekte und ihrer Wirksamkeit weitgehend aus (Klusmann, Köller, & Kunter, 2011).

Die aktuelle Studie untersucht die Frage, welche Effekte verschiedene Selbsterkundungsverfahren auf kognitive, affektive und motivationale Aspekte der Studienwahl bei Oberstufenschüler/innen haben. Untersucht werden studiengangübergreifende im Vergleich zu studiengangspezifischen Selbsterkundungsverfahren, die einer von Studierenden durchgeführten Studienberatung gegenübergestellt werden. Im Rahmen dieses Beitrags geht es insbesondere um die Fragen, welche Informationen die Schüler/innen aus den verschiedenen Verfahren gewinnen, wie sie diese Informationen bewerten und inwieweit die Verfahren zu einer Überprüfung oder Bestätigung der bisherigen Studienwünsche und Berufsziele oder zu einer Umorientierung führen. Darüber hinaus soll gezeigt werden, inwiefern die Bewertung und die Wirkung von kognitiven und Persönlichkeitsmerkmalen der Schüler/innen abhängig ist. Die Datenbasis bildet eine Studie, die im Sommer 2013 an zehn Gymnasien in Schleswig-Holstein durchgeführt wurde. Insgesamt nahmen N = 529 Schülerinnen und Schüler des 11. und 12. Jahrgangs teil. Die Studie war als Quasi-Experiment mit Prä-Post-Vergleich angelegt. Zwischen der Bearbeitung der Prä- und Post-Fragebögen bearbeiteten die Schüler/innen randomisiert eines von drei Selbsterkundungsverfahren, das studiengangspezifische Career Counselling for Teachers (CCT, Mayr 2011), eines von zwei studiengangübergreifenden Verfahren „was-studiere-ich.de“ (Schuler, Hell, Päßler, & Schüpbach, 2013) oder EXPLORIX® (Fux, Stoll, Bergmann, Hell, & Eder, 2013) oder sie erhielten eine halbstandardisierte allgemeine Studienberatung, die von Studierenden durchgeführt wurde. Die Auswahl der Self-Assessments orientierte sich an ihrem Verbreitungsgrad, ihrer wissenschaftlichen Fundierung, Wirtschaftlichkeit und Durchführbarkeit. Alle Self-Assessments wurden individuell am Computer bearbeitet, die Studienberatung wurde in der Gruppe durchgeführt.

Zur Untersuchung der Effekte der Selbsterkundungsverfahren und der Studienberatung wurden Varianzanalysen berechnet. Erste Analysen zeigen, dass Schüler/innen durch die allgemeine Studienberatung Wissen über Studiengänge, Studienbedingungen und Unterstützungsangebote generieren, während die Self-Assessments stärker zu einer Überprüfung der Passung von eigenen Kompetenzen und Studienanforderungen führen. Es zeigte sich, dass es signifikante Unterschiede in der affektiven Bewertung der gewonnenen Informationen gibt. So wird der Informationsgewinn durch die Studienberatung signifikant positiver bewertet als der der Self-Assessments, wobei das studiengangspezifische Verfahren am schlechtesten abschneidet. Der Informationsgewinn durch die Studienberatung wird als interessanter, hilfreicher, anspornender und weniger entmutigend und unpassend bewertet. Unterschiede zwischen den 11. und 12. Jahrgängen oder zwischen denjenigen, die ein mehr oder weniger klares Bild von ihren Studienwünschen haben, zeigten sich nicht. Die Überprüfung des Studien- bzw. Berufswunsches oder die Umorientierung nach der Bearbeitung eines Verfahrens bzw. nach der Studienberatung, ist im Mittel bei allen Verfahren eher gering ausgeprägt; Unterschiede zwischen den Verfahren sind nicht signifikant. Alle Selbsterkundungsverfahren einschließlich der Studienberatung scheinen eher die ursprünglichen Studien- und Berufswünsche zu bestätigen.

Klusmann, U., Köller, M. & Kunter, M. (2011). Anmerkungen zur Validität eignungsdiagnostischer Verfahren bei angehenden Lehrkräften. Zeitschrift für Pädagogik, 57, 711-721.

Kubinger, K. D., Frebort, M. & Müller, C. (2013). Self-Assessments im Rahmen der Studienberatung. Möglichkeiten und Grenzen. In K. D. Kubinger, M. Frebort, L. Khorramdel & L. Weitensfelder (Hrsg.), Self-Assessment: Theorie und Konzepte. Pabst: Lengerich.

Mayr, J. (2011). Persönlichkeitsansatz in der Lehrerforschung. Konzepte, Befunde und Folgerungen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), Handbuch der Forschung zum Lehrerberuf (S. 125–148). Münster: Waxmann.

Hell, B. (2009). Selbsttests zur Studienorientierung: nützliche Vielfalt oder unnützer Wildwuchs? In G. Rudinger & H. Hörsch (Hrsg.), Self-Assessments an Hochschulen: Von der Studienfachwahl zur Profilbildung. Göttingen: V & R unipress.

Joerin Fux, S., Stoll, F., Bergmann, C., Hell, B. & Eder, F. (2013). EXPLORIX® - Das Werkzeug zur Berufswahl und Laufbahnplanung - Deutschsprachige Adaptation und Weiterentwicklung des Self-Directed Search® (SDS) nach John Holland. Göttingen: Hogrefe.



## **Auswahlgespräche als Beratungstool in der Lehramtsausbildung**

Theoretischer Hintergrund: Seit einigen Jahren wird die Eignung für den Lehrerberuf an Hochschulen, in der Politik und in der Öffentlichkeit zunehmend diskutiert. Hintergrund sind vor allem Erkenntnisse über die Belastungen und die berufliche Zufriedenheit von Lehrkräften (z.B. Schaarschmidt, 2004). Insbesondere die Gestaltung des Zugangs zum Lehramtsstudium steht dabei im Blickpunkt. In Folge dessen wurde in letzter Zeit eine Vielzahl an zulassungsrelevanten Assessment- und Beratungsverfahren für Lehramtsstudiengänge entwickelt. Den aufwändigen Bemühungen zur Gestaltung solcher Verfahren stehen jedoch nur wenige Erkenntnisse über deren Wirksamkeit gegenüber. An der TUM School of Education werden Bewerberinnen und Bewerber für das gymnasiale Lehramtsstudium seit einigen Jahren zu einem zulassungsrelevanten Auswahlgespräch mit Hochschuldozierenden und Lehrkräften aus der Schulpraxis eingeladen. Ziel ist es, zur Reflexion über den Studien- und Berufswunsch anzuregen, Transparenz über die Anforderungen von Studium und Beruf zu schaffen und persönliche Potenziale und Entwicklungsmöglichkeiten aufzuzeigen. Grundlage des teilstrukturierten Gesprächs ist ein Kompetenzmodell, welches die fachliche, motivationale, pädagogisch-didaktische und persönliche Dimension heraushebt (KMK, 2004).

Fragestellungen: Dieser Beitrag thematisiert am Beispiel der TUM School of Education den Einsatz und die Wirkung von beratenden Auswahlgesprächen. Dabei wird den folgenden Fragen nachgegangen: (1) Welche Gesprächsinhalte werden in den Auswahlgesprächen diskutiert? (2) Wie werden die Auswahlgespräche von den Bewerberinnen und Bewerbern wahrgenommen? Die Studie ist Teil einer Dissertation über die Möglichkeiten und Grenzen von Auswahlgesprächen als Beratungstool im Kontext der Zulassung zum Lehramtsstudium.

Methode: Im Zuge des Auswahlverfahrens für den gymnasialen Lehramtsstudiengang „Naturwissenschaftliche Bildung“ zum Wintersemester 2011/12 wurden an der TUM School of Education N=22 Auswahlgespräche auf Tonband aufgenommen und inhaltsanalytisch ausgewertet. Neben den diskutierten Gesprächsinhalten, wurde erfasst welche Bewertungsindikatoren von den Kommissionen herangezogen wurden und welche Art von Rückmeldung die Bewerberinnen und Bewerber erhalten haben. Auf struktureller Ebene wurden Ablauf und Länge verschiedener Gesprächsteile, die Ausdrucksart und der Fragenbezug erfasst. Die Gesprächsatmosphäre und Wirkung der Auswahlgespräche wurden per Fragebogen (vgl. Mayr, 2002) erfasst (N=79). Deskriptive und korrelative Analysen wurden durchgeführt.

Ergebnisse: Die Ergebnisse zeigen, dass die vier Dimensionen des Kompetenzmodells sowohl im Fragen- als auch im Bewertungsteil der Auswahlgespräche abgebildet werden können. Die genauen Gesprächsinhalte variieren jedoch zwischen verschiedenen Kommissionen und Bewerberinnen bzw. Bewerbern. Im fachlichen Bereich (39,3%) werden insgesamt die häufigsten Fragen gestellt. Dazu gehören z.B. Fragen nach fachlichen Interessen, schulischen Leistungen und Erwartungen an das Fachstudium. Im motivationalen Bereich (17,6%) werden unter anderem Studien- und Berufswahlmotive sowie die genaue Ausrichtung der Motivation (z.B. Schulart) diskutiert. Auch pädagogisch-didaktische Themen (32,1%), wie z.B. die Reflexion von pädagogischen Vorerfahrungen und Vorstellungen vom Unterrichten, werden ausführlich behandelt. Im persönlichen Bereich (10,5%) wird insbesondere die psychische Belastbarkeit angesprochen. Das Feedback in den Auswahlgesprächen erfolgt vor allem entwicklungsfördernd, positiv und informierend. Die Bewerberinnen und Bewerber bewerten die Gesprächsatmosphäre und die Wirkung der Auswahlgespräche insgesamt sehr positiv. Dabei unterscheiden sich die Bewertungen unterschiedlicher Kommissionsgruppen nicht signifikant. Es konnten Gesprächsmerkmale identifiziert werden, die mit der Bewertung des Auswahlgesprächs korrelieren (z.B. Dauer des Feedbacks und reflexive Wirkung des Auswahlgesprächs;  $r = .511$ ).

Diskussion: Insgesamt sprechen die Ergebnisse dafür, dass Auswahlgespräche ein vielversprechender Ansatz für eine individuelle Eignungsberatung sind, die von Bewerberinnen und Bewerbern geschätzt wird. Es bleibt zu prüfen, wie Auswahlgespräche bei unterschiedlichen Beratungsbedürfnissen optimal eingesetzt werden können. Dabei geraten zunehmend weiterführende Konzepte mit dem Ziel der Anregung eines kontinuierlichen Reflexionsprozesses in den Fokus.

Kultusministerkonferenz, KMK (2004). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland: Standards für die Lehrerbildung: Bildungswissenschaften.

Mayr, J. (2002). Qualitätssicherung durch Laufbahnberatung: Zur Rolle von Selbsterkundungs-Verfahren. In: Brunner, H., Mayr, E., Schratz, M., & Wieser, I. (Hrsg.): Lehrerinnen- und Lehrerbildung braucht Qualität. Und wie!? Innsbruck: Studienverlag, S. 413-434.

Schaarschmidt, U. (2005). Halbtagsjobber? Psychische Gesundheit im Lehrerberuf – Analyse eines veränderungsbedürftigen Zustandes. Weinheim, Basel: Beltz.

Birgit Brouër (Christian-Albrechts-Universität zu Kiel), Anna Zaunbauer (Christian-Albrechts-Universität zu Kiel)

***Studienwahl Lehramt – Entscheidungsprozess und unterstützende Maßnahmen***

---

**GEBF** TAGUNG  
2014

**B03 d**

**Symposiums-  
beitrag**

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G168

**Self-Assessments als Instrumente der Studienberatung**

2

## **Validierung von Messinstrumenten zur Erfassung naturwissenschaftlicher Kompetenzen**

Chair(s): Olaf Köller (IPN Kiel), Nele Kampa (IPN Kiel)

Diskutant/in: Andreas Frey (Friedrich-Schiller-Universität Jena)

In dem Symposium werden Messinstrumente in den Naturwissenschaften in nationalen und internationalen Studien aus unterschiedlichen Perspektiven betrachtet. Daher besteht das Symposium aus zwei pädagogisch-psychologischen und zwei fachdidaktischen Beiträgen und greift verschiedenste methodische Herangehensweisen auf. Die Beiträge führen von einer wissenschaftlichen Betrachtung der Testaufgaben (Beiträge Wagner und Kauertz) über Fragen der Dimensionalität von Testungen in den Naturwissenschaften (Beitrag Walpuski) zur Validierung der Erstellung eines Kompetenzstufenmodells (Beitrag Kampa) und bieten somit Validitätsargumente vom Beginn bis zum „Ende“ der Testentwicklung.

Im ersten Beitrag wird die Inhalts- und Konstruktvalidität von Testungen naturwissenschaftlicher Kompetenz in Deutschland anhand von drei Large-Scale Assessments geprüft. Die Testaufgaben der PISA-Studie, von NEPS und den Bildungsstandards werden mittels Experteneinschätzungen mit Blick auf die Testkonstruktion verglichen. Es wird der Frage nachgegangen, inwiefern die unterschiedlichen Tests auch unterschiedliche Facetten naturwissenschaftlicher Kompetenz messen. Der Vergleichbarkeit der Test wird des Weiteren über umfangreiche Dimensionalitätsanalysen nachgegangen.

Im zweiten Beitrag werden kontextbeschreibende Aufgabenmerkmale von kompetenzorientierten Testaufgaben systematisch variiert und somit dessen Einfluss auf die Aufgabenschwierigkeit ermittelt. Inhaltsbezogene und situationsbezogene sowie sprachliche Merkmale werden sowohl geratet als auch mittels IRT-Modelle näher untersucht.

In einem weiteren Beitrag werden sowohl durch qualitative als auch quantitative Analysen die Testaufgaben zum Kompetenzbereich Kommunikation der Bildungsstandards in den Naturwissenschaften validiert. In den qualitativen Analysen werden die Prozesse bei der Aufgabenlösung näher untersucht. In den quantitativen Analysen liegt das Augenmerk auf der Abgrenzung zu weiteren Kompetenzbereichen und zu sprachlichen Fähigkeiten.

Im abschließenden Beitrag wird ein Angoff-Verfahren zur Erstellung eines Kompetenzstufenmodells für naturwissenschaftliche Kompetenz in der 13. Jahrgangsstufe in Schleswig-Holstein validiert. Erste Ergebnisse werden zu prozeduralen und internen Validitätsargumenten geliefert. Zum Zeitpunkt der Tagung werden zudem Daten zu zwei Messzeitpunkten vorliegen, sodass die Validitätsargumente Konsequenzen und external mit einbezogen werden kann.

Die beschriebenen Beiträge vereinen sowohl unterschiedliche Forschungsmethoden als auch Interdisziplinarität (Erziehungswissenschaft/Psychologie und Didaktik der Naturwissenschaften) und zeichnen somit ein umfassendes Bild der Erfassung von naturwissenschaftlicher Kompetenz.



Helene Wagner (IPN Kiel), Katrin Schöps (IPN Kiel), Inga Hahn (IPN Kiel), Marcus Pietsch (Leuphana Universität Lüneburg), Silke Rönnebeck (IPN Kiel)

## ***PISA, Bildungsstandards und das Nationale Bildungspanel (NEPS). Vergleich der Rahmenkonzepte und Validierung des NEPS-Testinstruments in den Naturwissenschaften***

Einführung: Die vorliegende Studie beschäftigt sich mit der Validierung des Naturwissenschaftstests des Nationalen Bildungspanels (NEPS) für die neunte Jahrgangsstufe. Zusätzlich zur Validierung des Tests besteht unser Anliegen in der Überprüfung der Vergleichbarkeit der Leistungstests, die in Deutschland in der neunten Klassenstufe in den Naturwissenschaften bundesweit eingesetzt werden. Aus diesem Grund überprüfen wir die Validität des NEPS-Naturwissenschaftstests anhand der nationalen Tests zur Erfassung der Bildungsstandards in den Fächern Biologie, Chemie und Physik sowie anhand des Naturwissenschaftstests des Programme for International Student Assessment (PISA). Zur Verwirklichung dieser Zielstellung analysieren wir die Äquivalenz der Testkonzeptionen und Kompetenzmodelle des NEPS und der Testkonzeptionen und Kompetenzmodelle von PISA und der nationalen Bildungsstandards (BiSta) in den Fächern Biologie, Chemie und Physik. Das Ziel dieser Analysen besteht darin, die Kompetenzskalen für Naturwissenschaften im NEPS auf die Studien PISA und die Prüfung der nationalen Bildungsstandards zu beziehen. Somit lassen sich künftig beispielsweise die Ergebnisse der Erhebungen im NEPS an einem internationalen Referenzmaßstab abbilden oder in den nationalen Standards verorten.

Fragestellungen: 1. Konzeptionelle Äquivalenz: Inwieweit ist die theoretische NEPS-Testkonzeption für die Domäne Naturwissenschaften vergleichbar mit der Testkonzeption der naturwissenschaftlichen Grundbildung des PISA und der Erfassung der nationalen Bildungsstandards in den Fächern Biologie, Chemie und Physik?

2. Dimensionale Äquivalenz: Inwieweit ist die faktorielle Struktur des latenten Konstrukts „naturwissenschaftliche Kompetenz“ im NEPS vergleichbar mit der faktoriellen Struktur der naturwissenschaftlichen Grundbildung des PISA-Tests aus 2012 und der Tests zur Erfassung der nationalen Bildungsstandards in den Fächern Biologie, Chemie und Physik?

Methoden :

Konzeptionelle Äquivalenz: Zur Untersuchung der konzeptionellen Äquivalenz wurden Ratingskalen entwickelt, auf welchen die Items des NEPS-Naturwissenschaftstests hinsichtlich der Passung zu den Testkonzeptionen der Vergleichsstudien eingeschätzt werden können. Für diese Einschätzung wurden sieben Expertinnen und Experten gewonnen, die mit den Large-Scale-Assessments in Deutschland vertraut sind. Zur Auswertung der Ergebnisse des Experten-Reviews wird die Generalisierbarkeitstheorie (Cronbach et al., 1972) angewendet. Die einzelnen Varianzquellen (Rater, Items, Rater x Items) und deren Gewichte werden hier empirisch bestimmt. Vorteil dieses Ansatzes im Vergleich zu den Alternativverfahren besteht in einer tieferen Analyse der Fehlervarianz. Zur Einschätzung der konzeptionellen Äquivalenz wird ein Generalisierbarkeitskoeffizient berechnet und die Zuordnung der Items zu Testdimensionen beschrieben.

Dimensionale Äquivalenz: Zur Untersuchung der dimensional Äquivalenz werden die Daten der Studien sowohl getrennt als auch zusammen nach Modellen der Item-Response-Theorie skaliert. Zur Beantwortung der Fragestellung werden folgende Korrelationen näher betrachtet: (a) Korrelationen der Personenfähigkeiten, die durch die gemeinsame IRT-Skalierung der Testscores aus beiden Tests ermittelt wurden, (b) Korrelationen der Teildimensionen innerhalb der Tests, (c) Korrelationen der Teildimensionen zwischen den Tests.

Ergebnisse: Die Auswertung des Experten-Reviews zeigte eine ausreichende Interrater Reliabilität auf allen Dimensionen und keine Verzerrung durch den systematischen rater bias. Die meisten NEPS-Items konnten den Kategorien der PISA- bzw. BiSta- Rahmenkonzeption zugeordnet werden. Allerdings zeigte das Experten-Review, dass nicht alle Kategorien der PISA- und BiSta-Rahmenkonzeption durch NEPS-Items abgedeckt wurden.

Bei der Untersuchung der dimensional Äquivalenz der Studien wurde eine hohe Korrelation der Tests PISA und BiSta mit NEPS festgestellt. Die Prüfung auf die Eindimensionalität zeigte eine bessere Passung des zweidimensionalen Modells im Vergleich zum eindimensionalen Modell. Dagegen konnte die Untersuchung der partiellen Korrelationen der Residuen von PISA und NEPS Items bzw. von BiSta und NEPS Items nach der Methode von Huynh die Geltung der Annahme der lokalen stochastischen Unabhängigkeit der Rasch-Skalierung bestätigen.

Im Symposium werden Analysen und Ergebnisse zur Untersuchung der konzeptionellen und dimensional Äquivalenz vorgestellt und diskutiert.

Bühner, M. (2010). Einführung in die Test- und Fragebogenkonstruktion (3. Aufl.). München: Pearson-Education.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley.

Huynh, H., Michaels, H. R. & Ferrara, S. (1995). A comparison of three statistical procedures to identify clusters of items with local dependency. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

## **Der Zusammenhang zwischen Aufgabenkontext und Aufgabenschwierigkeit - Ergebnisse aus der Domäne Physik**

### Theoretischer Hintergrund

Physik-Kompetenzmodelle beschreiben Aufgabenmerkmale (Schecker & Parchmann, 2006), die Grundlage für die Aufgabenkonstruktion in Tests sind (Kauertz et al., 2010).

Naturwissenschaftliche Grundbildung umfasst die Anwendung von Physikwissen und Arbeitsweisen zur Lösung von Problemen, die in Natur und Technik durch menschliches Handeln entstehen oder dieses betreffen (OECD, 1999). Kontexte sind Situationen, die Anlass zu solcher physikalischer Analyse bieten. Sie werden bisher in der naturwissenschaftsdidaktischen Forschung hinsichtlich ihres motivierenden Potenzials und ihrer Lernwirksamkeit betrachtet (vgl. Bennett, Lubben, & Hogarth, 2007).

Zur Messung von Physikkompetenz ist eine Operationalisierung von Kontext durch Aufgabenmerkmale nötig. Eine höhere Physikkompetenz drückt sich dann u.a. in der Fähigkeit aus, Probleme zu lösen, die in eine Situation eingebettet sind, die eine physikalische Analyse notwendig macht, dies aber nicht unmittelbar erkennbar ist. Mit einer geringeren Physikkompetenz wird dasselbe Problem nur in einer Situation gelöst, in der dies deutlich erkennbar ist. Aufgabenmerkmale, die Aspekte von Kontexten abbilden, lassen sich logisch nach Merkmalen der Situation, Merkmalen des Inhalts und Merkmalen, die den Bezug zwischen beiden beschreiben, unterscheiden (Dorschu, 2013).

Aufgabenmerkmale nach Dorschu (2013)

Situationsbezogen (affektiv): Situationales Interesse (Krapp, 1999), Bekanntheit (van Vorst, 2012), Glaubwürdigkeit (in Anlehnung an Kuhn, 2010)

Situationsbezogen (Oberfläche): Themenfeld (Alltag, Technik, Gesellschaft, Sport), Textliche Darstellung (Lesbarkeit [Amstad, 1978])

Inhaltsbezogen: Inhaltsbereich (Mechanik, Elektrizitätslehre, Wärmelehre)

Inhalts-Situations-bezogen: Sprachähnlichkeit (Fachsprache vs. Alltagssprache+Füllsätze)

### Fragestellung

Der Zusammenhang zwischen Aufgabenmerkmalen und Physikkompetenz entscheidet sich an der Frage, ob Aufgabenmerkmale die Aufgabenschwierigkeiten beeinflussen. Eine höhere Schwierigkeit bedeutet bei sonst gleichen Bedingungen eine höhere Kompetenz. Untersucht wird daher die Frage: Welchen Einfluss haben kontextbeschreibende Aufgabenmerkmale auf die Aufgabenschwierigkeit?

Methode: Zur Untersuchung dieser Frage wurden insgesamt 12 Situationen und zu vier der Situation jeweils 15 Aufgaben entwickelt, die sich in inhalts-situations-bezogenen und inhaltsbezogenen Merkmalen unterscheiden. Weitere Kompetenzdimensionen (vgl. Kauertz et al., 2010) wurden konstant gehalten. Die Untersuchung erfolgte in zwei Teilen. Im ersten Teil wurden affektive Merkmale der Situationen untersucht, indem die Situationsbeschreibungen von N = 518 Schülerinnen und Schülern (Alter: 15,2 Jahre; weiblich: 50,3%) hinsichtlich dieser Merkmale geratet wurden. Zur Skalierung wurde ein partial-credit-Rasch-Modell genutzt. Die dreidimensionale Modellstruktur wurde durch Vergleiche unterschiedlich genesteter Modelle, explorativer Faktorenanalyse und Korrelationen zwischen den Skalen untersucht. Im zweiten Teil wurden Aufgaben zu vier Situationen mit gleichen Mittelwerten im Rating N = 829 Schülerinnen und Schülern (Alter: 15,8 Jahre; weiblich: 39,9%) vorgelegt und das Ergebnis dichotom (richtig/falsch) kodiert. Die Aufgaben wurden im Multimatrix-Design administriert und mit einem dichotomen, eindimensionalen Raschmodell skaliert. Die Itemparameter wurden hinsichtlich der Aufgabenmerkmale varianzanalytisch untersucht. Begleitend wurde eine Validierungsstudie durchgeführt.

Ergebnisse: Die Skalierung in beiden Teilen zeigte eine hohe Passung für die meisten Aufgaben (Infit/Outfit: MNSQ > 1.3; ZSTD > 2.0; 4 Items ausgeschlossen) und akzeptable Reliabilitäten (Teil 1: .89 >  $\alpha$  > .85; Teil 2:  $\alpha$  = .69). Die affektiven situationsbezogenen Aufgabenmerkmale lassen sich durch Vergleich der Deviance pro Freiheitsgrad (168.6 dreidimensional vs. 176.1 für eindimensional), hohe Faktorladungen (> .59) auf die intendierten Skalen und geringe Korrelationen ( $r < .39$ ) empirisch unterscheiden. Zwischen den Situationen zeigen sich signifikante Unterschiede im Rating der Aufgabenmerkmale ( $p < .001$ ;  $.43 < \eta^2_{part} < .73$ ). Im zweiten Teil lässt sich die Validität durch begleitende Tests und Fragebögen nachweisen (vgl. Dorschu, 2013). Die Sprachähnlichkeit als inhalts-situations-bezogenes Aufgabenmerkmal hat keinen Einfluss auf die Aufgabenschwierigkeit. Für das Aufgabenmerkmal Inhaltsbereich lässt sich ein kleiner Effekt nachweisen ( $\eta^2_{part} = 0.113$ ;  $p < .001$ ).

Schlussfolgerung: Die Ergebnisse zu den affektiven situationsbezogenen Aufgabenmerkmalen zeigen einen ähnlichen Effekt wie er in älteren Kontextprojekten erzielt wurde. Mit der Sprachähnlichkeit ist es nicht gelungen, die schwierigkeitsgenerierende Wechselwirkung zwischen Situation und Inhalt abzubilden. Neuere Untersuchungen weisen darauf hin, dass eine Fokussierung auf tieferliegende Modellbildungsprozesse eine höhere Erklärungsmächtigkeit haben könnte (Löffler & Kauertz, in Vorb.).

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, 91(3), 347–370. doi:10.1002/sce.20186

Dorschu, A. V. (2013). Die Wirkung von Kontexten in Physikkompetenztestaufgaben. Berlin: Logos Verlag.

Kauertz, A., Fischer, H. E., Mayer, J., Sumfleth, E. & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 135-153.

Löffler, P. & Kauertz, A. (in Vorb.). Applying physics models in context-based tasks in physics education. Proceedings zur ESERA-Tagung in Nikosia.

OECD [ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT] (1999). Measuring student knowledge and skills. A New Framework for Assessment. Paris: OECD Publication Service.

Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. In: *Zeitschrift für Didaktik der Naturwissenschaften* 12 (2006), S. 45-66.



## Validierung von Testaufgaben im Kompetenzbereich Kommunikation am Beispiel der Chemie

Bei der Evaluation der Bildungsstandards für das Fach Chemie werden vier Kompetenzbereiche unterscheiden: Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung. Für die Evaluation werden die Bereiche separat operationalisiert und erfasst. Im Rahmen dieses Beitrages werden ergänzende Validierungsmaßnahmen zum Kompetenzbereich Kommunikation vorgestellt sowie eine Untersuchung zur empirischen Trennbarkeit der Kompetenzbereiche. Der Aufgabenkonstruktion wird ein dreidimensionales Kompetenzstrukturmodell zugrunde gelegt, das sich bereits für die Aufgabenkonstruktion bewährt hat und das sich in seinen wesentlichen Annahmen auch empirisch bestätigen ließ (Kauertz et al. 2010). Als schwierigkeitsbestimmende Dimensionen werden die Komplexität der Aufgaben und die zur Lösung notwendigen kognitiven Prozesse angenommen. Die dritte Dimension ist eine inhaltliche Dimension, auf der Kompetenzbereiche und deren Untergliederungen in Kompetenzteilbereiche abgetragen werden, um so fachspezifische Ausprägungen abbilden zu können. Da die Bildungsstandards für den Kompetenzbereich Kommunikation keine weitere Untergliederung vorsehen, wurde diese Definition für die Evaluation der Standards auf Basis verschiedener Kommunikationstheorien und naturwissenschaftsdidaktischen Forschungsarbeiten vorgenommen. Unter Kommunikationskompetenz wird in den Bildungsstandards zunächst allgemein das sach- und fachbezogene Erschließen und Austauschen von Informationen verstanden (KMK, 2005). Aus gängigen Kommunikationstheorien und -modellen (z. B. Blumer, 1973, Habermas, 1971, Rautenstrauch, 2008, Schulz von Thun, 1981) lassen sich wichtige Kennzeichen der Kommunikation die Übermittlung von Informationen, die Nutzung von Zeichen und Symbolen, die Aufnahme von Informationen, die empfängerabhängige Informationsaufnahme und -interpretation sowie ein argumentativer Charakter ableiten. Diese Merkmale sind weitgehend deckungsgleich mit den Anforderungen der Standards, sodass folgende Kompetenzteilbereiche als Untergliederung des Kompetenzbereiches Kommunikation angenommen wurden: (1) Erschließen von Informationen (2) Weitergeben von Informationen und (3) Argumentieren (Kobow et al., 2011). In einem ersten Schritt wurde mit dem Think-Aloud-Verfahren und Interviews an einer kleinen Stichprobe (N = 12) geprüft, ob die angenommenen Prozesse bei der Aufgabenbearbeitung überhaupt angesprochen werden (Gehlen, 2012). Die Daten wurden mithilfe der qualitativen Inhaltsanalyse nach Mayring (2008) ausgewertet. Die Auswertung zeigt, dass Aufgaben in der Regel dann nicht korrekt gelöst werden können, wenn die vorausgesetzten kognitiven Prozesse oder die intendierte Komplexität sich in der Bearbeitung nicht widerspiegeln. Zusätzlich wurden Testaufgaben zur Kommunikationskompetenz im Fach Chemie mit Hilfe von Tests zum Fachwissen, zur Bewertungskompetenz sowie zur Sprachfähigkeit divergent validiert (N=496). Diese Testinstrumente wurden ausgewählt, da sie die größte Nähe zur Kommunikationskompetenz aufweisen sollten. Die Nähe zwischen Fachwissen und Kommunikation wird erwartet, da in den Standards zu beiden Bereichen der Umgang mit Informationen enthalten ist. Es werden z. B. teilweise die gleichen Operatoren (erklären, beschreiben) genutzt. Darüber hinaus ist das Lernen einer Fachsprache (in den Bildungsstandards im Bereich Kommunikation verankert) untrennbar mit dem Lernen von Fachwissen verknüpft (Merzyn, 2008). Die Nähe zwischen den Bereichen Bewertung und Kommunikation resultiert aus dem Umstand, dass Argumentieren nicht nur eine kommunikative Fähigkeit ist, sondern oft auch die Bewertung eines Sachverhaltes mit einschließt. Ebenso kann angenommen werden, dass die allgemeine Sprachkompetenz ähnlich, aber nicht deckungsgleich zur fachlichen Kommunikationskompetenz ist. Die Auswertung erfolgte im Rahmen der probabilistischen Testtheorie mit Hilfe der IRT-Software Conquest. Ergebnisse für den Kompetenztest Kommunikation zeigen zunächst akzeptable Kennwerte für den Test (Item Separation Reliability: .976; EAP/PV Reliability: .711.). Außerdem passt das dreidimensionale Modell basierend auf den Kompetenzteilbereichen besser zu den Daten als das eindimensionale Modell (Deviance, AIC, BIC, Likelihoodquotiententest). Dies gilt auch, wenn man eine gemeinsame und getrennte Skalierung der Kompetenzbereiche vergleicht. Auch hier erklärt ein mehrdimensionales Modell die Daten besser.

Blumer, Herbert (1973). Der Methodologische Standort des Symbolischen Interaktionismus. In: J. Matthes (Hrsg.): Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit. 5. Aufl. Opladen: Westdeutscher Verlag, S. 80-146.

Gehlen, C. (2012). Validierung eines schriftlichen Tests zur Kommunikationskompetenz im Fach Chemie mittels Interviews. Hausarbeit zur Erlangung des 1. Staatsexamens für das Lehramt an Gymnasien und Gesamtschulen. Universität Duisburg-Essen.

Habermas, Jürgen (1971): Vorbereitende Bemerkungen zu einer Theorie des kommunikativen Kompetenz. In: J. Habermas und N. Luhmann (Hrsg.): Theorie der Gesellschaft oder Sozialtechnologie. Was leistet Systemforschung? Frankfurt am Main: Suhrkamp, S. 101-141.

Kauertz, A., Fischer, H., Mayer, J., Sumfleth, E., & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe. Zeitschrift für Didaktik der Naturwissenschaften 16, 135-153.

KMK / Sekretariat der Ständigen Konferenzen der Kultusministerkonferenz der Länder in der Bundesrepublik Deutschland (2005). Beschlüsse der Kultusministerkonferenz Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss.

Kobow, I., Fischer, H. E., Hostenbach, J., Kauertz, A., Klebba, N., Klöpfel, K., Kremer, K., Mayer, J., Pant, H. A., Schwanewedel, J., Sumfleth, E., Walpuski, M. & Wächter, M. (2011). Evaluation der Standards in den Fächern Biologie, Chemie und Physik für die Sekundarstufe I. Band 7: Aufgabenkonstruktionsanleitung - Aufgabenentwicklung in den Fächern Biologie, Chemie und Physik für den Kompetenzbereich „Kommunikation“.

Mayring, P. (2008). Qualitative Inhaltsanalyse. Grundlagen und Techniken (10. Auflage). Weinheim: Beltz Verlag.

Merzyn, G. (2008). Sprache und Chemie lernen. Naturwissenschaften im Unterricht Chemie, 19 (106-107), 94-97.

Rautenstrauch, Christina (2008): Das Sprachspiel des Online-Tutoring. Zur Unterstützung von Wissenskommunikation in lernenden Online-Gemeinschaften. Eine Analyse im Kontext beruflicher Weiterbildung. Online verfügbar [http://duepublico.uni-duisburg-essen.de/servlet/DerivateServlet/Derivate-21269/ChRautenstrauch\\_151208.pdf](http://duepublico.uni-duisburg-essen.de/servlet/DerivateServlet/Derivate-21269/ChRautenstrauch_151208.pdf) , zuletzt geprüft am 23.10.2013.

Schulz von Thun, Friedemann (1981): Miteinander reden 1. Störungen und Klärungen. Reinbeck: Rowohlt.



## Validierung eines Angoff-Standard-Setting Verfahrens. Traditionelle und innovative Argumente

Kompetenzstufenmodelle und deren Validität: Kompetenzstufenmodelle sind wichtiges Kommunikationsmittel für die politische und interessierte Öffentlichkeit. Ihre Generierung wird als eine der wichtigsten Aufgaben in der Entwicklung von Testaufgaben gesehen (Cizek & Bunch, 2007). Kompetenzstufenmodelle entstehen in Standard-Setting Verfahren, welches durch die Setzung von einem oder mehreren cut-scores definiert wird (Cizek & Bunch, 2007). Eine kontinuierliche Schwierigkeitsskala wird in Stufen eingeteilt, welche unterschiedliche Leistungsniveaus einer Schülerpopulation abgrenzen sollen. Den Setzungen der Kompetenzstufen bzw. post-hoc Beschreibungen der Stufen liegen subjektive Beurteilungen von Expertinnen und Experten als auch technische, politische und empirische Abwägungen zu Grunde (AERA/APA/NCME, 1999). Sie befinden sich somit an der Schnittstelle zwischen beweisbaren Aspekten der Validierung und Aspekten der Konsequenzen (Pant et al., 2009). Eine Validierung dieses Verfahrens ist daher sowohl von immenser Wichtigkeit als auch besonders herausfordernd. Diese Validierung kann Evidenzen für die „interpretativen“ Argumente liefern, die die Angemessenheit und Haltbarkeit der Cut-Scores stützen können (Sireci et al., 2009).

Validierung eines Standard-Setting Verfahrens: In der vorliegenden Studie wurde die Two-Choice Variante (Nassif, 1978) bzw. die Extended Variante (Cizek, Bunch, & Koons, 2004) des Angoff Verfahrens zur Setzung von vier Kompetenzstufen in den Naturwissenschaften in der 13. Jahrgangsstufe angewendet (Angoff, 1971). Grundlage für das Verfahren sind Leistungsdaten in den Naturwissenschaften von 3775 Schülerinnen und Schülern in allgemeinen und beruflichen Gymnasien in Schleswig-Holstein. Neun Expertinnen und Experten (Vertreter aus bildungspolitischen Institutionen in Schleswig-Holstein sowie aus den Fachdidaktiken Biologie, Chemie und Physik) nahmen im Oktober 2013 an einer Standard-Setting Tagung teil. Das zu generierende Kompetenzstufenmodell sollte Kompetenzen zur Erlangung verschiedener naturwissenschaftsferner und naturwissenschaftsnaher Berufe und Studiengänge diskriminieren können. Die Expertinnen und Experten mussten die vier Kompetenzstufen auf Grundlage von 30 Multiple-Choice (MC) bzw. Complex-Multiple-Choice (CMC) Items festlegen. In der ersten Rating-Runde gaben die Expertinnen und Experten individuelle Einschätzungen ab, in der zweiten Runde wurde nach weiteren individuellen Einschätzungen eine Konsensualisierung herbeigeführt. In der dritten Runde wurden anhand der Items auf den unterschiedlichen Kompetenzstufen Niveaubeschreibungen generiert. Vor und während der Tagung wurden die Rating-Ergebnisse aus den jeweiligen Standard-Setting Runden gesammelt und insgesamt vier Fragebögen an die Expertinnen und Experten administriert. Zudem liegen Daten der Schülerinnen und Schüler zu deren Abiturprofilen, Noten in den naturwissenschaftlichen Fächern, berufliche Interessen und Selbsteinschätzungen vor. Des Weiteren werden zukünftige Daten zeigen (Januar 2013), welche Berufe und Studienfächer die Schülerinnen und Schüler ca. vier Monate nach dem Abitur gewählt haben. Anhand dieser Daten können Validierungsargumente in den Kategorien prozedural, internal, external und Konsequenzen gesammelt werden (Kane, 1994; Pant et al., 2009). Insbesondere konsequenzbezogene Argumente werden selten in die Validierung von Kompetenzstufenmodellen bzw. Standard-Setting Verfahren einbezogen.

Erste Ergebnisse können den prozeduralen und internen Validierungsargumenten zugeordnet werden. Bezogen auf die prozeduralen Validierungsargumente gaben die Expertinnen und Experten an, dass sie ihren Arbeitsauftrag verstanden haben. Jedoch fühlten sie sich unter Zeitdruck und hatten Probleme damit, sich die Schülerpopulation auf den jeweiligen Kompetenzstufen vorzustellen und waren sich nur mäßig sicher bei der Setzung der cut-scores.

Bezogen auf die internen Validierungsargumente wird deutlich, dass die Expertinnen und Experten die Leistungen der Schülerinnen und Schüler zunächst stark überschätzt und dieses neue Wissen teilweise genutzt haben. In der zweiten Ratingrunde wurden die Ratings der Kompetenzstufen I bis III statistisch signifikant verringert. Die Fachdidaktikerinnen und Fachdidaktiker unterschieden sich in ihren Ratings nicht von den Vertreterinnen und Vertretern aus den bildungspolitischen Institutionen. Weitere Analysen werden die externalen und konsequenzbezogenen Validierungselemente aufgreifen. So kann die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen mit den belegten Wochenstunden in den Naturwissenschaften und Abiturprofilen in Bezug gesetzt werden. Die Daten zum tatsächlich erlangten Beruf/Studium der Schülerinnen und Schüler liegen Anfang des kommenden Jahres vor und können Aufschluss über die Vorhersagekraft des Kompetenzstufenmodells für den beruflichen Werdegang geben.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). Standards for Educational and Psychological Testing. Washington DC, American Educational Research Association.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 508-600). Washington, DC: American Council on Education.

Cizek, G. J., & Bunch, M. B. (2007). Standard setting: A guide to establishing and evaluating performance standards on tests. Thousand Oaks, CA: Sage.

Cizek, G. J., Bunch, M. B., & Koons, H. (2004). Setting performance standards: Contemporary methods. Educational Measurement: Issues and Practice, 23(4), 31-50.

Kane, M. T. (1994). Validating the performance standards associated with passing scores. Review of Educational Research, 64, 425-461.

Nassif, P. M. (1978). Standard-setting for criterion-referenced teacher licensing tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Toronto.

Pant, A. H., Rupp, A. A., Tiffin-Richards, S. P., Köller, O. (2009). Validity issues in standard-setting studies. Studies in Educational Evaluation, 35, 95-101.

Sireci, S. G., Hauger, J., Wells, C. S., Shea, C., & Zenisky, A. L. (2009). Evaluation of the standard setting on the 2005 Grade 12 National Assessment of Educational Progress mathematics test. Applied Measurement in Education, 22, 339-358.



Kathrin Kaub (Universität des Saarlandes), F.-Sophie Wach (Universität des Saarlandes), Roland Brünken (Universität des Saarlandes)

## Lehrerkompetenzen in Pädagogik und Fachdidaktik

Chair(s): Kathrin Kaub, F.-Sophie Wach, Roland Brünken (alle: Universität des Saarlandes)

Diskutant/in: Hermann Josef Abs (Universität Duisburg-Essen)

Es wird allgemein angenommen, dass ein guter Lehrer über ein umfassendes pädagogisches und fachdidaktisches Wissen verfügen muss. Eine breite Wissensbasis wird demnach als eine wesentliche Voraussetzung zur Bewältigung der beruflichen Anforderungen betrachtet (Baumert & Kunter, 2006). Gleichzeitig wird angenommen, dass sich das Professionswissen in verschiedene Wissensbereiche (pädagogisches bzw. bildungswissenschaftliches Wissen, fachdidaktisches und fachwissenschaftliches Wissen) unterteilen lässt, die wiederum jeweils mehrdimensionale Konstrukte abbilden. Bislang relativ gut untersucht sind die Wissensstrukturen im Bereich des pädagogischen bzw. bildungswissenschaftlichen Wissens und der Mathematikdidaktik. Die Befunde aus der Mathematikdidaktik weisen u.a. darauf hin, dass neben dem Fachwissen insbesondere fachdidaktische Wissenskomponenten mit einer erfolgreichen Gestaltung von Unterricht in Zusammenhang stehen (Baumert et al., 2010). Inwiefern solche differentiellen Effekte auch in anderen Bereichen wie beispielsweise dem Fremdsprachenunterricht vorzufinden sind, ist bislang empirisch nicht überprüft. In diesen Bereichen stehen die Entwicklung und empirische Fundierung geeigneter Instrumente zur Erfassung fremdsprachendidaktischen Professionswissens noch am Anfang.

In dem vorliegenden Symposium soll einerseits ein vertiefender Einblick in die Struktur des bildungswissenschaftlichen sowie des fremdsprachen- und naturwissenschaftsdidaktischen Wissens (angehender) Lehrkräfte gegeben werden und andererseits die Zusammenhänge zwischen den einzelnen Wissenskomponenten mit anderen Kompetenzmerkmalen wie beispielsweise der professionellen Unterrichtswahrnehmung, den professionsbezogenen Überzeugungen sowie der Instruktionsqualität im Schulunterricht näher beleuchtet werden.

Ein erster Beitrag von Kunina-Habenicht et al. (Goethe Universität Frankfurt a.M.) beschäftigt sich im Rahmen der BiLWiss Studie mit der Validierung der mehrdimensionalen Struktur des *bildungswissenschaftlichen Wissens* und überprüft die Entwicklungsverläufe der verschiedenen unterrichtsnahen und -ferneren Wissenskomponenten über die schulpraktische Ausbildungsphase (Referendariat) hinweg.

Äquivalent untersucht ein zweiter Beitrag von Tachtsoglou, Roters, König (Universität zu Köln) & Nold (TU Dortmund) anhand der TEDS-LT Daten die Faktorenstruktur des *fremdsprachendidaktischen Wissens* im Fach Englisch und liefert damit eine empirische Fundierung zur Erfassung und Prüfung des fachdidaktischen Wissens angehender Englischlehrkräfte.

Steffensky, Fiebranz, Meschede (IPN Kiel) & Möller (WWU Münster) fokussieren in einem dritten Beitrag ausgehend von den Daten aus dem ViU Projekt verschiedene Komponenten der professionellen Kompetenz speziell im naturwissenschaftlichen Grundschulunterricht. Untersucht werden die Zusammenhänge zwischen dem *naturwissenschaftsdidaktischen Wissen*, der professionellen Wahrnehmung der Lernunterstützung sowie den Überzeugungen zum naturwissenschaftlichen Lehren und Lernen.

Ein vierter Beitrag von Biermann et al. (Universität des Saarlandes) stellt Videodaten aus dem SioS-L Projekt vor und verbindet das pädagogische Unterrichtswissen mit der videoanalytisch erfassten Instruktionsqualität im Schulpraktikum.

Mit den vorgestellten Themenschwerpunkten werden Beiträge aus den Disziplinclustern Lehrerbildung, Didaktik der Naturwissenschaften und der Fremdsprachendidaktik gebündelt. Die Beiträge werden hinsichtlich der Relevanz für das Forschungsfeld sowie möglicher Implikationen von Prof. Hermann Josef Abs (Universität Duisburg-Essen) diskutiert.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469-520.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. American Educational Research Journal, 47(1), 133-180.



Olga Kunina-Habenicht (Goethe-Universität Frankfurt), Hendrik Lohse-Bossenz (PädQUIS® gGmbH – Kooperationsinstitut der Freien Universität Berlin), Mareike Kunter (Goethe-Universität Frankfurt), Theresa Dicke (Universität Duisburg-Essen), Doris Holzberger (Staatsinstitut für Schulqualität und Bildungsforschung), Christina Linninger (Goethe-Universität Frankfurt), Franziska Schulze-Stocker (Westfälische Wilhelms-Universität Münster), Detlev Leutner (Universität Duisburg-Essen), Ewald Terhart (Universität Münster)

### **Empirische Struktur des bildungswissenschaftlichen Wissens und dessen Veränderung im Referendariat**

**THEORETISCHER HINTERGRUND:** Die Lehrerbildung in Deutschland ist in zwei Phasen strukturiert. Während in der universitären Phase vorwiegend theoretisches Wissen aufgebaut wird, steht im Referendariat die Vermittlung praxisbezogener Inhalte im Vordergrund. Unter bildungswissenschaftlichem Wissen (BW) werden fachspezifische universitäre Studieninhalte zusammengefasst, an deren Vermittlung i.d.R. mehrere Disziplinen (Erziehungswissenschaft, Psychologie und Soziologie) beteiligt sind. Die Anteile der Bildungswissenschaften variieren im Vergleich zur Gesamtstudiendauer je nach Bundesland und Universität (Bauer, Diercks, Rösler, Möller, & Prenzel, 2012; Terhart, Lohmann, & Seidel, 2010). Meist herrscht große Heterogenität an angebotenen Inhalten sowie hohe Wahlfreiheit seitens der Studierenden, welche sich auch in der Kritik an Bildungswissenschaften äußern, denen häufig fehlende Kumulativität und Beliebigkeit vorgeworfen werden (Lersch, 2006; Merzyn, 2002).

Da die Bedeutsamkeit verschiedener bildungswissenschaftlicher Themen meist nur innerhalb der Disziplinen diskutiert wird, liegt in der Literatur keine anerkannte Theorie zur Struktur von BW vor (Wilson, Floden, & Ferrini-Mundy, 2002). Im Rahmen des Projekts „Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung“ (BilWiss) konnte in einer Delphi-Befragung mit 49 Expert(e)innen aus verschiedenen Fachdisziplinen und Phasen der Lehrerbildung ein Konsens über 104 zentrale bildungswissenschaftliche Themen aus neun Inhaltsbereichen (u.a. Unterricht, Diagnostik, Bildungstheorie, Lehrerberuf) erzielt werden (Kunina-Habenicht et al., 2012). Aufbauend auf den Ergebnissen dieser Delphi-Studie wurde im Rahmen des BilWiss-Projekts ein neuer Wissenstest konstruiert, der alle relevanten Delphi-Themen berücksichtigt. Dieser Wissenstest geht über bestehende Verfahren zur Messung von pädagogisch-psychologischem Wissen (König & Blömeke, 2010; Voss, Kunter, & Baumert, 2011) hinaus, da er theoretisches deklaratives und konzeptuell-analytisches universitäres Wissen fokussiert, und daher auch nicht unmittelbar unterrichtsbezogene disziplinübergreifende Themengebiete (bspw. Bildungstheorie, Wissen über Lehrerberuf) berücksichtigt.

**FRAGESTELLUNG:** In diesem Beitrag werden zwei Forschungsfragen untersucht. Die erste bezieht sich auf die empirische Struktur bzw. Dimensionalität des BW, während die zweite die Veränderung des BW bzw. seine Stabilität im Verlauf des Referendariats fokussiert. Bezüglich der empirischen Struktur des bildungswissenschaftlichen Wissens (BW) ist anzunehmen, dass BW – aufgrund der hohen inhaltlichen und strukturellen Heterogenität in der Organisation der Bildungswissenschaften – empirisch kein einheitliches sondern mehrdimensionales Konstrukt darstellt. In Bezug auf Wissensentwicklung sind differentielle Entwicklungsprofile für eher unterrichtsferne und eher unterrichtsnahe Inhaltsbereiche denkbar, da im Referendariat (genauer in Studienseminarveranstaltungen) handlungsbezogenes unterrichtsnahes und seltener theoretisches Wissen thematisiert wird. Daher vermuten wir, dass in eher unterrichtsfernen Bereichen sich keine Veränderungen oder ggf. auch negative Leistungstendenzen zeigen, während positive Leistungszuwächse am wahrscheinlichsten für unterrichtsnahe Bereiche auftreten werden.

**METHODE:** Zur Beantwortung der Forschungsfragen werden längsschnittliche Daten aus dem Forschungsprojekt BilWiss mit zwei Messzeitpunkten herangezogen. Zum ersten Messzeitpunkt (t1) bearbeiteten 3273 Referendare (zum Referendariatsbeginn) im Frühjahr 2011 die lange Fassung des Wissenstests. Am Ende des Referendariats bearbeitete eine Teilstichprobe (n=267) zusätzlich im Winter 2013/14 (t2) eine kurze Fassung des Wissenstests. Die lange Fassung des Wissenstests beinhaltet 289 Items und wurde im Rotationsdesign eingesetzt, während die Kurzform 56 Items umfasst.

**ERGEBNISSE:** Empirisch konnte eine sechs-dimensionale Struktur ermittelt werden, wobei die Skalierung mittels eindimensionaler generalisierter Partial-Credit-Modelle (Muraki, 1992) erfolgte. Die sechs ermittelten Dimensionen (Unterrichtsdidaktik, Lernen/Entwicklung, Diagnostik, Schulpädagogik, Lehrerberuf, Bildungstheorie) weisen in der langen Testfassung ausreichend hohe Reliabilitäten auf (zwischen 0,65 und 0,76), während die Reliabilitäten im Kurztest z.T. etwas niedriger waren (zwischen 0,58 und 0,81).

Hinsichtlich der zweiten Fragestellung wurde eine vergleichsweise geringe Stabilität des Wissens festgestellt (Korrelationen zwischen den Leistungen in derselben Domäne zu t1 und t2 liegen zwischen 0,11 und 0,42). Es kann nicht ausgeschlossen werden kann, dass dies z.T. durch den Einsatz verschiedener Testversionen bedingt wird. Die Entwicklung des Wissens ist weitestgehend konform mit den Hypothesen: Leistungen in unterrichtsfernen Bereichen bleiben unverändert (Bildungstheorie, Diagnostik, Lehrerberuf), während unterrichtsnahe Domänen keine (Unterricht, Schulpädagogik) oder positive Entwicklungstendenzen (Lernen/Entwicklung ( $d=0,16$ )) aufweisen. Diese Ergebnisse werden diskutiert in Bezug auf methodische Herausforderungen und Implikationen für Lehrerbildung.

Bauer, J., Diercks, U., Rösler, L., Möller, J., & Prenzel, M. (2012). Lehramtsausbildung in Deutschland: Wie groß ist die strukturelle Vielfalt? *Unterrichtswissenschaft*, 40(2).

König, J., & Blömeke, S. (2010). Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung. Berlin: Humboldt-Universität.

Kunina-Habenicht, O., Kunter, M., Leutner, D., Terhart, E., Baumert, J., Lohse, H., . . . Schulze, F. (2012). Welche bildungswissenschaftlichen Inhalte sind wichtig in der Lehrerbildung? Ergebnisse einer Delphi-Studie. *Zeitschrift für Erziehungswissenschaft*, 15(4), 649–682. doi: 10.1007/s11618-012-0324-6

Lersch, R. (2006). Lehrerbildung im Urteil der Auszubildenden. Eine empirische Studie zu beiden Phasen der Lehrerausbildung. In C. Allemann-Ghionda (Ed.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern*. Zeitschrift für Pädagogik. Beiheft 51 (pp. 164–181). Weinheim u.a.: Beltz.

Merzyn, G. (2002). Stimmen zur Lehrerausbildung. Ein Überblick über die Diskussion. Hohengehren: Schneider Verlag.

Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement*, 16(2), 159–176. doi: 10.1177/014662169201600206

Terhart, E., Lohmann, V., & Seidel, V. (2010). Die bildungswissenschaftlichen Studien in der universitären Lehrerbildung. Eine Analyse aktueller Studienordnungen und Modelhandbücher an Universitäten in Nordrhein-Westfalen.

Unveröffentlichter Bericht. Institut für Erziehungswissenschaft. Westfälische Wilhelms-Universität Münster.

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical and psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. doi: 10.1037/a0025125

Wilson, S., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education*, 53(3), 190–204. doi: 10.1177/0022487102053003002



## **Faktorenanalytische Prüfung des fachdidaktischen Wissens angehender Englischlehrkräfte**

Auf theoretisch-konzeptioneller Ebene ist das Konstrukt des fremdsprachendidaktischen Professionswissens gut entwickelt. Im Gegensatz zu anderen Domänen (z.B. Mathematik), steht jedoch die empirische Fundierung im Bereich Fremdsprachenlernen und -lehren noch am Anfang (Trautmann, 2010; Nold, 2013). An dieser Stelle setzt die Studie Teacher Education and Development Study: Learning to Teach (TEDS-LT; Roters et al., 2011, Jansing et al., 2013) an, die die fachspezifischen Komponenten des professionellen Wissens von Englischstudierenden in der 1. Ausbildungsphase überprüft. Aufbauend auf den Vorarbeiten von TEDS-LT präsentiert der vorliegende Beitrag eine theoriegeleitete Verfeinerung und Neuausrichtung des Konstrukts „Fachdidaktisches Wissen angehender Englischlehrkräfte“, das den unterschiedlichen inhaltlichen Anforderungen, die Englischlehrkräfte beim Unterrichten bewältigen müssen, entspricht und somit zwei grundsätzliche Wissensbereiche – fremdsprachliche Lehr- und Lernprozesse – unterscheidet.

Basierend auf den Vorarbeiten in TEDS-LT sowie den qualitativen Studien von Grossman (1990) und Tsui und Nicholson (1999) wurde ein zweidimensionales Konstrukt fachdidaktischen Wissens entwickelt, das anhand einer sekundären Analyse basierend auf den Daten aus TEDS-LT getestet wurde. Die Testung erfolgte zweistufig durch eine Rasch-Skalierung des Testinstruments und eine explorative Faktorenanalyse. Anhand der Rasch-Skalierung sollte zunächst die Reliabilität des gesamten Testinstruments zur Erfassung des fachdidaktischen Wissens im Fach Englisch überprüft werden. Ziel der explorativen Faktorenanalyse ist die empirische Überprüfung der Zweidimensionalität des Testinstruments zum Konstrukt „Fachdidaktisches Wissen angehender Englischlehrkräfte“.

Zur Neuskalierung des Testinstruments wurden Daten von 417 Studierenden von sechs Universitäten bzw. Hochschulen verwendet. Die Analysen wurden mithilfe der Software Conquest (Wu, Adams & Wilson, 1997) und Mplus (Version 6, Muthén & Muthén, 1998-2010) durchgeführt. Insgesamt erwies sich der Test mit den 38 Testaufgaben bzw. Items als reliabel. Der EAP-Wert für die Reliabilität der Skala betrug 0,67 und die ermittelte empirische Varianz lag bei 0,43. Der kleine Wert der empirischen Varianz gilt als Indiz dafür, dass die Studierenden in den sechs Hochschulen eine relativ homogene Gruppe bilden, die einen ähnlichen Umfang von fachdidaktischem Wissen, wie es hier getestet wird, aufweisen. Mithilfe von Mplus (Version 6, Muthén & Muthén, 1998-2010) wurde eine explorative Faktorenanalyse basierend auf dem Geomin-Rotationsverfahren zur Testung der Zweidimensionalität des Konstrukts durchgeführt. Die zweifaktorielle Lösung zeigt einen akzeptablen Modellfit ( $\chi^2 = 1012,722$ ,  $df = 463$ ,  $p = 0,00$ ; CFI = 1; RMSEA = 0,028; SRMR = 0,075). Mit Ausnahme von zwei Items (13 und 35) laden alle anderen Items eindeutig entweder auf dem ersten oder auf dem zweiten Faktor relativ hoch. Darüber hinaus stimmt diese auf Basis der Faktorladungen erstellte Item-Zuordnung mit den zwei Dimensionen überein, die anhand von theoretischen Vorüberlegungen für diese 38 Items vorgenommen worden war. Angesichts der sehr guten Übereinstimmung von erwarteter und empirischer Zuordnung der Items zu den beiden Dimensionen bzw. Faktoren sprechen die Ergebnisse der explorativen Faktorenanalyse für die Validität des mit dem vorliegenden Testinstrument gemessenen Konstrukts. Dies wurde dann erneut durch eine Raschskalierung überprüft. Diese zeichnete sich dadurch aus, dass die messfehlerfreie Interkorrelation zwischen den beiden Subdimensionen niedrig (0,52) ausfiel. Insgesamt bestätigen unsere Analysen die Zweidimensionalität des Konstrukts „Fachdidaktisches Wissen angehender Englischlehrkräfte“.

Anderson, L.W. (Ed.), Krathwohl, D.R. (Ed.), Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives. New York: Longman.

Jansing, B., Haudeck, H., Keßler, J.-U., Nold, G. & Stancel-Piątak, A. (2013). Professionelles Wissen im Studienverlauf: Lehramt Englisch. In: S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H.

Haudeck, J.-U. Keßler, J.-U. & K. Schwippert (Hrsg.), Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT (77-106). Münster und New York: Waxmann.

Muthén, L.K. & Muthén, B.O. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nold, G. (2013). Fremdsprachendidaktik. In M. Byram & A. Hu (Eds.), Routledge Encyclopedia of Language Teaching and Learning (S. 253-257). London, New York: Routledge.

Roters, B., Nold, G., Haudeck, H. Keßler, J.-U. & Stancel-Piątak, A. (2011). Professionelles Wissen von Studierenden des Lehramts Englisch. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen: Erste Ergebnisse aus TEDS-LT (S. 77-99). Münster: Waxmann.

Trautmann, M. (2010). Professionsforschung in der Fremdsprachendidaktik. In W. Hallet & F. G. Königs (Hrsg.), Handbuch Fremdsprachendidaktik (S. 346-350). Seelze-Velber: Klett Kallmeyer.

Tsui, A.B.M. & Nicholson, S. (1999). A Hypermedia Database and English as a Second Language Teacher Knowledge Enrichment. Journal of Information Technology for Teacher Education, 8(2), 215-237.

Wu, M.L. & Adams, R.J. & Wilson, M. R. (1997). ConQuest: Mutli-Aspect Test Software Computer program]. Camberwell, Vic.: Australian Council for Educational Research.



## **Zusammenhänge zwischen naturwissenschaftsdidaktischem Wissen und professioneller Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht**

Die professionelle Wahrnehmung gilt als wichtige Komponente der Lehrerexpertise (Sherin, 2007). Sie umfasst das selektive Erkennen und Interpretieren lernrelevanter Situationen im Unterricht und ist Voraussetzung für das adaptive Handeln und die Unterrichtsreflexion der Lehrkraft (Bromme, 2008). Empirische Studien zeigen, dass erfahrene Lehrkräfte Unterricht professioneller wahrnehmen als Studierende (Seidel & Prenzel, 2007; van Es & Sherin, 2002; Wolters, Meschede, Steffensky, & Möller, eing.). Bezüglich der Verortung der professionellen Wahrnehmung im Kontext professioneller Kompetenzen gibt es unterschiedliche Ansätze: Von manchen Autoren wird die Fähigkeit des selektiven Erkennens und Interpretierens lernrelevanter Unterrichtssituationen als eigene Komponente der professionellen Kompetenz von Lehrkräften (Sherin & van Es, 2009) verstanden, andere beschreiben die professionelle Wahrnehmung als situierten Bestandteil des Professionswissens (Kersting et al., 2012; Roth et al., 2011). Über den Zusammenhang zwischen Komponenten des Professionswissens, die eher als theoretisch-formales Wissen beschrieben werden können, und der professionellen Wahrnehmung ist bisher wenig bekannt. Lediglich für den Mathematikunterricht konnte gezeigt werden, dass die professionelle Wahrnehmung (hier als „usable knowledge in context“ beschrieben) der Lehrkraft positiv mit ihrem fachbezogenen Professionswissen zusammenhängt (Kersting et al., 2012).

Im Rahmen des BMBF geförderten ViU-Projekts wird (unter anderem) der Frage nach Zusammenhängen zwischen der professionellen Wahrnehmung der Lernunterstützung (diese umfasst Maßnahmen der kognitiven Aktivierung und der inhaltlichen Strukturierung) im naturwissenschaftlichen Grundschulunterricht und dem fachdidaktischen Wissen sowie fachdidaktischen Überzeugungen zum naturwissenschaftlichen Lehren und Lernen nachgegangen. Für die Erfassung der professionellen Wahrnehmung wurde ein videobasiertes Instrument eingesetzt, bei dem sechs Videoszenen anhand von szenenspezifischen Rating-Items eingeschätzt werden müssen (Wolters et al., eingereicht). Das fachdidaktische Professionswissen wurde mit einem auch im ViU-Projekt entwickelten Paper-Pencil-Test und die Überzeugungen zum naturwissenschaftlichen Lehren und Lernen mit einem von Kleickmann (2008) entwickelten Fragebogen erhoben. Die Instrumente wurden an einer Stichprobe aus 434 Probanden eingesetzt, die aus Bachelor- und Masterstudierenden des Grundschullehramtes (Schwerpunkt Sachunterricht) sowie Sachunterrichtslehrkräften bestand. Die Analysen zeigen signifikante Zusammenhänge zwischen den verschiedenen Komponenten der professionellen Kompetenz. So finden sich positive moderate Zusammenhänge zwischen dem fachdidaktischen Wissen und der professionellen Wahrnehmung; in ähnlichen Größenordnungen liegen diese auch bei den einzelnen Gruppen vor (Lehrkräfte, MA- und BA-Studierende). Positive moderate Zusammenhänge finden sich auch zwischen der professionellen Wahrnehmung und den konstruktivistischen Lehr-Lern-Überzeugungen, während zu den transmissiven Lehr-Lern-Überzeugungen negative Zusammenhänge bestehen. Die Ergebnisse zeigen, dass die professionelle Wahrnehmung erwartungsgemäß mit dem mit Paper-Pencil-Tests erfassten fachdidaktischen Wissen zusammenhängt, beide Aspekte der professionellen Kompetenz aber empirisch abgrenzbar sind. Ein zentraler Unterschied ergibt sich sicherlich im Grad der Situietheit der verschiedenen Erhebungsmethoden. Die verschiedenen Lehr-Lern-Überzeugungen wirken möglicherweise als eine Art Filter für die professionelle Wahrnehmung der Lernunterstützung im Unterricht (Borko et al., 2011). Im Vortrag werden die detaillierten Ergebnisse berichtet und diskutiert.

- Borko, H., Koellner, K., Jacobs, J. & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM Mathematics Education*, 43, 175–187.
- Bromme, R. & Haag, L. (2008). Forschung zur Lehrerpersönlichkeit. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 803-819). Wiesbaden: VS Verlag Sozialwissenschaften.
- Kersting, N.B., Givvin, K.B., Thompson, B.J., Santagata, R. & Stigler, J.W. (2012). Measuring Usable Knowledge: Teachers' Analyses of Mathematics Classroom Videos Predict Teaching Quality and Student Learning. *American Educational Research Journal*, 49, 568-589.
- Kleickmann, T. (2008). Zusammenhänge fachspezifischer Vorstellungen von Grundschullehrkräften zum Lehren und Lernen mit Fortschritten von Schülerinnen und Schülern im konzeptuellen Verständnis. Münster: Inaugural-Dissertation.
- Roth, K. J., Garnier, H., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, 48, 117-148
- Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen bei Lehrpersonen mit Hilfe von Videosequenzen. *Zeitschrift für Erziehungswissenschaft, Sonderheft*, 8, 201-216.
- Sherin, M.G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (S. 383-395). Hillsdale, NJ: Erlbaum.
- Sherin, M. G. & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60 (1), 20–37.
- van Es, E.A. & Sherin, M.G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10, 571-596.
- Wolters, M., Meschede, N., Steffensky, M. & Möller, K. (eing.). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht - Theoretische Beschreibung und empirische Erfassung.

## ***Pädagogisches Professionswissen und Instruktionsqualität im ersten Schulpraktikum***

Es wird übereinstimmend davon ausgegangen, dass professionelles Wissen die Voraussetzung für professionelles Handeln im Lehrerberuf darstellt. Eine hohe Wissensbasis ermöglicht kompetentes, flexibles und adaptives Handeln und hat demnach eine entscheidende Bedeutung für den Lernerfolg von Schülerinnen und Schülern (Lipowsky, 2006). In Anlehnung an aktuelle Kompetenzmodelle (vgl. z.B. Baumert & Kunter, 2011) wird das Professionswissen in die Facetten Fachwissen, fachdidaktisches und allgemeines pädagogisches Wissen unterschieden, denen jeweils eine unterschiedlich hohe Bedeutung für den Lernerfolg von Schülern zugewiesen wird (Baumert & Kunter, 2006; Lipowsky, 2006).

Für die theoretische Annahme einer effektiveren Instruktionsqualität durch eine höhere Wissensbasis liegt jedoch erst ansatzweise empirische Evidenz vor. Kunter et al. (2013) konnten wie auch Baumert et al. (2010) zeigen, dass höheres fachdidaktisches Wissen mit einer höheren kognitiven Aktivierung und einer stärkeren Unterstützung der Lernenden einhergeht; auf das Klassenraummanagement hatte das fachdidaktische Wissen jedoch keinen Einfluss. Hinsichtlich des Fachwissens in Mathematik konnte bisher kein Zusammenhang mit der Instruktionsqualität gezeigt werden (Baumert et al., 2010).

Inwiefern das fachunabhängige pädagogische Wissen mit der Instruktionsqualität assoziiert ist, wurde bisher empirisch noch nicht geprüft und wird im vorliegenden Beitrag untersucht.

Den Rahmen der Untersuchung liefert die längsschnittlich angelegte Studie SioS-L (Studie zu individuellen und organisationalen Einflüssen auf den Studien- und Ausbildungserfolg in der Lehrerbildung; Brünken, Spinath, & Bedersdorfer, 2009). In SioS-L werden die Studierenden über einen Zeitraum von insgesamt 6 Jahren in einem längsschnittlichen Kohortendesign über ihr Studium und ihre zweite Ausbildungsphase hinweg begleitet. Eine Teilstichprobe (N=89) innerhalb des Projekts wird bei ihren Unterrichtsversuchen in den Schulpraktika videographiert; derzeit liegen hierzu Daten aus dem ersten sogenannten Orientierungspraktikum vor. Dieses findet für alle Studierende nach dem 1. bzw. 2. Semester statt. Die Instruktionsqualität wurde von je zwei unabhängigen Ratern auf den Dimensionen Klarheit/Strukturiertheit, Umgang mit Störungen/Störungsprävention und Schülerorientierung/Lernprozessbegleitung eingeschätzt (ICC zwischen .61 und .78). Das pädagogische Wissen wurde zeitlich ca. 3-6 Monate vor dem Praktikum mit zwei Tests erfasst: (1) wurde der Test zum Pädagogischen Unterrichtswissen (PUW, König & Blömeke, 2010) eingesetzt, (2) ein eigener Wissenstest entwickelt, der sich an den Inhalten des Bildungswissenschaftlichen Studiums der Universität des Saarlandes orientiert, die sich eng an die KMK-Standards anlehnen (Kultusministerkonferenz, 2004).

Betrachtet man die (um die Abiturnote bereinigten) Korrelationen des Instruktionsverhaltens mit der Gesamtleistung der beiden Tests, so zeigen sich keine substantiellen Zusammenhänge. Tendenzielle Zusammenhänge zeigen sich jedoch für einzelne Wissensbereiche: Der Bereich Strukturierung des PUW geht positiv mit der im Unterricht gezeigten Klarheit und Strukturiertheit einher ( $r=.18$ ,  $p=.08$ ), ein positiver Zusammenhang zeigt sich ebenfalls zwischen dem Wissen über Heterogenität und der Schülerorientierung/Lernprozessbegleitung im Unterricht ( $r=.18$ ;  $p=.09$ ). Für den eigenen Wissenstest sind ebenfalls tendenzielle Zusammenhänge zwischen Wissen im Bereich Lehren und Lernen (vgl. Kompetenzbereich 2 der KMK-Standards) mit der Skala Umgang mit Störungen/Störungsprävention sichtbar ( $r=.20$ ,  $p=.06$ ).

Die gefundenen Zusammenhänge können als ein erstes Indiz dafür betrachtet werden, dass die Wissensbasis zu Beginn des Studiums mit Qualitätsmerkmalen guter Unterrichtsgestaltung im Schulpraktikum in Zusammenhang steht. Insbesondere die Zusammenhänge unterschiedlicher Wissensbereiche mit korrespondierenden Dimensionen des Instruktionsverhaltens stützen diese Annahme. Die Ergebnisse werden – vor dem Hintergrund, dass die Daten in einer sehr frühen Studienphase erhoben wurden – im Hinblick auf deren Aussagekraft und mögliche Implikationen für die weitere Forschung diskutiert.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469-520.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, S. Klusmann & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV (S. 29-53). Münster: Waxmann.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. American Educational Research Journal, 47(1), 133-180.

Brünken, R.; Spinath, F. & Bedersdorfer, H.-W. (2009). Studie zu individuellen und organisationalen Einflüssen auf den Studienerfolg in der Lehrerbildung. Projektantrag im Rahmen des BMBF-Förderschwerpunkts „Professionalisierung des Pädagogischen Personals.“ Saarbrücken: Universität des Saarlandes.

König, J. & Blömeke, S. (2010). Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerbildung. Berlin: Humboldt-Universität.

Kultusministerkonferenz. (2004). Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004. Retrieved 05.08.2013, from [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf)

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional Competence of Teachers: Effects on Instructional Quality and Student Development. Journal of Educational Psychology.

Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. Zeitschrift für Pädagogik, 52 (51. Beiheft), 47-71.





Chair(s): Johanna Schmid (Eberhard Karls Universität Tübingen, IDeA-Zentrum), Jan Kühnhausen (DIPF, IDeA-Zentrum)

Diskutant/in: Ulrich Ebner-Priemer (Karlsruher Institut für Technologie)

Dieses interdisziplinäre Symposium befasst sich mit der Frage nach dem Potential körperlicher Bewegung für die Erkennung und Förderung behavioraler, affektiver und kognitiver Korrelate bei Kindern, Jugendlichen und Erwachsenen. Es werden dabei empirische Arbeiten präsentiert, die dieser Frage aus unterschiedlichen theoretischen Perspektiven (z.B. klinische Psychologie, kognitive Psychologie, Mathematikdidaktik) und mit vielfältigen methodischen Forschungsdesigns (z.B. querschnittlich, experimentell, längsschnittlich) nachgehen.

Die ersten beiden Beiträge fokussieren Kinder und Jugendliche mit Symptomen der Aufmerksamkeitsdefizit-/ Hyperaktivitätsstörung (ADHS):

Es wird berichtet, inwiefern objektiv mittels Accelerometern erfasste Bewegung zur ökonomischen Erkennung von Hyperaktivität genutzt werden kann. Hierzu wurde eine Studie durchgeführt, bei der die Bewegungen von Kindern mit und ohne diagnostizierter ADHS während der Bearbeitung einer Computeraufgabe verglichen werden. Anhand des Analyseansatzes der Support Vector Machines wird versucht, die vorliegenden und auf Fremdbberichten beruhenden Diagnosen durch die Bewegungsdaten zu replizieren.

Im zweiten Beitrag wird eine intensive Längsschnittstudie vorgestellt, welche den Zusammenhang zwischen Bewegungsverhalten im Alltag und abendlicher Stimmung von Jugendlichen untersucht. Die Ergebnisse dieser Studie zeigen, dass körperliche Bewegung im Alltag von Jugendlichen und insbesondere bei Jugendlichen mit erhöhten Hyperaktivitätssymptomen positive Effekte auf die unmittelbare Reduktion tagtäglicher negativer Stimmung haben kann.

Die weiteren beiden Beiträge folgen experimentellen Untersuchungsdesigns und beschäftigen sich mit dem Nutzen körperlicher Bewegung zur Förderung kognitiver Kompetenzen in verschiedenen Lebensphasen:

Der dritte Beitrag präsentiert Ergebnisse zweier experimenteller Studien mit Erst- und Zweitklässlern. Diese Studien bauen auf Theorien der embodied cognition (z.B. Barsalou, 2008) auf und zeigen übereinstimmend, dass basisnumerische Repräsentationen, welche entscheidend für spätere numerische Leistungen sind, effektiv durch räumlich-körperliche Trainings verbessert werden können. Eine mögliche didaktische Implementierung dieser Trainingsform in den Mathematik-Förderunterricht wird aufgezeigt.

Der vierte Beitrag präsentiert Ergebnisse eines Trainings zur Steigerung der Fitness bei älteren Erwachsenen und dessen Auswirkungen auf die kognitive Leistung. Es wird gezeigt, dass regelmäßige Bewegung eine positive Wirkung sowohl auf die körperliche Fitness, als auch die kognitive Leistung über die Lebensspanne hinweg hat.

Zusammenfassend zeigt das Symposium einerseits die Bedeutung der Analyse körperlicher Bewegung in Bezug auf die Erkennung / Diagnostik behavioraler Merkmale und andererseits die vielfältigen Einsatzgebiete von körperlicher Bewegung, die in der Steigerung des Wohlbefindens und der kognitiven Leistungen liegen. Dies hat bedeutende praktische Implikationen für die klinische Psychologie aber auch für die Mathematikdidaktik und für die Konzeption von Trainingsmaßnahmen im Entwicklungsverlauf.



## **Erkennung von Hyperaktivität auf Basis von Bewegungsdaten**

Die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) ist eine psychische Störung, die im Alltag von Betroffenen zu erheblichen Einschränkungen führen kann. Da sie vor allem im jungen Alter auftritt, wirken sich diese Einschränkungen häufig besonders negativ auf die akademische Leistung der Betroffenen aus. Zusätzlich führt die hohe Prävalenz von ADHS (fast 5% der Kinder und Jugendlichen in Deutschland) zu hohen Belastungen des Gesundheitssystems (Huss et al., 2008). Eine effektive Behandlung von ADHS ist somit von hoher Bedeutung, sowohl für die betroffenen Kinder als auch für die Gesellschaft. Daher sind effektive Methoden zur Erkennung von ADHS wichtig. Aufgrund der hohen Anzahl von Betroffenen, sollten diese Methoden möglichst zeitsparend und einfach anwendbar sein. Laut DSM (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association, 2000) gibt es drei Kernsymptome von ADHS, auf die sich alle Diagnosen beziehen: Unaufmerksamkeit, Impulsivität und Hyperaktivität. Der hier vorliegende Beitrag fokussiert auf das Kernsymptom der Hyperaktivität. Hyperaktivität drückt sich unter anderem in übermäßigem Zappeln und der Unfähigkeit, ruhig zu sitzen, aus (Gawrilow, 2012). Diese Symptome sind für Außenstehende sichtbar und können daher verhältnismäßig objektiv identifiziert werden. Da sich Hyperaktivität also in vermehrter bzw. veränderter Bewegung ausdrückt, wird hier ein Versuch unternommen, diese mit Hilfe von Bewegungsmessern zu erkennen. Im Rahmen einer Studie trugen 97 Kinder (8-12 Jahre) einen Bewegungsmesser, während sie am Computer Aufgaben bearbeiteten. Für einen Teil dieser Kinder liegt eine diagnostizierte Hyperaktivität vor, während für einen anderen Teil bekannt ist, dass sie nicht unter Hyperaktivität leiden. In diesem Beitrag wird versucht diese Diagnosen auf Basis der während der Aufgabenbearbeitung gesammelten Bewegungsdaten zu replizieren. Hierfür werden die Daten mit Hilfe von Support-Vector-Machines (Cortes & Vapnik, 1995) analysiert. Support-Vector-Machines sind eine effektive und flexible Methode für Klassifizierungsprobleme und konnten von uns für eine ähnliche Problemstellung bereits erfolgreich angewendet werden (Kühnhausen et al., 2013). Sie erlauben eine präzise, auf Daten basierende, Einteilung von Individuen in bestimmte Gruppen. In der vorliegenden Arbeit werden Support-Vector-Machines benutzt, um unterschiedliche Muster im Bewegungsverhalten von hyperaktiven und nicht hyperaktiven Kindern zu finden und diese Unterschiede zur Erkennung von Hyperaktivität zu nutzen. Erste Auswertungen von während der Testungen aufgenommenen Videos zeigen bereits, dass sich hyperaktive Kinder erkennbar anders verhalten und bewegt haben, als nicht hyperaktive Kinder. Mit einer solchen Methode könnten Symptome von Hyperaktivität in großen Gruppen, wie zum Beispiel ganzen Schulklassen, relativ zeit- und kosteneffizient erkannt werden. Des Weiteren könnte hierdurch ein Risiko für Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung frühzeitig erkannt werden.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC: Author.

Cortes, C., and Vapnik, V. N. (1995). Support-vector networks. *Machine Learning*, 20, 273-297. doi:10.1007/BF00994018

Gawrilow, C. (2012). Lehrbuch ADHS. München: Reinhardt UTB.

Huss, M., Hölling, H., Kurth, B. M., & Schlack, R. (2008). How often are German children and adolescents diagnosed with ADHD? Prevalence based on the judgment of health care professionals: results of the German health and examination survey (KiGGS). *European Child & Adolescent Psychiatry*, 17(1), 52-58.

Kühnhausen, J., Leonhardt, A., Dirk, J., & Schmiedek, F. (2013). Physical activity and affect in elementary school children's daily lives. *Frontiers in Psychology*, 4, 456. doi:10.3389/fpsyg.2013.00456



## **Körperliche Bewegung und Stimmung bei Jugendlichen mit Symptomen der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS)**

Theoretischer Hintergrund: Negative Stimmung und depressive Störungen sind häufige komorbide Auffälligkeiten von Kindern und Jugendlichen mit ADHS (z.B. Blackman, Ostrander, & Herman, 2005; Daviss, 2008; DuPaul, Barkley, & McMurray, 1994). Zahlreiche Studien weisen auf Zusammenhänge zwischen körperlicher Bewegung und Stimmung hin: Je aktiver Menschen sind, umso besser ist ihre Stimmung. Ebenso sind Interventionen zur Steigerung der körperlichen Bewegung wirksam zur Reduktion depressiver Symptome im Erwachsenenalter (Conn, 2010). Bislang gibt es jedoch noch wenige Studien, die den Zusammenhang zwischen körperlicher Bewegung und Stimmung bei Kindern und Jugendlichen mit ADHS untersuchen. Diese Studien deuten auf die Wirksamkeit von Interventionen zur Verbesserung der Stimmung bei Kindern und Jugendlichen mit ADHS hin (Jensen und Kenny, 2004; Kiluk, Weden, & Culotta, 2009; Verret et al., 2012). Der Zusammenhang zwischen körperlicher Bewegung und Stimmungsschwankungen von Kindern und Jugendlichen mit Symptomen der ADHS im Alltag wurde allerdings bisher noch nicht untersucht. Die vorliegende Studie greift diese Forschungslücke auf und untersucht den Zusammenhang zwischen täglicher körperlicher Bewegung und Stimmung innerhalb und zwischen Jugendlichen.

Methode: Design. Die Studie folgt einem intensiven längsschnittlichen Untersuchungsdesign. Die teilnehmenden Jugendlichen trugen an 10 aufeinander folgenden Tagen Schrittzähler und beurteilten jeweils am Abend ihre aktuelle Stimmung.

Teilnehmende und Hintergrundvariablen. Die Stichprobe bestand aus 38 Jugendlichen (52.6% Mädchen, M = 14.4 Jahre). Alle Teilnehmenden besuchten eine Gesamtschule und gehörten mehrheitlich der Mittelschicht an. Ein Sechstel der Teilnehmenden (15.6%) zeigte eine erhöhte Depressivität (Allgemeine Depressionsskala Kurzversion, Mayer & Hautzinger, 2001; Cut-Off Werte basierend auf Lehr, Hillert, Schmitz, & Sosnowsky, 2008).

Messinstrumente: Depressive Stimmung. Die Teilnehmenden berichteten ihren aktuelle Stimmung täglich abends auf einer 5-stufigen Skala von 1 (gar nicht) bis 5 (sehr) anhand von drei Items: „traurig“, „deprimiert“ und „mutlos“. Die drei Items wurden gemittelt und zu einem Kennwert zusammengefasst.

Körperliche Bewegung. Tägliche körperliche Bewegung wurde durch Schritte erfasst und mit piezoelektrischen Schrittzählern (OMRON Walking Style II - HJ-113; OMRON Healthcare, Mannheim, Germany) gemessen.

Hyperaktivität. Die Teilnehmenden schätzen ihre Hyperaktivität anhand eines Items („Ich bin ständig in Bewegung und zappelig“) des Strength and Difficulties Questionnaire (SDQ; Goodman, 1997) auf einer 3-stufigen Skala von 0 (stimmt nicht) bis 2 (stimmt genau) ein.

Ergebnisse: Mehrebenenanalysen zeigten einen moderaten Zusammenhang zwischen körperlicher Bewegung und Stimmung innerhalb von Personen: Die Teilnehmenden berichteten geringere depressive Stimmung an Tagen mit individuell mehr körperlicher Bewegung als gewöhnlich, Estimate (SE) = -0.29 (0.11),  $p < .05$ . Dieser Zusammenhang zwischen körperlicher Bewegung und Stimmung innerhalb von Personen war stärker für Jugendliche mit höherer Hyperaktivität, Estimate (SE) = -0.38 (0.17),  $p < .05$ . Zwischen Personen zeigte sich kein signifikanter Zusammenhang für körperliche Bewegung und Stimmung: Jugendliche, die durchschnittlich mehr körperlicher Bewegung zeigten, berichteten nicht weniger negative Stimmung als Jugendliche mit weniger körperlicher Bewegung (Estimate (SE) = 0.03 (0.26), ns) und Hyperaktivität moderierte diesen Zusammenhang nicht (Estimate (SE) = -0.58 (0.39), ns).

Diskussion: Die Ergebnisse der Studie deuten darauf hin, dass körperliche Bewegung im Alltag von Jugendlichen, die nicht im Kontext strukturierter Interventionsprogramme stattfindet, positive Effekte auf die Reduktion tagtäglicher negativer Stimmung haben kann – insbesondere bei Jugendlichen mit erhöhter ADHS-Symptomatik. Weitere Forschung ist nötig um der Frage nachzugehen, ob allgemein erhöhte körperliche Bewegung bzw. körperliche Fitness unter spezifischen Voraussetzungen in positivem Zusammenhang mit verbesserter Stimmung steht. Zudem sollen aktuelle Ergebnisse weiterführend diskutiert werden.

Blackman, G. L., Ostrander, R. & Herman, K. C. (2005). Children with ADHD and depression: A multisource, multimethod assessment of clinical, social, and academic functioning. *Journal of Attention Disorders*, 8, 195–207. doi:10.1177/1087054705278777

Conn, V. S. (2010). Depressive Symptom outcomes of physical activity interventions: Meta-analysis findings. *Annals of Behavioral Medicine*, 39, 128–138. doi:10.1007/s12160-010-9172-x

Daviss, W. (2008). A review of co-morbid depression in pediatric ADHD: Etiologies, phenomenology, and treatment. *Journal of Child and Adolescent Psychopharmacology*, 18, 565–571. doi:10.1089/cap.2008.032

DuPaul, G. J., Barkley, R. A., & McMurray, M. B. (1994). Response of children with ADHD to Methylphenidate: Interaction with internalizing symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 894–903. doi: 10.1097/00004583-199407000-00016

Jensen, P. S., & Kenny, D. T. (2004). The effects of yoga on the attention and behavior of boys with attention-deficit/hyperactivity disorder (ADHD). *Journal of Attention Disorders*, 7, 205–216. doi:10.1177/108705470400700403

Kiluk, B. D., Weden S., & Culotta, V. P. (2009). Sport participation and anxiety in children with ADHD. *Journal of Attention Disorders*, 12, 499–506. doi:10.1177/1087054708320400

Verret, C., Guay, M.-C., Berthiaume, C., Gardiner, P., & Béliveau, L. (2012). A physical activity program improves behavior and cognitive functions in children with ADHD: An exploratory study. *Journal of Attention Disorders*, 16, 71–80. doi:10.1177/1087054710379735



Tanja Link (Eberhard Karls Universität Tübingen), Korbinian Möller (Eberhard Karls Universität Tübingen),  
Hans-Christoph Nuerk (Eberhard Karls Universität Tübingen)

## ***Embodied Trainings – räumlich-körperliche Bewegung zur Verbesserung basisnumerischer Repräsentationen***

Zahlen sind ein fester Bestandteil des menschlichen Lebens und aus dem Alltag kaum wegzudenken (Moeller, Pixner, Klein, Cress, & Nuerk, 2009). Die Tatsache, dass ca. 5% der deutschen Zweit- bis Drittklässler starke Probleme im Rechnen haben (Fischbach et al., 2013), verdeutlicht, dass frühe Interventionen wichtig und notwendig sind. In der kognitiven Entwicklungspsychologie ebenso wie in der Fachdidaktik der Mathematik wird die Bedeutsamkeit basaler numerischer Repräsentationen wiederholt hervorgehoben. Einige Studien betonen hierbei den Zusammenhang zwischen in der frühen Kindheit entwickelten zahlsspezifischen Kompetenzen und späterer curricularer Mathematikleistung (z.B. Jordan, Kaplan, Ramineni, & Locuniak 2009; Moeller, Pixner, Kaufmann, & Nuerk, 2011). Die Zusammenhänge verdeutlichen, dass die Grundlagen basisnumerischer Fähigkeiten bereits vor dem Mathematikerstunterricht der Grundschule erworben werden und entscheidend für spätere numerische Leistungen sind.

Das vorliegende Projekt konzentriert sich auf die Förderung von zwei Basiskomponenten numerischer Kognition: die räumliche Größenrepräsentation bzw. den sog. mentalen Zahlenstrahl sowie die Repräsentation der Platz x Wert-Struktur des Arabischen Zahlensystems (Nuerk, Graf, & Willmes, 2006). Diese wurden basierend auf einem neuen Interventionsansatz (siehe Fischer, Moeller, Bientzle, Cress, & Nuerk, 2011 für eine theoretische Ableitung des Konzepts) aufbauend auf Theorien der sog. embodied cognition (z.B. Barsalou, 2008) in zwei Studien systematisch trainiert.

In einer ersten Studie (Link, Moeller, Huber, Fischer, & Nuerk, 2013) wurden Erstklässler darin trainiert die Position von Zahlen auf einem leeren Zahlenstrahl, von welchem nur Start- und Endpunkt (0 und 100) markiert waren anzugeben. In der räumlich-körperlichen Experimentalbedingung fand dieses Training auf einem auf den Boden geklebten bis zu 3 Meter langen Zahlenstrahl statt, bei dem die Kinder durch die Bewegung entlang des Zahlenstrahls die Größe der Zahlen durch die zurückgelegte Distanz (vom Startpunkt bis zur geschätzten Position der anzugebenden Zahl) räumlich und körperlich erfahren konnten. Verglichen wurden die Trainingseffekte mit einem Kontrolltraining in dem derselbe numerische Inhalt an einem Tablet-PC jedoch ohne aufgabenspezifische Bewegung trainiert wurde. Nach beiden Trainings konnten spezifische Trainingseffekte auf die Verbesserung der räumlichen Zahlenstrahlrepräsentation und darüber hinaus Transfereffekte auf Additionsaufgaben beobachtet werden. Spezifische und differentielle Trainingseffekte waren den Erwartungen entsprechend nach dem räumlich-körperlichen Training größer als nach dem Training am Tablet-PC.

In einer weiteren Studie (Link, Schwarz, Huber, Fischer, Cress, Nuerk, & Moeller, in Vorbereitung) wurden Zweitklässler ebenfalls mit einer Zahlenstrahlaufgabe trainiert, wobei hier, analog zur Repräsentation von Zehnern und Einern in der Platz x Wert-Struktur arabischer Zahlen, eine getrennte Eingabe dieser erfolgte. Die Kinder sollten einen Balken durch Betreten verschiedener Felder einer Tanzmatte in Zehner- bzw. Einerschritten bis zur vorgegebenen Zielposition anwachsen lassen. Die vorderen Felder der Tanzmatte wurden hierfür auf eine Treppenstufe gelegt, sodass das Erreichen dieser Zehnerfelder mit größerem körperlichen Aufwand verbunden war als das Erreichen der ebenerdig platzierten Einerfelder. Zwei Kontrollbedingungen sollten sicherstellen, dass die Trainingseffekte der Experimentalbedingung nicht ausschließlich auf den numerischen Inhalt der Trainingsaufgabe bzw. auf die enthaltene Bewegung und das Medium Tanzmatte zurückgeführt werden konnten. Die Ergebnisse entsprachen den Erwartungen: im Vergleich zu den Kontrolltrainings zeigte sich die deutlichste Verbesserung der räumlichen Zahlenstrahlrepräsentation nach dem Experimentaltraining. Auch Transfereffekte auf nicht direkt trainierte Aufgaben (z.B. Addition, nonsymbolischer Größenvergleich) waren nach dem Experimentaltraining am stärksten ausgeprägt. Da die Kontrolltrainings sich in ihrer Wirkung nicht signifikant voneinander unterschieden, kann davon ausgegangen werden, dass es die Kombination aus numerischem Aufgabeninhalt und aufgabenspezifischer Bewegung war, die zu deutlichen Lerneffekten führte.

Die beiden vorgestellten Studien liefern damit weitere Evidenz für die Wirksamkeit räumlich-körperlicher Trainings zur Verbesserung numerischer Kompetenzen. Geringe Anschaffungskosten und bedeutsame Lerneffekte, die bereits nach nur 3 Trainingseinheiten zu beobachten waren, sprechen dabei für eine Implementierung dieser Trainingsform in den Mathematik-Förderunterricht.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645. doi: 10.1146/annurev.psych.59.103006.093639

Fischbach, A., Schuchardt, K., Brandenburg, J., Kleszczewski, J., Balke-Melcher, C., Claudia Schmidt, C., Büttner G., Grube, D., Mähler, C., & Hasselhorn M. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen*, 2, 65-76. doi: 10.1024/2235-0977/a000035

Fischer, U., Moeller, K., Bientzle, M., Cress, U., & Nuerk, H.-C. (2011). Embodied spatial numerical training of number magnitude representation – An intervention study. *Psychonomic Bulletin & Reviews*, 18, 177-183. doi: 10.3758/s13423-010-0031-3

Jordan, N.C., Kaplan, D., Ramineni, C., & Locuniak, M.N. (2009). Early math matters: Kindergarten number competence and later arithmetic outcomes. *Developmental Psychology*, 45, 850-867. doi: 10.1037/a0014939

Link, T., Schwarz, E.J., Huber, S., Fischer, U., Cress, U., Nuerk, H.-C., & Moeller, K. (in Vorbereitung). *Mathe mit der Matte – Verkörperlichtes Training der Zehner-Einer Struktur des mentalen Zahlenstrahls.*

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H.-C. (2013). Walk the number line – An embodied training of numerical concepts. *Trends in Neuroscience and Education*, 2, 74-84. doi: 10.1016/j.tine.2013.06.005

Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2009). Children's early mental number line: Logarithmic or rather decomposed linear? *Journal of Experimental Child Psychology*, 103, 503-515. doi: 10.1016/j.jecp.2009.02.006

Moeller, K., Pixner, S., Klein, E., Cress, U., & Nuerk, H.-C. (2009). Zahlenverarbeitung ist nicht gleich Rechnen – Eine Beschreibung basisnumerischer Repräsentationen und spezifischer Interventionsansätze. *Prävention und Rehabilitation*, 21, 121-136.

Nuerk, H.-C., Graf, M., & Willmes, K. (2006). Grundlagen der Zahlenverarbeitung und des Rechnens. In *Sprache, Stimme, Gehör: Zeitschrift für Kommunikationsstörungen, Schwerpunktthema Dyskalkulie* (Guest Editor: Landerl, K.), 30, 147-153. Doi: 10.1055/s-2006-951751

GEBF TAGUNG  
2014

B06 c

Symposiums  
beitrag

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G165



Erkennung und Förderung von Verhalten durch körperliche Bewegung

Maike Kleemeyer (MPI für Bildungsforschung), Sabine Schaefer (MPI für Bildungsforschung),  
Ulman Lindenberger (MPI für Bildungsforschung)

## ***Wirkt sich ein sportliches Training positiv auf die kognitiven Leistungen älterer Erwachsener aus? Erste Ergebnisse einer Fitnessintervention***

Es konnte bereits mehrfach gezeigt werden, dass sich eine Verbesserung der körperlichen Fitness, insbesondere im Bereich der Ausdauer, bei alten Erwachsenen positiv auf kognitive Leistungen und neurologische Prozesse auswirkt (Colcombe & Kramer, 2003; Smith et al., 2010). Insbesondere Aufgaben, die exekutive Kontrolle erfordern, scheinen von einer verbesserten körperlichen Fitness zu profitieren. Die zugrunde liegenden neuronalen Mechanismen sind jedoch noch weitgehend unklar. Die vorliegende Studie untersucht, ob trainingsinduzierte Fitnessverbesserungen mit Veränderungen der kognitiven Leistungsfähigkeit zusammenhängen.

Untrainierte, aber gesunde ältere Erwachsene (60 bis 75 Jahre alt) wurden zufällig auf zwei Interventionsgruppen aufgeteilt. Beide Gruppen trainierten über einen Zeitraum von 6 Monaten drei Mal die Woche für jeweils eine knappe Stunde auf Fahrradergometern. Die Gruppen unterschieden sich in der umgesetzten Trainingsintensität: Die weniger intensiv trainierende Gruppe (n = 27) fuhr mit einem konstant niedrigen Widerstand von nur 10 Watt, während die intensiv trainierende Gruppe (n = 25) in Abhängigkeit der Herzrate mit einer individuell adjustierten Belastung fuhr, die meist deutlich höher lag. Vor dem Beginn des Trainings (Prätest), unmittelbar nach der 6monatigen Trainingsphase (Posttest 1) und 6 Monate nach Ende des Trainings (Posttest 2) durchliefen alle Probanden eine Batterie kognitiver Tests (mit Maßen für die kognitive Geschwindigkeit, exekutive Funktionen, das episodische Gedächtnis, Vokabelwissen und Logik), eine sportmedizinische Untersuchung mit Fitnessmessung, mehrere magnetresonanztomographische Messungen und diverse motorische Tests. Beide Gruppen konnten ihre Ausdauerleistung durch das Training verbessern, wobei die Leistungsverbesserungen bei den intensiv trainierenden Probanden für einige Leistungsbereiche deutlicher ausfielen als bei den weniger intensiv trainierenden Probanden. Auch in diversen kognitiven Aufgabenbereichen kam es in beiden Gruppen zu Leistungsverbesserungen. Interaktionen mit dem Gruppenfaktor fanden sich für Aufgaben, die im Bereich der exekutiven Funktionen anzusiedeln sind und die erfordern, zwischen zwei Antwortdimensionen hin und her zu switchen oder den Inhalt des Arbeitsgedächtnisses fortlaufend zu aktualisieren. In diesen Bereichen verbesserte die intensiv trainierende Gruppe ihre Leistungen deutlicher als die weniger intensiv trainierende Gruppe. Diese Befunde stimmen mit den Ergebnissen der Metaanalyse von Colcombe und Kramer (2003) überein. Im Ausblick soll auch auf neuronale Mechanismen eingegangen werden, die diesen Veränderungen zugrunde liegen könnten.

Colcombe, S. & Kramer, A.F.(2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, 14 (2), 125–130.

Smith, P. J., Blumenthal, J. A., Hoffman, B. M., Cooper, H., Strauman, T. A., Welsh-Bohmer, K., ... Sherwood, A. (2010). Aerobic exercise and neurocognitive performance: a meta-analytic review of randomized controlled trials. *Psychosomatic Medicine*, 72(3), 239–252. doi:10.1097/PSY.0b013e3181d14633

## **Entwicklung und Validierung einer Skala zur Erfassung computerbezogener Motivationen bei Erwachsenen**

Chair(s): Martin Senkbeil (IPN Kiel)

Theoretischer Hintergrund und Fragestellung:

Bislang fehlt es an theoretisch fundierten Modellen und Skalen, die individuelle Unterschiede in computerbezogenen Kompetenzen (ICT Literacy) aufklären können. Da die Bedeutung motivationaler Variablen für akademisches wie außerakademisches Leistungsverhalten vielfach belegt ist (Eccles & Wigfield, 2002), war das Ziel unserer Untersuchung, eine ökonomische Skala zur Erfassung computerbezogener Motivationen zu entwickeln und zu validieren. Hierbei wurden für die Konzeptualisierung intrinsischer (enjoyment value) und extrinsischer Motivationsaspekte (utility value) in Anlehnung an den aus der Medienforschung stammenden „Uses & Gratifications“ (U&G)-Ansatz verschiedene Anreizfaktoren berücksichtigt (Papacharissi & Rubin, 2000).

Fragestellung:

Grundlegende Annahme war, dass sich die in U&G-Studien identifizierten Anreizfaktoren (z.B. Computer als Lern- und Arbeitswerkzeug, Informationssuche, sozialer Austausch, Unterhaltung, Eskapismus) jeweils einer intrinsischen bzw. extrinsischen Motivationskomponente zuordnen lassen (vgl. Luo, Chea & Chen, 2011). Dabei wurde erwartet, dass extrinsische Motivationsaspekte (z.B. Computer als Lern- und Arbeitswerkzeug oder zur Informationssuche) einen signifikanten Beitrag zur Aufklärung computerbezogener Kompetenzen (ICT Literacy) leisten, während intrinsische Motivationsaspekte (Computer zur Unterhaltung oder Entspannung) aufgrund ihrer hedonistischen Orientierung von ICT Literacy weitgehend unabhängig sein sollten. Weiterhin wurden spezifische Validitätsaspekte der Skala unter Berücksichtigung theoretisch einschlägiger Referenzkonstrukte (z.B. ICT Literacy, Merkmale der Computernutzung, Indikatoren der sozialen Herkunft und Persönlichkeitsmerkmale) untersucht.

Methode:

Die Messeigenschaften der neu konstruierten Skala sowie die erwähnten Validitätsaspekte wurden in zwei Studien, die im Rahmen des Nationalen Bildungspanels durchgeführt wurden, mit Erwachsenen (N = 215 und N = 462) überprüft. Alle Zusammenhänge wurden mit Mplus latent modelliert.

Ergebnisse:

Konfirmatorische Faktorenanalysen wiesen in beiden Studien auf ein reliables und faktoriell valides Instrument zur Erfassung computerbezogener Motivationen hin. Weiterhin zeigten sich im Sinne der Konstruktvalidität hypothesenkonforme und gut interpretierbare Zusammenhänge mit den Validierungsvariablen. Erwartungsgemäß korrelierte die extrinsische Motivation (utility value) positiv mit ICT Literacy sowie einem hohen Bildungsabschluss und den Persönlichkeitsmerkmalen Offenheit für Erfahrungen und Need for Cognition, während sich für die intrinsische Motivation (enjoyment value) keine oder sogar leicht negative Zusammenhänge mit den Validierungsmerkmalen ergaben. Vertiefende Analysen deuteten zusätzlich darauf hin, dass der starke Zusammenhang zwischen sozialer Herkunft und ICT Literacy ( $r = .53$ ) zumindest partiell über extrinsische und intrinsische Motivationsaspekte vermittelt wird. Insgesamt weisen die Ergebnisse auf ein valides wie ökonomisches Instrument zur Erfassung computerbezogener Motivationen bei Erwachsenen hin. Abschließend werden im Beitrag Limitationen dieser Studie sowie Implikationen für zukünftige Untersuchungen diskutiert.

Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132.

Luo, M. M., Chea, S. & Chen, J.-S. (2011). Web-based information service adoption: A comparison of the motivational model and the uses and gratifications theory. *Decision Support Systems*, 51, 21--30

Papacharissi, Z. & Rubin, A. M. (2000). Predictors of internet use. *Journal of Broadcasting & Electronic Media*, 44, 175--196.

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G131



***Ich kann nicht – ich will nicht:***

***Kognitive und motivational-/ affektive Faktoren im Umgang mit Computern bei Jungen und Mädchen***

Ausgangspunkt dieser Studie ist die in der Literatur beschriebene Feststellung, dass sich Jungen und Mädchen in ihrem Nutzungsverhalten und in ihrem Wissen über Computer unterscheiden (Vgl. Vekiri & Chronaki, 2008; Volman et al., 2004). Die aktuelle Diskussion über Unterschiede im computerbezogenen Nutzungsverhalten und Wissen von Jungen und Mädchen ist verbunden mit Bemühungen, Jungen und Mädchen in ihren computerbezogenen Kompetenzen gleichermaßen zu fördern. Als Ursache für die beschriebenen Unterschiede werden zunehmend motivationale Faktoren in den Mittelpunkt gerückt: Mädchen nutzen seltener als Jungen die sich ihnen bietenden Gelegenheiten mit einem Computer (bspw. außerhalb der Schule), was auf ein niedrigeres computerbezogenes Selbstkonzept und Interesse der Mädchen zurückgeführt wird (Margolis & Fisher, 2003; Vekiri & Chronaki, 2008). Obwohl motivationale Unterschiede im Umgang mit Computern bereits gut dokumentiert sind, gibt es nur wenige Studien, welche sowohl motivationale als auch affektive Faktoren in Zusammenhänge zum computerbezogenen Nutzungsverhalten und Wissen darstellen und in einen theoretischen Rahmen einbetten (Vgl. Sainz & Eccles, 2012). In Ergänzung an bisherige Forschungsbefunde werden zunächst sowohl motivationale (Selbstkonzept, Interesse) als auch affektive Faktoren (Angst) in Relation zum außerschulischen Nutzenverhalten von und Wissen über Computer betrachtet und in einen motivationstheoretischen Rahmen eingebunden. Die Zusammenhänge werden dann differenziert für Jungen und Mädchen untersucht. Die Ergebnisse einer querschnittlich angelegten Studie von 446 SchülerInnen geben erste Hinweise auf entsprechend zu differenzierende Zusammenhänge. Mittels Strukturgleichungsmodellen (Mediations- und Moderationsanalysen) konnte gezeigt werden, dass der Zusammenhang von computerbezogenen Selbstkonzept und Nutzung des Computers außerhalb der Schule primär durch das Interesse am Computer sowie computerbezogener Angst vermittelt wird. Hingegen mediiieren die betrachteten motivationalen und affektiven Faktoren die Zusammenhänge zwischen computerbezogenen Selbstkonzept und Wissen nicht. Die Zusammenhänge werden in Multigruppen-Analysen für Jungen und Mädchen differenziert betrachtet. Demnach ist insbesondere das Interesse am Computer im unterschiedlichen Ausmaß sowohl bei Jungen als auch bei Mädchen ein wichtiger Mediator. Die Ergebnisse werden vor dem Hintergrund eines motivationalen theoretischen Rahmens (bspw. Erwartung-Wert-Theorie, Eccles et al., 2002) sowie eingebettet in eine Diskussion über die Relevanz bestehender Unterschiede für die zukünftige Kompetenzentwicklung von Jungen und Mädchen im Bereich der Informations-Kommunikationstechnologien erörtert.

Eccles, S.J., & Wigfield, A. (2002). Motivational Beliefs, Values & Goals. Annual Review of Psychology, 53, 109-132.

Margolis, J., & Fisher, A. (2003). Unlocking the clubhouse: Women in computing. Cambridge, MA: The MIT Press.

Sáinz, M., & Eccles, J. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. Journal of Vocational Behavior, 80, 486-499.

Vekiri, I., & Chronaki, A. (2008). Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. Computers & Education, 51(3), 1392-1404.

Volman, M., van Eck, E., Heemskerck, I., & Kuiper, E. (2005). New technologies, new differences. Gender and ethnic differences in pupils' use of ICT in primary and secondary education. Computers & Education, 45(1), 35-55.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G131



## **Kulturelles Kapital im 21. Jahrhundert – Ein Index zur Beschreibung digitaler Lernressourcen von Grundschülerinnen und Grundschülern**

Im Kontext der Diskussion um das Abschneiden von Schülergruppen verschiedener sozialen Lagen und im Gesamtkontext der Diskussion um Chancengerechtigkeit im Bildungssystem wurde der Erforschung von sozialen Disparitäten innerhalb der empirischen Schulleistungsfor-schung in den letzten 20 Jahren besondere Aufmerksamkeit zuteil. Diesbezüglich ist es neben der TIMSS-Studie (Baumert, Bos & Lehmann, 2000) und der PISA-Studie (Baumert et al., 2001) auch anderen nationalen Leistungsvergleichen gelungen, Unterschiede in den Leistungen der Schülerinnen und Schüler erstmals für die Bundesrepublik Deutschland auf einer breiten Datenbasis zu eruieren und auf Hintergrundmerkmale (z.B. Geschlecht, Migrationsstatus oder soziale Lage) der Schülerinnen und Schüler zurückzuführen. Die Perspektive, die für die Erfassung und Beschreibung von sozialen Disparitäten im Bildungswesen eingenommen wird, geht auf die theoretischen Konzepte Pierre Bourdieus zurück, der auf unterschiedliche Ausstattung von Personen und Personengruppen (z.B. Schülerfamilien) mit ökonomischem, kulturellem und sozialem Kapital referiert (Bourdieu, 1983, 1986). Auf dieser Grundlage wurden in den verschiedenen Schulstudien Messinstrumente entwickelt, die die Verfügbarkeit vorgenannter kulturellen Güter operationalisieren und somit u.a. für Deutschland soziale Disparitäten im Bildungssystem sichtbar machen sollen (vgl. Bos, Stubbe & Buddeberg, 2010; Ehmke & Siegle, 2005). Als Beispiel kann in diesem Kontext die Frage nach der Anzahl der Bücher im Haushalt, der Home Resources for Learning Index (HRL-Index) sowie der Index für den Economic, Social and Cultural Status (ESCS) genannt werden, die in IGLU und TIMSS sowie in PISA als Proxivariable für das kulturelle Kapital zur Verfügung stehen (vgl. Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Bos, Wendt, Köller & Selter, 2012; Klieme et al., 2010).

Im vorliegenden Beitrag soll der Frage nachgegangen werden, ob eine Ergänzung vorhandener Instrumente zur Erfassung von Hintergrundmerkmalen von Schülerinnen und Schülern durch den gesellschaftlichen und technischen Wandel hin zu einer Informations- und Wissensgesellschaft notwendig ist oder zumindest zusätzliche Erklärungskraft durch die Hinzunahme weiterer Variablen erreicht wird, was zukünftig in einer "wissens- und technologiebasierten Gesellschaft" (Drechsel, Prenzel & Seidel, 2009, S. 378) von Bedeutung sein wird. Dazu wird auf Grundlage des HRL-Index auf der Datenbasis von TIMSS 2011 und IGLU 2011 ein Index entwickelt, der die digitalen Lernressourcen von Kindern abbildet (Home Digital Resources Index [HDR-Index]). Durch einen regressionsanalytischen Vergleich wird die Qualität der Indizes als Prädiktoren für die mathematischen und naturwissenschaftlichen Fähigkeiten von Grundschülerinnen und Grundschülern evaluiert.

Die forschungsleitenden Fragestellungen, die sich auf die Evaluation der Qualität des HDR-Index' beziehen, können wie folgt zusammengefasst werden:

1. Weisen die Fit-Parameter auf eine ausreichende Güte des IRT-Modells hin?
2. In welchem Ausmaß korrelieren der HDR-Index und die mathematischen und naturwissenschaftlichen Kompetenzen? Wie viel Variation in den Schülerleistungen kann durch den HDR-Index erklärt werden?
3. Wie verhält sich die Varianzaufklärung des HDR-Index verglichen mit dem HRL-Index?

Auf Grundlage der Schüler- und Elternangaben von 188.058 Grundschülerinnen und Grundschülern aus TIMSS wird zunächst ein IRT-basierter HDR-Index berechnet. Um die poly-tomen Variablen im Index adäquat im Modell abbilden zu können, wird mit Rückgriff auf das Software-Programm ConQuest ein Partial Credit-Modell appliziert (Masters, 1982). Die Schätzung von Itemparametern wurde auf Grundlage eines Maximum-Likelihood-Ansatzes (Bock & Aitkin, 1981) und Personenparameter auf Basis einer Weighted Likelihood Estimation (Warm, 1989) realisiert. Die regressionsanalytischen Teile des Papiers wurden mit dem IDB-Analyzer durchgeführt, der für Sekundäranalysen im Zusammenhang mit IGLU/TIMSS von der IEA bereitgestellt wird (Rutkowski, Gonzalez, Joncas & von Davier, 2010).

Insgesamt zeigen die Fitstatistiken des IRT-Modells, dass sämtliche im Modell verarbeiteten Items guten Fit aufweisen und der HDR-Index auf internationalem Niveau moderat mit Schulleistungen korreliert. Ein regressionsanalytischer Vergleich der beiden Indizes zeigt, dass die Vorhersagekraft des Index in verschiedenen Ländern unterschiedlich ausgeprägt ist: Einerseits zeigen sich Länder, in denen der 'konservative' Index mehr Variation in den Schülerleistungen erklärt (z.B. Deutschland). Andererseits werden Länder offensichtlich, in denen ein umgekehrtes Befundmuster dominiert.

Baumert, J., Bos, W., & Lehmann, R. (Hrsg.). (2000). TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske+Budrich.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., et al. (Hrsg.). (2001). PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443-459.

Bos, W., Stubbe, T. C., & Buddeberg, M. (2010). Gibt es eine arbeitsbedingte Bildungsbenachteiligung? Die Operationalisierung verschiedener Indikatoren der sozialen Herkunft von Schülerinnen und Schülern in der empirischen Bildungsforschung. In D. H. Rost (Hrsg.), *Intelligenz, Hochbegabung, Vorschulbildung* (S. 165-208). Münster: Waxmann.

Bos, W., Tarelli, I., Bremerich-Vos, A., & Schwippert, K. (Hrsg.). (2012). IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.

Bos, W., Wendt, H., Köller, O., & Selter, C. (Hrsg.). (2012). TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183-198). Göttingen: Schwartz.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Hrsg.), *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood.

Drechsel, B., Prenzel, M., & Seidel, T. (2009). Nationale und internationale Schulleistungstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 354-380). Heidelberg: Springer.

Ehmke, T., & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. Indikatoren der sozialen Herkunft bei der Quantifizierung von sozialen Disparitäten. *Zeitschrift für Erziehungswissenschaft*, 8(4), 521-540.

Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., et al. (Hrsg.). (2010). PISA 2009. Bilanz nach einem Jahrzehnt. Münster: Waxmann.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142-151.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427-450.

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G131



## **Führt Argumentieren zu Wissenserwerb?**

### **Das Modell der argumentativen Wissenskonstruktion in Online-Diskussionen**

Online-Diskussionen werden in der Hochschullehre eingesetzt, um die aktive Beteiligung von Studierenden in Vorlesungen und Seminaren zu verbessern (Fischer, Stegmann, Wecker, & Kollar, 2011). Implizit steht dahinter die Annahme, dass Lernende in Online-Diskussionen ihre Standpunkte formulieren und begründen, die Position ihrer Lernpartner verstehen und kritisieren und Konflikte lösen. Im Rahmenmodell der argumentativen Wissenskonstruktion in Online-Diskussionen von Weinberger und Fischer (2006) werden unterschiedliche Dimensionen von Online-Diskussionen spezifiziert, die den Wissenserwerb durch Online-Diskussionen erklären sollen:

Die Partizipation stellt eine Voraussetzung für argumentative Wissenskonstruktion in Online-Diskussionen dar. Die Partizipation bedeutet hier, wie häufig sich Lernende in einer Lernumgebung anmelden und in welchem Ausmaß sie dort aktiv sind. Die Qualität einzelner Argumente wird anhand des Modells der Argumentation von Toulmin (1958) konzeptualisiert. Die Qualität der Argumentsequenzen wird im vorliegenden Modell anhand des Argument-Gegenargument-Synthese-Schemas bestimmt. Transaktivität beschreibt den Grad, zu dem Lernende auf den Wissensbeiträgen eines anderen Lernenden aufbauen bzw. darauf Bezug nehmen. Die epistemische Aktivität beschreibt, welche Konzepte Lernende anwenden und miteinander verknüpfen, bzw. inwieweit Lernende vom Thema abschweifen, weniger relevante Aspekte der Lernumgebung koordinieren oder welche Fehlkonzepte sie konstruieren.

Fragestellung: Vor dem Hintergrund des Modells der argumentativen Wissenskonstruktion wird in der Studie die Frage untersucht, inwiefern Partizipation, Transaktivität, Qualität der Argumentation (einzelne Argumente und Argumentsequenzen), epistemische Aktivitäten und individueller Wissenserwerb zusammenhängen. Es soll überprüft werden, inwiefern das Modell der argumentativen Wissenskonstruktion in Online-Diskussionen in der Lage ist den inhaltlichen Wissenserwerb zu erklären.

Methode: Für die Studie wurden 36 Online-Diskussionen von Dreiergruppen (N = 108 Lernende) analysiert. Die Teilnehmerinnen und Teilnehmer waren Erstsemester-Studierende der Pädagogik die im Rahmen einer Lehrveranstaltung an einer virtuellen Lernsitzung teilgenommen haben. Diese Daten sind z. T. Bestandteil von Analysen bereits publizierter Studien (Stegmann et al., 2007; Stegmann et al., 2012; Weinberger et al., 2010; Weinberger et al., 2005). Bei den hier re-analysierten Daten handelt es sich jeweils um die Kontrollbedingungen dieser Studien. Die Aufgabe der Lernenden war es, Problemfälle mit Hilfe einer psychologischen Theorie zu analysieren. Ausgewertet wurden sowohl die individuellen Analysen nach der Online-Diskussion als auch die Online-Diskussionen der Dreiergruppen. Die individuellen Analysen und Online-Diskussionen wurden von trainierten Kodierern in Segmente unterteilt. Die Segmentierung wurde auf der Ebene von Propositionen durchgeführt. Bezüglich der Segmentierung erreichten die Kodierer nach einem intensiven Training eine zufriedenstellende Übereinstimmung von 84%. Die Segmente wurden auf mehreren Dimensionen kodiert. Auf Grundlage der Kodierung wurden die folgenden Variablen gebildet (Cohens Kappa  $\geq$  .70): Partizipation (Anzahl Zeichen), epistemische Aktivitäten (Anzahl der korrekten Propositionen während der Onlinediskussion), domänenspezifisches Wissen (Anzahl der korrekten Propositionen im Nachtest), Transaktivität (Anzahl konflikt- und integrationsorientierter Konsensbildungen), Qualität der einzelnen Argumente (Anzahl der begründeten und/oder relativierten Behauptung) und Qualität der Argumentsequenzen (Anzahl der Gegenargumente).

Ergebnisse: Getestet wurde ein Pfadmodell in dem die Partizipation die Transaktivität, die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen vorhersagt. Die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen wurden wiederum als Prädiktoren für die epistemische Aktivität verwendet. Die epistemische Aktivität wurde schließlich als Prädiktor für den Wissenserwerb verwendet. Das Pfadmodell wurde mit Hilfe von AMOS überprüft. Das beschriebene Modell erreichte eine akzeptable Passung. Die Partizipation sagt signifikant und mit jeweils großer Effektstärke die Transaktivität (27% Varianzaufklärung), die Qualität der einzelnen Argumente (23% Varianzaufklärung) und die Qualität der Argumentsequenzen (15% Varianzaufklärung) vorher. Die Transaktivität, die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen klären 36% der Varianz der epistemischen Aktivität auf. Die Qualität der einzelnen Argumente ist ein signifikanter Prädiktor mit einer großen Effektstärke. Die Transaktivität ist ebenfalls ein signifikanter Prädiktor, weist jedoch nur eine kleine Effektstärke auf. Die Qualität der Argumentsequenzen weist ebenfalls eine kleine, nicht signifikante Effektstärke auf. Die epistemische Aktivität klärt signifikant 11% der Varianz des domänenspezifischen Wissens im Nachtest auf.

Fischer, F., Stegmann, K., Wecker, C. & Kollar, I. (2011). Online-Diskussionen in der Hochschullehre: Kooperationskripts können das fachliche Argumentieren verbessern. Zeitschrift für Pädagogik, 3, 326-337.

Stegmann, K., Wecker, C., Weinberger, A. & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. Instructional Science, 40(2), 297-323.

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning, 2(4), 421-447.

Toulmin, S. (1958). The uses of argument. Cambridge: Cambridge University Press.

Weinberger, A., Ertl, B., Fischer, F. & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. Instructional Science, 33, 1-30.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. Computers & Education, 46, 71-95.

Weinberger, A., Stegmann, K. & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). Computers in Human Behavior, 26, 506-515.

Chair(s): Anke Schmitz (Bergische Universität Wuppertal)

Diskutant/in: Cordula Artelt (Universität Bamberg)

Die Fähigkeit, geschriebene Worte und Texte verstehen zu können, ist ein universelles Kulturwerkzeug zur Teilhabe an vielen Bereichen des gesellschaftlichen Lebens (Klieme et al., 2010). Die differenzierte Diagnostik und Förderung individueller Leseleistungen stellt daher einen zentralen Bestandteil des Bildungsauftrages aller grund- und weiterbildenden Schulen dar.

Das Symposium thematisiert verschiedene lerner- und textseitige Aspekte des Leseverständnisses, mit einem Schwerpunkt auf Diagnostik und Intervention in der Primar- und Sekundarstufe. Die in den folgenden Beiträgen vorgestellten Forschungsfragen und Untersuchungsmethoden entspringen unterschiedlichen Disziplinen wie der Erziehungswissenschaft, Psychologie, Lehr-Lernforschung sowie der Sprachdidaktik. Entsprechend dieser Interdisziplinarität weisen sie ein breites Spektrum unterschiedlicher Zielsetzungen, Herangehensweisen sowie Untersuchungsmethoden auf, um der Komplexität des gemeinsamen Forschungsschwerpunktes Leseverständnis Rechnung zu tragen.

Im ersten Beitrag wird Frau Knoepke, ausgehend von einer kognitionspsychologischen Beschreibung der Teilprozesse der Leseverstehens auf Wort-, Satz- und Textebene, das computergestützte Testverfahren ProDi-L vorstellen, das interindividuelle Leistungsunterschiede im Leseverstehen von Schülerinnen und Schülern der 1. bis zur 4. Klasse prozessorientiert und differenziert erfasst. Frau Seifert wird im zweiten Beitrag ein längsschnittlich untersuchtes Interventionsprogramm vorstellen, das die Lesefähigkeiten von Grundschulkindern durch gleichzeitige Erweiterung der sprachlichen Fähigkeiten (LARS = Language And Reading Skills) fördert und in den Schulunterricht integriert werden kann. Frau Müller wird individuelle und interpersonale Effekte eines peer-gestützten Lesestrategietrainings in Klasse 4 zur Förderung des Leseverständnisses vorstellen. Frau Schmitz stellt abschließend eine Querschnittsuntersuchung vor, in welcher die differenzielle Wirkung von globaler Textkohäsion zur Förderung des Sachtextverständnisses bei Schülerinnen und Schülern der 9. Klasse analysiert wird. Es folgt eine abschließende Diskussion der zentralen Ergebnisse der Studien durch Frau Prof. Artelt.

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G135

## Prozessorientierte Diagnostik des Leseverstehens

### Theoretischer Hintergrund

Das Verstehen geschriebener Texte ist eine Schlüsselkompetenz, deren Erwerb für die erfolgreiche schulische sowie außerschulische Entwicklung von Kindern unabdingbar ist. Eine frühzeitige Diagnose von Leistungsschwächen und -defiziten im Leseverstehen ist daher sinnvoll, um frühzeitig Interventions- und Präventionsmaßnahmen einleiten zu können. Dabei sind die Ursachen für schwache Leseleistungen sehr heterogen. Sowohl Defizite im Bereich der visuellen Worterkennung (z.B. Perfetti, 1985; Richter, Isberner, Naumann, & Neeb, in press) als auch im Bereich syntaktischer und satzsemantischer Fertigkeiten (z.B. Graesser, Hoffman, & Clark, 1980) sowie in Lesestrategien auf der Textebene (z.B. Oakhill, Cain, & Bryant, 2003) können in schlechtem Leseverständnis resultieren. Die kognitionspsychologische Lesekompetenzforschung hat eine Reihe verschiedener kognitiver Teilprozesse auf Wort-, Satz-, und Textebene identifiziert, die gemeinsam das Leseverstehen konstituieren (Richter & Christmann, 2009; Müller & Richter, in Druck). Diese Unterscheidung von Teilprozessen des Leseverstehens muss in einem Diagnostikum, das interindividuelle Unterschiede in den Leseleistungen von Schülern/-innen erfassen soll, berücksichtigt werden: (1) Phonologische Rekodierungsprozesse, (2) Orthographische Vergleichsprozesse, (3) Zugriff auf Wortbedeutungen, (4) syntaktische und (5) semantische Integrationsprozesse sowie (6) Prozesse der Kohärenzbildung.

Prozessorientierte Diagnostik des Leseverstehens (ProDi-L): In diesem Beitrag soll das computergestützte Testverfahren ProDi-L (ProDi-L: Prozessbezogene Diagnostik des Leseverstehens im Grundschulalter; Richter, Isberner, Naumann, & Kutzner, 2012; Richter, Naumann, Isberner, Neeb, & Knoepke, in Druck) vorgestellt werden, das interindividuelle Leistungsunterschiede im Leseverstehen von Grundschüler/-innen der 1. bis zur 4. Klasse prozessorientiert und differenziert erfasst. Der computeranimierte Außerirdische "Reli" begleitet die Kinder durch sechs verschiedene Arten von Leseaufgaben, die auf jeweils einen Teilprozesse des Leseverstehens zugeschnitten sind: (1) Phonologische Vergleichsaufgabe, (2) Lexikalische Entscheidung, (3) Semantische Verifikationsaufgabe, (4) Grammatikalitätsurteile, (5) Sinnhaftigkeitsurteile zu einzelnen Sätzen und (6) Sinnhaftigkeitsurteile zu Satzfolgen. Als Maße der Effizienz werden die Antwortgenauigkeit, die die Zuverlässigkeit des untersuchten Prozesses abbildet, sowie die Antwortgeschwindigkeit, die den Routinisierungsgrad abbildet, erfasst. Es wurden Testitems mit unterschiedlicher Schwierigkeit (Lösungswahrscheinlichkeit bzw. mittlere Reaktionszeit) konstruiert, indem verschiedene aus psycholinguistischer Sicht schwierigkeitsgenerierende Merkmale variiert wurden. Auf diese Weise können sowohl inter- als auch intraindividuelle Leistungsunterschiede und -entwicklungen in Teilprozessen des Leseverstehens über den Verlauf der Grundschulzeit beobachtet werden. Zusätzlich wurden strikt parallel gestaltete Hörverstehensaufgaben (ProDi-H) konstruiert, in denen die kritischen Stimuli nicht visuell, sondern auditiv präsentiert werden. Dies ermöglicht einen direkten Vergleich in der Entwicklung von Lese- und korrespondierenden Hörverstehensfertigkeiten sowie die Untersuchung komplexer Zusammenhänge zwischen ihnen. Überprüfung der Simple View of Reading mithilfe von ProDi-L und ProDi-H: Neben der Schuleignungs- und Förderdiagnostik lassen sich ProDi-L und ProDi-H auch für die Bearbeitung sprach- und pädagogisch-psychologischer Fragestellungen einsetzen. Dies soll anhand einer Untersuchung demonstriert werden, in der die Gültigkeit der Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) für deutsche Grundschüler/-innen methodisch stringent überprüft wurde (Knoepke, Richter, Isberner, Neeb, & Naumann, in Druck). Nach der Simple View of Reading lassen sich individuelle Unterschiede im Leseverstehen durch das Produkt aus Dekodierfähigkeit (visuelle Worterkennung) und allgemeinem Sprachverstehen beschreiben. Um die Gültigkeit dieser Annahme zu überprüfen, wurden Lese- und Hörverstehen durch Aufgabe 6 (Sinnhaftigkeitsurteile zu Satzfolgen; ProDi-L und ProDi-H) operationalisiert, die visuelle Worterkennung durch Aufgabe 2 (Lexikalische Entscheidung, ProDi-L). In einer korrelativen Untersuchung mit 50 Dritt- und 74 Viertklässlern konnte gezeigt werden, dass das Produkt von Worterkennung und Sprachverstehen, entgegen den Annahmen der Simple View of Reading, kein Inkrement gegenüber ihrer additiven Verknüpfung liefert. Zudem konnten nur rund ein Drittel der Varianz im Leseverstehen durch individuelle Unterschiede in Worterkennung und Hörverstehen aufgeklärt werden, was auf weitere am Leseverstehen beteiligte und in der Simple View of Reading nicht berücksichtigte Prozesse schließen lässt. Insgesamt bietet die Untersuchung einen Einblick in die vielseitige Anwendbarkeit der vorgestellten Messinstrumente ProDi-L und ProDi-H.

Gough, P.B & Tunmer, W.E. (1986). Decoding, reading, and reading disability. Remedial and Special Education, 7, 6–10.

Graesser, A.C., Hoffman, N.L. & Clark, L.F. (1980). Structural components of reading time. Journal of Verbal Learning and Verbal Behavior, 19, 135-151.

Hoover, W.A. & Gough, P.B. (1990). The simple view of reading. Reading and Writing: An Interdisciplinary Journal, 2, 127–160.

Knoepke, J., Richter, T., Isberner, M.-B., Neeb, Y. & Naumann, J. (in Druck). Leseverstehen = Hörverstehen X Dekodieren? Ein stringenter Test der Simple View of Reading bei deutschsprachigen Grundschulkindern. In A. Redder & S. Weinert (Hrsg.), Sprachförderung und Sprachdiagnostik: Interdisziplinäre Perspektiven. Münster: Waxmann.

Oakhill, J.V., Cain, K. & Bryant, P.E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. Language and Cognitive Processes, 18, 443-468.

Perfetti, C.A. (1985). Reading ability. New York: Academic Press.

Richter, T., Naumann, J., Isberner, M.-B., Neeb, Y. & Knoepke, J. (in Druck). ProDi-L: Prozessbasierte Diagnostik von Lesefähigkeiten bei Grundschulkindern [Computergestützter Test]. Göttingen: Hogrefe.

Richter, T., Isberner, M.-B., Naumann, J. & Kutzner, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern. Zeitschrift für Pädagogische Psychologie, 26, 313-331.

Richter, T., Isberner, M.-B., Naumann, J. & Neeb, Y. (in press). Lexical quality and reading comprehension in primary school children. Scientific Studies of Reading.





## Ein Programm zur differenzierten Förderung des Lesesinnverstehens in heterogenen Klassen

Theoretischer Hintergrund: Etwa 13% der deutschen und 16% der österreichischen Kinder in der vierten Klasse nur grundlegendste Leseanforderungen erfüllen, d.h. sie können lediglich explizit angegebene Einzelinformationen in Texten identifizieren (Bos et al., 2007). Als eine zentrale Erklärungsvariable für schwache Lesekompetenzen bei Kindern werden Sprachdefizite (u.a. Wortschatz- und Grammatikprobleme) in der Unterrichtssprache Deutsch angeführt (Baumert, 2011). Dass Schwierigkeiten beim Schriftspracherwerb und das sprachliche Kompetenzniveau einander bedingen, ist durch einige Forschungsarbeiten belegt (z.B. Bialystok, 2007; Bishop & Snowling, 2004). So kann einerseits Lesen die sprachlichen Fertigkeiten verbessern (z.B. Snowling & Hulme, 2005; Stanovich, 1993), andererseits sind Fortschritte im Lesen stark von den sprachlichen Kompetenzen abhängig (Snowling & Hulme, 2005). Insbesondere das Lesesinnverstehen hängt in hohem Maße von den sprachlichen Fertigkeiten der Kinder ab (z.B. Perfetti, Landi, & Oakhill, 2005). Durch die nachgewiesene enge Interdependenz zwischen den Lese- und Sprach-Fertigkeiten haben einige Kinder (u.a. Kinder mit Migrationshintergrund, Kinder mit Sprachentwicklungsstörungen) ein erhöhtes Risiko, unzureichendes Lesesinnverstehen zu entwickeln. Damit die Entwicklung der schulischen und beruflichen Laufbahn von lesechwachen Kindern positiv beeinflusst werden kann, sollten Schwächen so früh wie möglich diagnostiziert und intensive, individuell abgestimmte Fördermaßnahmen angeboten werden (Gasteiger-Klicpera & Fischer, 2008; Torgesen, 2005). Um dies leisten zu können, ist ein fundiertes Verständnis der Leseentwicklung bei Kindern notwendig sowie ein sicheres diagnostisches Erkennen der einzelnen Teilfähigkeiten in den sich entwickelnden Lesestrategien.

Um die Stabilität der Effekte langfristig zu gewährleisten, ist es ratsam, die Förderung in den Schulalltag zu implementieren und somit zu einem veränderten Unterrichtskonzept im Sinne einer differenzierten, den Leistungen der Schüler/-innen entsprechenden Förderung zu gelangen (Gasteiger-Klicpera & Fischer, 2008; Strickland, 2002).

Innerhalb des Klassenverbandes stellt die individuelle Förderung der Schüler/-innen aufgrund sprach- und leistungsheterogener Zusammensetzung der Klasse für Lehrende oft eine große Herausforderung dar. Diese Heterogenität erfordert eine Differenzierung des Unterrichts den individuellen Bedürfnissen der Kinder entsprechend. Bezüglich des Leseunterrichts kommt dem Einbezug sprachfördernder Ansätze im inklusiven Klassensetting aufgrund der engen Interdependenz zwischen Sprach- und Lesefertigkeiten eine große Bedeutung zu. Insbesondere sollte die Förderung des Lesesinnverstehens durch Wortschatzarbeit bereichert werden (Baumert, 2011; Rupley, Logan, & Nichols, 1999).

Fragestellung: Innerhalb der Interventionsstudie LARS (Language And Reading Skills) wurde ein differenziertes Förderprogramm in den Schulalltag implementiert, um die Lese- und sprachlichen Fähigkeiten der Kinder zu fördern. Im Rahmen einer Evaluationsstudie wurde die Wirksamkeit des Programms bewertet. Es wurde untersucht, ob Kinder der zweiten Schulstufe, die mit dem Förderprogramm unterrichtet wurden, nach ca. einem Schuljahr eine bessere Leistungsentwicklung im Leseverständnis und in der Lesegeschwindigkeit zeigten als Kinder, die regulär beschult wurden.

Methode: Die Studie ist in einem quasi-experimentellen Prä-Posttest-Design mit Wartekontrollgruppe angelegt (N = 377 Kinder aus steirischen Grundschulklassen mit mind. 30% Migrationsanteil). 159 Kinder wurden zwei Mal wöchentlich im Rahmen der Deutsch- und Sachunterrichtsstunden (insgesamt im Durchschnitt 47,77 Schulstunden), abhängig von ihrem zuvor getesteten Leistungsniveau, mit in vier Schwierigkeitsstufen differenziertem Fördermaterial unterstützt. Weitere 218 Kinder dienten als Wartekontrollgruppe. Vor und nach der Interventionsphase wurden die literalen Fähigkeiten mittels standardisierter Testverfahren erhoben (ELFE 1-6; SLRT II).

Ergebnisse: Prä-Posttest-Vergleiche konnten zeigen, dass sich die geförderten Kinder im Vergleich zur Kontrollgruppe im Lesesinnverstehen auf Satzebene signifikant besser entwickelten ( $F(1,356) = 8.63, p < .01$ ; partielles  $\eta^2 = .02$ ), allerdings nicht auf Wortebene ( $F(1,356) = 1.05, n.s.$ ). In der Lesegeschwindigkeit zeigten sich deutliche Verbesserungen bei allen Kindern, diese unterschieden sich jedoch nicht zwischen der Interventions- und Wartekontrollgruppe (Interaktionseffekte Zeit x Gruppe im Wortlesen:  $F(1,356) = 0.15, n.s.$ ; Pseudowortlesen:  $F(1,356) = 0.18, n.s.$ ).

Das Leseförderprogramm LARS stellt somit eine wirksame Möglichkeit dar, eine Lese- und Sprachförderung in den Unterricht heterogener Klassen zu implementieren. Auch für Kinder mit eher schlechten Ausgangsbedingungen konnte durch eine systematische und differenzierte Sprach- und Leseförderung eine Möglichkeit geschaffen werden, ihre Lesekompetenzen weiter zu entwickeln.

Baumert, J. (2011). Expertenrat „Herkunft und Bildungserfolg“. Empfehlungen für bildungspolitische Weichenstellungen in der Perspektive auf das Jahr 2020. Berlin: Max Planck-Institut für Bildungsforschung Berlin.

Bialystok, E. (2007). Acquisition of literacy in bilingual children: A framework for research. *Language Learning*, 57, 45–77.

Bishop, D.V.M. & Snowling, M.J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, 130, 858–888.

Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.) (2007). IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.

Gasteiger-Klicpera, B. & Fischer, U. (2008). Evidenzbasierte Förderung bei Lese-Rechtschreibschwierigkeiten. In M. Fingerle & S. Ellinger (Hrsg.), *Sonderpädagogische Förderprogramme im Vergleich* (S. 67-84). Stuttgart: Kohlhammer.

Lenhard, W. & Schneider, W. (2006). ELFE 1-6: Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.

Moll, K. & Landerl, K. (2010). SLRT II – Verfahren zur Differentialdiagnose von Störungen der Teilkompetenzen des Lesens und Schreibens. Bern: Hans Huber.

Perfetti, C.A., Landi, N. & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Hrsg.), *The Science of Reading* (S. 227-247). Malden: Blackwell.

Rupley, W.H., Logan, J.W. & Nichols, W. D. (1999). Vocabulary instruction in a balanced reading program. *The Reading Teacher*, 52, 336-346.

Snowling, M.J. & Hulme, C. (2005). Learning to read with language impairment. In M.J. Snowling & C. Hulme (Hrsg.), *The Science of Reading* (S. 397-412). Malden: Blackwell.

Stanovich, K.E. (1993). Does reading make you smarter? Literacy and development of verbal intelligence. In H. Reese (Hrsg.), *Advances in child development and behavior* (S. 133-180). Orlando/FL: Academic Press.

Strickland, D.S. (2002). The importance of effective early intervention. In A.E. Farstrup & S. Samuels (Hrsg.), *What research has to say about reading instruction* (S. 69-86). DE: International Reading Association.

Torgesen, J.K. (2005). Recent discoveries on remedial interventions for children with dyslexia. In M.J. Snowling & C. Hulme (Hrsg.), *The Science of Reading* (S. 521-537). Malden: Blackwell.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G135



## **Peer-gestützte Leseförderung: Individuelle und interpersonale Effekte eines Lesestrategietrainings in Klasse 4**

Theoretischer Hintergrund: Wie durch die Internationale Grundschul-Lese-Untersuchung IGLU wiederholt belegt werden konnte, erreichen etwa 15% der Viertklässler/-innen in Deutschland nur Leseleistungen auf den niedrigsten Kompetenzstufen (Bos, Bremerich-Vos, Tarelli, & Valtin, 2012). Diese Kinder haben große Schwierigkeiten, verschiedene Informationen eines Textes zueinander in Beziehung zu setzen, um ein kohärentes Textverständnis aufzubauen. Zudem nutzen sie weder kognitive noch metakognitive Lesestrategien, um ihr Textverständnis zu überprüfen (Cain, Oakhill, & Bryant, 2004). Ein vielfach diskutiertes und erprobtes Verfahren zur Förderung des Leseverständnisses in der Grundschule sind Lesestrategietrainings (z.B. Guthrie et al., 2004; National Reading Panel, 2000; Palincsar & Brown, 1984). Ziel dieser Interventionen ist die Förderung des sinnentnehmenden Lesens und des selbstregulierten Umgangs mit Texten (von Suchodoletz, 2010). Das peer-gestützte Lernen in Dyaden aus einem leseschwachen und einem lesestarken Kind gilt dabei als effektive Methode zur Umsetzung von Lesestrategietrainings (z.B. Fuchs, Fuchs, Mathes, & Simmons, 1997; Topping, 2001; Van Keer & Verhaeghe, 2005). Doch trotz der dyadischen Gestaltung der Trainings erfolgt die Datenauswertung üblicherweise getrennt für die Gruppen der schwachen und starken Leser/-innen, ohne der durch das peergestützte Lernsetting bedingten Interdependenz Rechnung zu tragen.

Fragestellung: In diesem Beitrag sollen die Effekte eines peer-gestützten Lesestrategietrainings zur Förderung des Leseverständnisses in Klasse 4 vorgestellt und den Effekten innerhalb einer Kontrollgruppe gegenübergestellt werden. Ein besonderes Augenmerk liegt dabei auf der dyadischen Datenanalyse mit dem Akteur-Partner-Interdependenz-Modell (APIM, Kenny, 1996). Neben der Analyse der individuellen Trainingseffekte leseschwacher und lesestarker Kinder ermöglicht dieses Modell auch die Analyse interpersonaler Effekte, um der Interdependenz innerhalb der Dyaden gerecht zu werden.

Methode: Aufgrund ihrer Leistungen im Leseverständnistest ELFE 1-6 (Lenhard & Schneider, 2006) wurden 136 Viertklässler/-innen zur Teilnahme an der experimentellen Prä-Post-Test Studie ausgewählt. Pro Klasse wurden die fünf Kinder mit den schlechtesten Leseleistungen als Tutand(inn)en und die fünf Kinder mit den besten Leseleistungen als Tutor(inn)en ausgewählt. Die insgesamt 68 Dyaden wurden auf Klassenebene randomisiert der Lesestrategiegruppe (34 Dyaden) oder der Kontrollgruppe (Training des visuell räumlichen Arbeitsgedächtnisses, 34 Dyaden) zugewiesen (Blockversuchsplan mit parallelisierten Gruppen). Die Intervention umfasste jeweils 25 Sitzungen à 45 Minuten. Nach Abschluss der Trainings wurde erneut ELFE 1-6 durchgeführt.

Ergebnisse: Das APIM wurde als Multi-Group-Strukturgleichungsmodell gerechnet mit den ELFE-Prä-Test-Werten als Prädiktoren für die ELFE-Post-Test-Werte.

Für die Gruppe der Tutand(inn)en ergab sich nur in der Kontrollgruppe ein signifikanter und positiver Akteur-Effekt ( $B = 0.50$ ,  $SE = 0.29$ ,  $p < .05$ , einseitig), was auf stabile Leseverständnisleistungen über den Verlauf der Intervention hinweg hindeutet. Für die Tutor(inn)en ergaben sich in beiden Bedingungen signifikante Akteur-Effekte (EG:  $B = 0.94$ ,  $SE = 0.17$ ,  $p < .001$ , einseitig; KG:  $B = 0.43$ ,  $SE = 0.13$ ,  $p < .01$ , einseitig).

Ein signifikanter Partner-Effekt vom Prä-Test der Tutor(inn)en auf den Post-Test der Tutand(inn)en zeigte sich, wie erwartet, nur innerhalb der Strategiegruppe ( $B = 0.31$ ,  $SE = 0.17$ ,  $p < .05$ , einseitig). Demnach konnten die leseschwachen Kinder von den Leseleistungen ihrer Tutor(inn)en bei der Implementierung der Lesestrategien profitieren.

Der durchschnittliche Effekt des Treatments im Vergleich zur Kontrollgruppe wurde für die Tutor(inn)en nur knapp nicht signifikant ( $B = 1.10$ ,  $SE = .69$ ,  $p = .057$ , einseitig), d.h. die leseschwachen Kinder der Strategiegruppen zeigten nach der Intervention tendenziell bessere Leseverständnisleistungen als die leseschwachen Kinder innerhalb der Kontrollgruppe. Für die Tutor(inn)en ergaben sich keine Treatment-Effekte ( $B = -0.93$ ,  $SE = .64$ , n. s., einseitig).

Zusammengefasst liefern die Ergebnisse Hinweise für die Wirksamkeit des Lesestrategietrainings zur Förderung des Leseverständnisses leseschwacher Viertklässler/-innen. Die leseschwachen Kinder konnten von der dyadischen Zusammenarbeit profitieren: Ihr Leseverständnis veränderte sich in Abhängigkeit von den Leseleistungen der Tutor(inn)en. Mithilfe des APIM kann diese Annahme des peer-gestützten Lernens explizit und unter Berücksichtigung der Abhängigkeit der Kinder innerhalb der Dyaden getestet werden.

Bos, W., Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), IGLU 2011: Lesekompetenzen von Grundschulkindern im internationalen Vergleich. Münster: Waxmann.

Cain, K., Oakhill, J. & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96, 31-42.

Fuchs, D., Fuchs, L.S., Mathes, P.G. & Simmons, D.C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174-206.

Guthrie, J.T., Wigfield, A., Barbosa, P., Perencevich, K.C., Taboada, A., Davis, M.H., Scaffidi, N.T. & Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96, 403-423.

Kenny, D.A. (1996). Models of nonindependence in dyadic research. *Journal of Social and Personal Relationships*, 13, 279-294.

Lenhard, W. & Schneider, W. (2006). ELFE 1-6: Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.

Topping, K. (2001). Peer assisted learning: A practical guide for teachers. Newton: Brooklyn Books.

Van Keer, H. & Verhaeghe, J.P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *Journal of Experimental Education*, 73, 291-329.

von Suchodoletz, W. (2010). Konzepte in der LRS-Therapie. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 38, 329-339.



## ***Können global kohäsive Texte das Verständnis von Sachtexten fördern? Interaktionen zwischen Textkohäsion und individuellen Lernvoraussetzungen***

Theoretischer Hintergrund: Die Lehrpläne vieler Unterrichtsfächer belegen, dass Lernende nahezu täglich vor der anspruchsvollen Aufgabe stehen, Wissen aus Sachtexten zu erwerben. Deshalb ist die Verständlichkeit von Unterrichtstexten ein zentraler Faktor für Unterrichtsqualität. Aufgrund der zuletzt fokussierten Verständlichkeit von Texten in medialen Darbietungsformen, des Verstehens von Text-Bild-Kombinationen oder Hypertexten (vgl. Naumann, Richter, Christmann & Groben, 2008), fokussiert diese Studie das Kriterium der Textverständlichkeit von traditionellen gedruckten Sachtexten. Die Ergebnisse sollen im Hinblick auf die Heterogenität unterrichtlicher Lernmilieus neue Erkenntnisse für die Passung von Lernmaterialien an individuelle Lernvoraussetzungen liefern und Empfehlungen für die Konzeption von Unterrichtsmaterialien seitens Schulbuchverlagen oder Lehrpersonen generieren.

Fragestellung: Im Zentrum der Studie steht die übergeordnete Fragestellung, ob eine erhöhte globale Textkohäsionsdichte das Verständnis von Sachtexten fördern kann. Unter globaler Textkohäsion werden linguistische Textverknüpfungen auf der syntaktischen und semantischen Textebene subsumiert, die über benachbarte Sätze hinweg Textpassagen im Textganzen miteinander verbinden (vgl. de Beaugrande & Dressler, 1981). Das Merkmal globale Textkohäsion wird untersucht, da es in seiner isolierten Form bisher nur selten fokussiert und meist Kombinationen globaler und lokaler Textkohäsion undifferenziert betrachtet wurden (vgl. Ozuru, Dempsey, & McNamara, 2009). Zudem wird diesem Merkmal eine verständnisfördernde Wirkung zugesprochen, da es für Lernende einen sogenannten ‚roten Faden‘ zur Orientierung bietet (vgl. Kintsch, 1998; Schnotz, 2006). Allerdings mangelt es an empirischen Studien im deutschen Forschungsraum, um für hier stattfindende unterrichtliche Lernprozesse aufzuklären, ob die im US-amerikanischen Forschungsraum mehrfach nachgewiesene differenzielle Wirkung von Textkohäsion im Hinblick auf individuelle kognitive Lernvoraussetzungen auch auf das Merkmal globale Textkohäsion zutrifft.

Methode: In einem experimentellen Design (between) lesen 700 Lernende der 9. Jahrgangsstufe entweder einen global kohäsiven oder einen global inkohäsiven Sachtext, dessen Manipulation sich an theoretische Grundlegungen nach de Beaugrande und Dressler (1981) anlehnt. Beispiele globaler Kohäsion sind Überschriften, deiktische Textverweise und die Explikation relevanter Textinformationen. Der Test, mit dem das Textverständnis erfasst wurde, orientiert sich an den kognitionspsychologischen Subskalen der Lesekompetenz aus PISA (vgl. OECD, 2010) und beinhaltet Fragen zum Suchen und Extrahieren von Informationen, Aufgaben des Typs Kombinieren und Interpretieren sowie Items, die ein Reflektieren und Bewerten erfordern. Diese Dreiteilung erlaubt eine spezifische Wirkungsanalyse der globalen Kohäsion auf die kognitiven Operationen des Leseverstehens (vgl. Kintsch, 1998). Die Ausprägung des thematischen Vorwissens wird durch einen 10-minütigen Wissenstest erhoben. Weiterhin werden in der Studie motivationale Aspekte wie die Lesemotivation und das Leseselbstkonzept erfasst. Letztere wurden in vergleichbaren experimentell angelegten Studien bis dato ausgeklammert, obwohl sie wichtige Einflussfaktoren des Textverständnisses bilden (vgl. Schaffner, Schiefele & Schneider, 2004). Die Lesefähigkeit wird durch den standardisierten Lesegeschwindigkeits- und Verständnistest LGVT 6-12 erfasst (vgl. Schneider, Schlagmüller & Ennemoser, 2007).

Ergebnisse: Regressionsanalysen des Pretests (N = 154) verdeutlichen, dass die globale Kohäsion insbesondere textbezogenes Verständnis fördern kann, jedoch nicht das Reflektieren und Bewerten eines Textes tangiert. Es zeigten sich bei den textbezogenen Items Interaktionseffekte von Lesefähigkeit, thematischem Vorwissen und globaler Textkohäsion. Während Lernende ohne gute Lesefähigkeiten nicht von globaler Kohäsion profitierten, konnten Lernende mit ausgeprägter Lesefähigkeit die Kohäsivität der Texte für sich nutzen ( $t(4, 140) = .243, p < .01, \Delta R^2 = .02$ ). Erwartungswidrig zeigte sich bei der Berücksichtigung des thematischen Vorwissens, dass auch hier ausschließlich Lernende mit ausgeprägtem Vorwissen profitierten ( $t(4, 140) = .404, p < .001, \Delta R^2 = .03$ ), wohingegen Lernende ohne ausgeprägtes Vorwissen keinen Nutzen aus der globalen Kohäsion ziehen konnten (vgl. Schmitz & Gräsel, 2013). Im Hinblick auf die Förderung des Textverständnisses von Lernenden mit schlechteren Verständnisvoraussetzungen ist zu hinterfragen, weshalb die global kohäsive Variante nicht zu einem deutlich besseren Textverständnis führte. Untersucht werden sollte, ob die mit der Steigerung der globalen Kohäsionsdichte verbundene erhöhte Wortanzahl zu einer kognitiven Überlastung führt und ob motivationale Einflussfaktoren wie die Lesemotivation und das Leseselbstkonzept die Wirkung der Kohäsion bedingen.

de Beaugrande, R. & Dressler, W.U. (1981). Einführung in die Textlinguistik. Konzepte der Sprach- und Literaturwissenschaft. Tübingen: Niemeyer.

Kintsch, W. (1998). Comprehension: A Paradigm for Cognition. Cambridge: Cambridge University Press.

Naumann, J., Richter, T., Christmann, U. & Groeben, N. (2008). Working memory capacity and reading skill moderate the effectiveness of strategy training in learning from hypertext. *Learning and Individual Differences*, 18, 197-213.

Ozuru, Y., Dempsey, K. & McNamara, D.S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Language and Instruction*, 19, 228-242.

Schmitz, A. & Gräsel, C. (2013, eingereicht). Bei welchen Lernenden fördert globale Textkohäsion das Verständnis von Sachtexten? Eine Studie zu Wechselwirkungen zwischen globaler Textkohäsion und kognitiven Verständnisvoraussetzungen. *Unterrichtswissenschaft*.

Schneider, W., Schlagmüller, M. & Ennemoser, M. (2007). Lesegeschwindigkeits- und Verständnistest für die Klassen 6-12 (LGVT 6-12). Hogrefe, Göttingen.

Schnotz, W. (2006). Was geschieht im Kopf des Lesers? Mentale Konstruktionsprozesse beim Leseverständnis aus der Sicht der Psychologie und der kognitiven Linguistik. In: H. Blühdorn, E. Breindl & U.H. Waßner (Hrsg.), *Text-Verstehen. Grammatik und darüber hinaus* (S. 222-238). Walter de Gruyter: Berlin, New York.



Chair(s): Tim Heemsoth (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)

Diskutant/in: Eveline Wuttke (Goethe-Universität Frankfurt am Main)

In konstruktivistischen Lehr-Lern-Theorien wird angenommen, dass das Reflektieren von Fehlern ein großes Lernpotenzial bietet (VanLehn, 1990). Allerdings gibt es bisher nur unzureichend Belege dafür, dass das Reflektieren von Fehlern den Lernprozess tatsächlich unterstützt, außerdem ist unklar wie und unter welchen Bedingungen das Lernen aus Fehlern gelingen kann (Wuttke & Seifried, 2011). Grundsätzlich werden auch geeignete kognitive, affektiv-motivationale und handlungsadaptive Komponenten beim Umgang mit Fehlern als Vorbedingung eines effektiven Lernprozesses angenommen. Vor diesem Hintergrund geht das Symposium den Fragen nach, ob das Reflektieren fremder oder eigener Fehler das Lernen stärker fördert, welche Strategien im Umgang mit Fehlern von Vorteil sind, und welche Einstellungen und Haltungen Lehrkräfte sowie Schülerinnen und Schüler zum Umgang mit Fehlern haben und wodurch diese erklärt werden können. Insgesamt greift das Symposium damit aktuelle Forschungsfragen zum Lernen aus Fehlern auf. Es vereint Arbeiten aus der Wirtschaftspädagogik, der Mathematikdidaktik sowie der pädagogischen Psychologie und thematisiert ökologisch valide Unterrichtsettings ebenso wie Domäne übergreifende Bedingungsanalysen.

Im ersten Beitrag von Lütznier, Förster-Kuschel, Ryssel und Fürstenau wird eine Interventionsstudie präsentiert, die der Frage nachgeht, ob das Lernen betriebswirtschaftlicher Zusammenhänge durch das Korrigieren von Fehlern in fremden Concept Maps oder durch das Korrigieren von Fehlern in eigenen Concept Maps besser gefördert wird.

Im zweiten Beitrag von Heemsoth und Heinze wird ebenfalls eine Interventionsstudie vorgestellt, die sich jedoch ausschließlich auf das Lernen aus eigenen Fehlern konzentriert. Es wird untersucht, durch welche Art der Reflektion eigener Fehler der Aufbau von Wissen in der Bruchrechnung bestmöglich gelingt. Dabei wird eine Lernumgebung, in der die Lernenden das Rational hinter ihren eigenen Fehlern reflektieren, mit einer Lernumgebung, in der lediglich die richtigen Lösungen reflektiert werden, verglichen.

Im dritten Beitrag von Eichelmann, Narciss und Körndle werden die Einstellungen der Schülerinnen und Schüler in Bezug auf das Lernpotenzial, das von der Reflektion fremder Fehler ausgeht, mit den Einstellungen der Lehrkräfte verglichen. Weiterhin wird untersucht, ob typische Fehler in der Bruchrechnung besser erinnert werden, wenn zuvor mit fremden Fehlern oder klassischen Aufgaben in der Bruchrechnung gelernt wurde.

Der abschließende vierte Beitrag von Tulis, Steuer und Dresel konzentriert sich auf die Frage, inwieweit sich kognitive, affektiv-motivationale und handlungsadaptive Komponenten im Umgang mit Fehlern auf unterschiedliche Domänen zurückführen lassen. Der Beitrag erweitert damit bisherige Befunde, die sich fast ausschließlich auf die Domäne Mathematik konzentrierten.

Die Beiträge des Symposiums werden von Eveline Wuttke im Hinblick auf ihre Bedeutung für Theorie und Lehr-Lern-Praxis diskutiert.



## **Lernen aus eigenen vs. Lernen aus fremden Fehlern – Nutzung von Concept Maps im Planspielunterricht**

Fehler sind Teil von Lernprozessen. Um aus Fehlern lernen zu können, genügt es jedoch nicht, Fehler zu machen, sondern es ist wichtig, mit ihnen konstruktiv umzugehen. Lernende sollten hierfür Fehler als solche erkennen, verstehen und die Möglichkeit erhalten, sie zu korrigieren (Oser & Spychiger, 2005). Insbesondere in komplexen Lehr-Lernarrangements, wie z. B. Planspielen, machen Schüler häufig Fehler und es ist wichtig, sie durch einen produktiven Umgang mit Fehlern im Verständnis komplexer Zusammenhänge zu unterstützen. Dafür können Lernstrategien, wie z. B. Concept Mapping, eingesetzt werden. Diese lassen sich mit Lernen aus Fehlern verknüpfen, indem beispielsweise fehlerhafte Concept Maps (sog. Fehlermaps) korrigiert werden.

Eine interessante, aber bislang kaum erforschte Frage, die sich im Zusammenhang mit dem Lernen aus Fehlern im schulischen Kontext stellt, ist, ob Schüler besser aus eigenen Fehlern oder aus fremden Fehlern lernen. Bisherige Studien, die in sehr unterschiedlichen Kontexten durchgeführt worden sind, weisen auf eine Vorteilhaftigkeit des Lernens aus fremden Fehlern hin (vgl. Bartlett, 1982; KC, Staats & Gino, 2012).

Vor diesem Hintergrund wurde in einer Interventionsstudie untersucht, ob sich auch im Zusammenhang mit komplexen Lehr-Lernarrangements und mit Concept Maps das Lernen aus fremden Fehlern vom Lernen aus eigenen Fehlern unterscheidet. Ausgehend von bisherigen Forschungsergebnissen wurde die Hypothese formuliert, dass die Korrektur einer fremden Fehlermap lernwirksamer ist als die Korrektur einer eigenen Fehlermap.

An der Studie nahmen 47 Schüler zweier neunter Klassen einer Mittelschule teil. Der Untersuchung lag ein Pretest-Posttest-Design zugrunde. Die Schüler wurden dabei einer von zwei Treatmentgruppen (Lernen aus eigenen bzw. fremden Fehlermaps) zugewiesen. Alle Teilnehmer spielten zunächst das Planspiel "Easy Business™", welches anhand eines Modellunternehmens betriebswirtschaftliche Zusammenhänge verdeutlicht. Nach der Durchführung erhielt jeder Teilnehmer der Gruppe „eigene Fehlermap“ seine eigene Concept Map, die er vor Beginn des Planspiels erstellt hatte. Diese galt es zu korrigieren. Die Teilnehmer der Gruppe „fremde Fehlermap“ sollten eine fehlerhafte Experten Map korrigieren. Vor Beginn des Planspiels und im Anschluss an die Map-Korrektur nahmen alle Teilnehmer an einem Wissenstest teil.

Im Ergebnis wurde für beide Gruppen ein signifikanter Wissenszuwachs nachgewiesen ( $F = 88.356$ ;  $df = 1$ ;  $p = 0.000$ ;  $d = 1.6$ ). Die Gruppe „fremde Fehlermap“ erzielte den größeren Wissenszuwachs, wobei der Unterschied zur Gruppe „eigene Fehlermap“ nicht signifikant war ( $F = 2.353$ ;  $df = 1$ ;  $p = 0.132$ ;  $d = 0.57$ ). Hinsichtlich des Umgangs mit den Fehlern zeigte sich, dass die Gruppe „fremde Fehlermap“ signifikant mehr Fehler erkannte ( $F = 10.152$ ;  $df = 1$ ;  $p = 0.003$ ;  $d = 0.96$ ) und richtig korrigierte ( $F = 26.061$ ;  $df = 1$ ;  $p = 0.000$ ;  $d = 1.53$ ). Ein signifikanter Einfluss des Umgangs mit den Fehlern auf den Wissenszuwachs konnte jedoch nicht nachgewiesen werden (Indikator „Anteil erkannter Fehler“:  $F = 1.163$ ;  $df = 1$ ;  $p = 0.287$ ,  $d = 0.59$ , Indikator „Anteil richtig korrigierter Fehler“:  $F = 1.098$ ;  $df = 1$ ;  $p = 0.301$ ,  $d = 0.60$ ).

Für die aufgestellte Hypothese lassen sich insgesamt keine signifikanten Effekte nachweisen, die für eine Vorteilhaftigkeit der Korrektur fremder Fehler in einer Expertenmap gegenüber der Korrektur eigener Fehler sprechen. Tendenziell führt die Korrektur der fehlerhaften Expertenmap aber zu einem höheren Wissenszuwachs. Weiterhin zeigt sich, dass den Schülern das Korrigieren der fehlerhaften Expertenmap deutlich besser gelingt als das Korrigieren der eigenen Fehlermap, womit die Ergebnisse von Bartlett (1982) repliziert werden konnten. Ursachen dafür gilt es in Folgeuntersuchungen zu ergründen.

Bartlett, E. J. (1982). Learning to revise: Some component processes. In M. Nystrand (Ed.) What writers know: The language, process and structure of written discourse (pp. 345- 364). New York: Academic Press.

KC, D., Staats, B. R. & Gino, F. (2012). Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. Harvard Business School Working Paper, 12-065.

Oser, F. & Spychiger, M. (2005). Lernen ist schmerzhaft. Weinheim: Beltz Verlag



## Wie gelingt das Lernen aus eigenen Fehlern? Eine Interventionsstudie im Mathematikunterricht

In der Lehr-Lern-Forschung gibt es Hinweise darauf, dass das Reflektieren des Rationalen, das eigenen Fehlern zu Grunde liegt, das Lernen fördern kann (etwa Ben-Zeev, 1998; VanLehn, 1999). Ein Fehler mit Rational liegt etwa vor, wenn Strategien, die bis dato für das Lösen bestimmter Probleme hinreichend geeignet waren, auf neue Probleme angewendet werden, welche jedoch anderen Strategien bedürfen. Es gibt Hinweise darauf, wie Instruktionen ausgestaltet sein können, damit sie auf das Reflektieren des Rationalen eigener Fehler abzielen (Guldiman & Zutavern, 1999). Bisherige Studien konzentrierten sich vornehmlich auf das Reflektieren fremder Fehler (Große & Renkl, 2007; Siegler, 2002) oder berücksichtigten keine spezifische Anleitung zum Reflektieren (Keith & Frese, 2005). Gleichzeitig fehlt es im Rahmen der Forschung zum Lernen aus Fehlern grundsätzlich an Studien, die im Rahmen ökologisch valider Unterrichtssettings durchgeführt wurden. Die vorliegende Studie geht der Frage nach, ob Schülerinnen und Schüler im Unterricht zur Bruchrechnung aus ihren eigenen Fehlern besser lernen, wenn sie das Rational hinter ihren Fehler reflektieren, als wenn sie lediglich die richtige Lösung reflektieren. Die Bruchrechnung wurde als Lerninhalt gewählt, da Kompetenzen in der Bruchrechnung ein zentraler Prädiktor für Mathematikleistungen darstellen (Siegler et al., 2012). Gleichzeitig gibt es in der Fachdidaktik umfangreiche Kenntnisse über Schülerfehler in der Bruchrechnung und deren Ursachen (etwa Padberg, 2009).

An der quasi-experimentellen Interventionsstudie mit Prä-Post-Follow-up-Design nahmen insgesamt 174 Schülerinnen und Schüler aus neun Klassen der Jahrgangsstufen 7 und 8 teil. Die Intervention wurde im regulären Mathematikunterricht über 11 Unterrichtsstunden durchgeführt. Alle Schülerinnen und Schüler bearbeiteten im Verlaufe der Intervention drei parallele Tests, zu denen sie jeweils direkt im Anschluss ein Feedback zu richtigen und falschen Lösungen erhielten. Auf dieser Grundlage reflektierten die Schülerinnen und Schüler ihre eigenen Fehlern stets in zwei Lernumgebungen, denen sie anfangs randomisiert zugeordnet wurden: In der fehlerorientierten Lernumgebung ( $n = 87$ ) wurden die eigenen Fehler über Prompts reflektiert, die auf das Rational des Fehlers abzielten (z.B. „Erkläre, was Du Dir dabei gedacht hast“). In der lösungsorientierten Lernumgebung ( $n = 87$ ) wurden die eigenen Fehler über korrespondierende (richtige) Lösungsbeispiele mit Prompts reflektiert (z.B. „Wie wird die Aufgabe gelöst? Beschreibe.“). Zu allen drei Messzeitpunkten wurde das prozedurale und konzeptionelle Wissen in der Bruchrechnung erfasst, außerdem wurde ein Test zum negativen Wissen (in Anlehnung an Oser & Spychinger, 2005) eingesetzt, in dem die Schülerinnen und Schüler fremde Fehler identifizieren und erklären sollten. Als Testleistung ging jeweils die prozentuale Lösungsrate der Schülerin bzw. des Schülers ein. Es wurden ANCOVAs mit Messwiederholung berechnet, wobei die Zeit mit drei Stufen (Prä-, Post- und Follow-up-Test) als Messwiederholungsfaktor und die Lernumgebung als Gruppenfaktor eingingen. Als Kontrollvariablen dienten die Mathematiknote, das Geschlecht, das Alter, der Jahrgang, die Klassenstufe und die Schulform.

Es zeigen sich signifikante Interaktionen von Zeit\*Lernumgebung sowohl beim prozeduralen ( $F(2, 334) = 4.97, p < .05, \eta^2 = .029$ ) als auch beim konzeptionellen Wissen ( $F(2, 334) = 3.26, p < .05, \eta^2 = .019$ ). Hingegen zeigen sich keine signifikanten Interaktionen bezüglich des negativen Wissens. Eine Analyse der bedingten Haupteffekte zeigt, dass die Schülerinnen und Schüler in der fehlerorientierten Lernumgebung sowohl im Posttest ( $F(1, 167) = 5.99, p < .05, \eta^2 = .035$ ) als auch im Follow-up-Test ( $F(1, 167) = 4.12, p < .05, \eta^2 = .024$ ) ein signifikant höheres prozedurales Wissen aufweisen als die Schülerinnen und Schüler in der lösungsorientierten Lernumgebung. Bezüglich des konzeptionellen Wissens zeigt sich ein gleich gerichteter Unterschied lediglich zum Follow-up-Test ( $F(1, 167) = 4.02, p < .05, \eta^2 = .023$ ). Theoretische Implikationen sowie Empfehlungen für die Unterrichtspraxis werden im Vortrag diskutiert.

Ben-Zeev, T. (1998). Rational errors and the mathematical mind. *Review of General Psychology*, 2, 366–383.

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction*, 17, 612–634.

Guldiman, T., & Zutavern, M. (1999). „Das passiert uns nicht noch einmal!“ – Schülerinnen und Schüler lernen gemeinsam den bewußten Umgang mit Fehlern. In W. Althof & F. Oser (Hrsg.), *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern: Beiträge und Nachträge zu einem Interdisziplinären Symposium aus Anlaß des 60. Geburtstags von Fritz Oser* (S. 233-258). Opladen: Leske + Budrich.

Keith, N., & Frese, M. (2005). Self-Regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology*, 90, 677–691.

Oser, F., & Spychinger, M. (2005). *Lernen ist schmerzhaft: Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Weinheim: Beltz.

Padberg, F. (2009). *Didaktik der Bruchrechnung. Für Lehrerbildung und Lehrerfortbildung* (4., erw., stark überarb.). Heidelberg: Spektrum.

Siegler, R. S. (2002). Microgenetic studies of self explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment. Transition processes in development and learning* (pp. 31-58). Cambridge, New York: Cambridge University Press.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23, 691–697.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions. Learning, development, and conceptual change*. Cambridge, Mass: MIT Press.



## **Fehler-Finde-und-Korrektur-Aufgaben zum Bruchrechnen: Lernen Schüler die falschen Lösungen?**

Aus der behavioristischen Tradition heraus sollen Fehler im Lernprozess möglichst vermieden oder bei ihrem Auftreten bestraft oder ignoriert werden, damit sie sich nicht wiederholen (Miller, 1983). Demgegenüber stehen neue Ansätze (z.B. Theorie des negativen Wissens von Oser & Spychiger, 2005) und empirische Erkenntnisse (z.B. Adams et al., 2013), nach denen dem Einbeziehen von Fehlern positive Effekte für den Lernprozess zugesprochen werden. Für den Matheunterricht wurden dafür Aufgabenformen konstruiert, in denen typische Fehler gefunden, erklärt und korrigiert werden sollen (z.B. Melis, 2004, Große & Renkl, 2007, Eichelmann, Narciss & Schnaubert, 2013). Es stellt sich die Frage, inwieweit sich die traditionell, kritische Sichtweise des Behaviorismus auf Fehler bei der Einschätzung solcher Aufgaben durch Lehrer und Schüler widerspiegelt. Zur Einschätzung durch Lehrer gibt es bereits erste Befunde. Tsamir & Tirosh (2005) befragten Lehrer, ob und wie sie fehlerhafte Aufgaben im Unterricht als Lerngelegenheit einsetzen würden: Während einige der Lehrer den Einbezug von Fehlern in den Unterricht uneingeschränkt befürworteten, gab es auch Lehrer, die fehlerhafte Aufgaben als verwirrend und kompliziert für Schüler einschätzten oder sie nur bei fortgeschrittenen Schülern einsetzen würden. Die Sichtweise der Schüler und die tatsächliche Auswirkung auf die Häufigkeit von typischen Fehlern wurden bisher noch nicht empirisch untersucht.

Folgende Fragestellung soll daher im Mittelpunkt stehen: Sehen Lehrer und Schüler Fehler-Finde-und-Korrektur als eine Gefahr oder als Potential für den Lernprozess? Befürchten sie insbesondere, dass die in den Aufgaben präsentierten Fehler von den Schülern übernommen werden oder sehen sie eher die Schutzfunktion des „negativen Wissens“? An die Einschätzung der Schüler und Lehrer anknüpfend sollen zudem empirische Ergebnisse zur Häufigkeit typischer Fehler präsentiert werden: Machen Schüler nach der Bearbeitung von Fehler-Finde-und-Korrektur-Aufgaben mehr typische Fehler als Schüler, die normale Bruchrechenaufgaben bearbeitet haben?

Die Daten zu den Schülern wurden im Rahmen einer Studie zu motivationalen und kognitiven Effekten von Fehler-Finde-und-Korrektur-Aufgaben erhoben (Eichelmann, Schnaubert, & Narciss, 2013). 59 Mädchen und 51 Jungen der 6.-8. Klasse im Alter von 11-15 Jahren (MW=12.1) nahmen teil. Dabei wurde die Anzahl typischer Fehler vor und nach dem Bearbeiten von Fehler-Finde-und-Korrektur-Aufgaben und normalen Rechenaufgaben zum Bruchrechnen erhoben. Zudem wurden die Schüler befragt, inwieweit sie glauben, dass die Fehler aus den Fehler-Finde-und-Korrektur-Aufgaben übernommen werden oder zum Aufbau negativen Wissens dienen. Die Einschätzung erfolgte auf einer vierstufigen Ratingskala (von „trifft zu“ bis „trifft nicht zu“).

Die Einschätzung der Lehrer wurde im Rahmen einer Online-Befragung zu Fehler-Finde-und-Korrektur-Aufgabenerhoben. An der Befragung nahmen 150 Mathematik-Lehrer teil (68.9% weiblich, 31.1% männlich). Analog zu den Schülern wurden die Lehrer befragt, inwiefern sie glauben, dass die Schüler Fehler aus den Fehler-Finde-und-Korrektur-Aufgaben übernehmen oder anhand der Aufgaben lernen können, die Fehler zu vermeiden.

Die Einschätzung der Lehrer fällt insgesamt positiv aus. Die Befürchtung, dass die Schüler die Fehler aus den Aufgaben übernehmen, teilen nur 16.7% der befragten Lehrer, der größte Teil der Lehrer (83.3%) stimmt dem (eher) nicht zu. Dagegen stimmen 69.4% (eher) der Aussage zu, dass die Schüler durch die Aufgaben sehen, welche Fehler passieren können und sie dann selbst vermeiden. Auf Seiten der Schüler ist die Skepsis jedoch größer: Obwohl der überwiegende Teil der Schüler (eher) nicht glaubt, dass die Fehler aus den Aufgaben übernommen werden, sehen doch immerhin 33.6% diese Gefahr. Trotzdem sehen auch noch 65.5% der Schüler die Möglichkeit, dass durch die Aufgaben Fehler in Zukunft vermieden werden können. Bei der Häufigkeit der typischen Fehler zeigen sich keine Unterschiede zwischen Schülern, die normale Bruchrechenaufgaben und Schülern, die Fehler-Finde-und-Korrektur-Aufgaben bearbeitet haben ( $T(108)=-.79, p=.43$ ).

Die aus dem behavioristischen Ansatz abgeleiteten Bedenken gegen den Einbezug von Fehlern lassen sich also empirisch weder bei der Fehlerhäufigkeit noch bei der Sichtweise der Lehrer wiederfinden.

Adams, D.M., McLaren, B.M., Mayer, R.E., Gogvadze, G. & Isotani, S. (2013). Erroneous examples as desirable difficulty. *Artificial Intelligence in Education*, Volume 7926, 803-806.

Eichelmann, A., Narciss, S. & Schnaubert, L. (2013). Learning from errors through tasks-with-typical-errors. Paper presented at the 15th Biennial Conference EARLI in Munich, 27-31 August 2013.

Große, C. S. & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning & Instruction*, 17, 612-634.

Melis, E. (2004). Errors as a source of learning in mathematics. Presented at *Cognition and Exploratory Learning in the Digital Age (CELDA)*.

Miller, P.H., 1983. *Theories of developmental psychology*. Freeman, San Francisco, USA.

Oser, F. & Spychiger, M. (2005). Lernen ist schmerzhaft. Zur Theorie des negativen Wissens und zur Praxis der Fehlerkultur. Weinheim: Beltz.

Tsamir, P., & Tirosh, D. (2005). In-service elementary mathematics teachers' views of errors in the classroom. *Focus on Learning Problems in Mathematics*, 27 (3), 30-42.



## ***Adaptiver individueller Umgang mit Fehlern: Eine Frage des Unterrichtsfaches?***

Der adaptive Umgang mit Fehlern stellt eine Grundvoraussetzung für erfolgreiches Lernen aus Fehlern dar. Als adaptiv gelten dabei einerseits die Aufrechterhaltung günstiger affektiv-motivationaler Bedingungen, andererseits Lernhandlungen, die auf eine tiefere Auseinandersetzung mit der Fehlerursache, damit verbundenen fehlerhaften Konzepten und deren Korrektur abzielen (Dresel, Schober, Ziegler, Grassinger, & Steuer, 2013). Darüber hinaus ist anzunehmen, dass „fehler-freundliche“ Überzeugungen adaptive Fehlerreaktionen günstig beeinflussen (vgl. Keith & Frese, 2005). Basierend darauf werden drei Komponenten einer adaptiven Haltung gegenüber Fehlern unterschieden: die kognitive (z.B. Fehler als Lernchance auffassen), affektiv-motivationale (z.B. Aufrechterhaltung aktivierender Emotionen) und handlungsadaptive Komponente (z.B. Analyse der Fehlerursache). Bisher wurde der Umgang mit Fehlern überwiegend im Fach Mathematik untersucht (vgl. Oser & Spychiger, 2005). Einige wenige Studien beziehen sich auch auf Deutsch, Englisch (z.B. DESI-Konsortium, 2008) oder Wirtschaftsunterricht (vgl. Wuttke & Seifried, 2012). Unklar (für praktische Implikationen jedoch essentiell) ist, ob der individuelle Umgang mit Fehlern primär als domänenübergreifendes Konstrukt angesehen, oder – wie eine Reihe anderer motivationaler Konstrukte (z.B. Bong, 2001; Möller & Köller, 2001) – domänenspezifisch betrachtet werden muss. Dieser Frage wurde in vorliegender Studie anhand von Selbstberichten von 600 Schüler(inne)n aus 40 Klassen der 6.-7. Jahrgangsstufe für die Unterrichtsfächer Mathematik, Deutsch und Englisch nachgegangen. Analysen auf korrelativer Basis ergaben Hinweise auf Unterschiede zwischen Fächern, während sich anhand von Mittelwertvergleichen kaum signifikante Unterschiede zeigen ließen. Zusammengefasst kann für die untersuchte Altersgruppe eine gewisse Domänenspezifität für den individuellen Umgang mit Fehlern angenommen werden. Eine domänenspezifische Betrachtung erscheint demnach empfehlenswert.

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, 93(1), 23-34.

DESI-Konsortium (Hrsg.). (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.

Dresel, M., Schober, B., Ziegler, A., Grassinger, R. & Steuer, G. (2013). Affektiv-motivational adaptive und handlungsadaptive Reaktionen auf Fehler im Lernprozess. *Zeitschrift für Pädagogische Psychologie*, 27(4), 255–271.

Keith, N. & Frese, M. (2005). Self-Regulation in Error Management Training: Emotion Control and Metacognition as Mediators of Performance Effects. *Journal of Applied Psychology*, 90(4), 677-691.

Möller, J. & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology*, 93(4), 826-835.

Oser, F. & Spychiger, M. (2005). *Lernen ist schmerzhaft: Zur Theorie des Negativen Wissens und Praxis der Fehlerkultur*. Weinheim: Beltz.

Wuttke, E. & Seifried, J. (Hrsg.). (2012). *Learning from errors at school and at work*. Opladen: Verlag Barbara Budrich.



## **Unterrichtsprozesse und Qualitätsmerkmale: Erfassung, Perspektiven auf Unterricht und Wirkungen**

Chair(s): Klieme Eckhard (DIPF / Goethe-Universität Frankfurt / IDeA-Forschungszentrum)

Diskutant/in: Kathrin Krammer (Pädagogische Hochschule Luzern (CH))

Ein zentrales Ziel der Unterrichtsforschung ist es, Merkmale von Unterricht zu identifizieren, die mit günstigen motivationalen und kognitiven Entwicklungen von Schülerinnen und Schülern in Beziehung stehen. Diese Forschung begann mit dem Prozess-Produkt-Paradigma, in dem Prozess-Merkmale des Unterrichts (z.B. verbale Interaktionsmuster) mit Schüler-Outcomes als Produkte in Beziehung gesetzt wurden (z.B. Dunkin & Biddle, 1974). Anschließend wurden kognitiv-konstruktivistische und sozio-konstruktivistische Modelle zum Lehren und Lernen in dieses Paradigma integriert und Lernen wurde als ein aktiver und konstruktiver Prozess angesehen, der im Miteinander mediiert wird (z. B. Bransford, Brown, & Cocking, 2000).

Das Symposium widmet sich der auf diesen Theorien aufbauenden empirischen Analyse von Unterrichtsprozessen und -merkmalen. Ziel dieses vielschichtigen und komplexen Aufgabenfeldes ist es, aus unterschiedlichen Perspektiven die Erfassung von Unterrichtsprozessen und Qualitätsmerkmalen sowie deren Wirkungen auf motivationale und kognitive Merkmale von Schülerinnen und Schülern eingehender zu prüfen. Im ersten Beitrag werden auf Basis von Videographierungen Lehrer-Schüler-Interaktionen im Physikunterricht der Sekundarstufe dahingehend analysiert, welche Differenzierungsmaßnahmen bei spezifischen Schülerprofilen zu günstigen Entwicklungen führen. Lehrer-Schüler-Interaktionen, die mit positiven motivationalen und kognitiven Entwicklungen in Beziehung stehen, wurden in den letzten Jahren mit Hilfe eines Modells zu drei Basisdimensionen von Unterrichtsqualität beschrieben: kognitive Aktivierung, unterstützendes Klima und strukturierte Klassenführung. Diese Basisdimensionen konnten in verschiedenen Studien unabhängig voneinander nachgewiesen werden (z.B. Baumert et al., 2010; Klieme, Pauli, & Reusser, 2009; Pianta & Hamre, 2009). Dennoch bedarf es hierzu weiterer Forschung. Der zweite Beitrag untersucht, auf welche Weise die drei Unterrichtsqualitätsdimensionen mit dem Einsatz metakognitiver Strategien von Grundschulkindern verknüpft sind und inwiefern dieser Zusammenhang durch die intrinsische Motivation der Schülerinnen und Schüler mediiert wird. Der dritte Beitrag widmet sich der Erfassung von Unterrichtsqualitätsmerkmalen aus drei Perspektiven: Schülerinnen und Schülern, Lehrkräften und externen Beobachtern. Hierbei wird untersucht, inwieweit diese Perspektiven innerhalb einer jeweiligen Basisdimension zusammenhängen und welche Perspektive auf welcher Dimension schulische Leistungen vorherzusagen vermögen. Während der bisherige Forschungsschwerpunkt in der mathematisch-naturwissenschaftlichen Domäne liegt, prüft der letzte Beitrag, ob sich die Basisdimensionen fachunabhängig erfassen lassen und legt dabei seinen Fokus auf die Fächer Deutsch und Englisch. Zusammen genommen ergänzen sich die Beiträge, indem sie Unterrichtsprozesse und -qualitätsmerkmale in unterschiedlichen Fachdisziplinen (Naturwissenschaften, Deutsch, Englisch), erfasst über unterschiedliche Perspektiven (Schülerinnen und Schüler, Lehrkräfte, Beobachter), eingehender betrachten und Wirkungen auf motivationale und kognitive Lernergebnisse von Schülerinnen und Schülern untersuchen. Das Symposium verbindet somit fachdidaktische, erziehungswissenschaftliche und pädagogisch-psychologische Perspektiven und leistet somit einen wichtigen Beitrag zur interdisziplinär ausgerichteten quantitativen Unterrichtsforschung.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180.

Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How people learn*. Washington, DC: National Academy Press.

Dunkin, M. & Biddle, B. (1974). *The study of teaching*. New York: Holt, Rhinehart & Winston.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster: Waxmann.

Pianta, R. C. & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109–119.

Montag,  
03.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G150



## ***Das Zusammenspiel von Mikroprozessen und Schülervoraussetzungen im Physikunterricht: Ergebnisse einer Videostudie***

### 1. Theoretischer Hintergrund

Individuelle Schülervoraussetzungen wie Vorwissen, Interesse und Selbstkonzept spielen eine wichtige Rolle für das erfolgreiche Lernen (Shuell, 1996). Die Lernvoraussetzungen einzuschätzen, aber auch das Einsetzen von Differenzierungsmaßnahmen unter Berücksichtigung dieser Voraussetzungen sind wichtige Merkmale eines qualifizierten Unterrichts (Becker, 2007), denn jeder Schüler benötigt für sich spezifische Unterstützung bezüglich seiner Fähigkeiten (Shavelson & Seidel, 2006). Unterrichtsprozesse sollen hierbei förderlich wirken. Insbesondere eine aktive Beteiligung an Interaktionen wird als positiv für den Erfolg und die Lernentwicklung eingestuft (Mercer, 1996; Webb, 2009). Auch ein offener Umgang mit Fehlern wird als bedeutend für das Lernen und die Schülerentwicklung angesehen (Santagata, 2005).

Die ersten Analysen der IPN-Videodaten haben fünf verschiedene Schülerprofile im Physikunterricht ergeben: stark, überschätzend, unterschätzend, uninteressiert und schwach (Seidel, 2006). Diese Profile unterscheiden sich in ihren kognitiven Fähigkeiten und der Ausprägung im motivational-affektiven Bereich. Der vorliegende Beitrag betrachtet eine spezielle Schülergruppe, die unterschätzenden Lerner, welche sich durch hohe kognitive Fähigkeiten und ein niedriges Selbstkonzept auszeichnen. Durch Re-Analysen der Daten wird untersucht, ob und inwiefern Lehrer in Unterrichtsprozessen des Physikunterrichts auf diese Schülergruppe besonders eingehen und diese sich so im motivational-affektiven Bereich entwickeln und ihre Leistungen entsprechend ihrer hohen kognitiven Fähigkeiten besser ausschöpfen können.

Durch Untersuchung der Differenzierungsmaßnahmen der Lehrpersonen im Hinblick auf diese spezielle Schülergruppe wird im vorliegenden Artikel ein wichtiger Aspekt von Unterrichtsqualität im naturwissenschaftlichen Unterricht betrachtet.

### 2. Fragestellungen

Im vorliegenden Beitrag wurden die Mikroprozesse zwischen Lehrpersonen und unterschätzenden Schülern im Physikunterricht untersucht. Dabei wurden folgende Fragestellungen betrachtet:

(1) Gestalten Lehrer Mikroprozesse im Unterricht gezielt als Unterstützung für unterschätzende Schülerinnen und Schüler?

Angenommen wurde, dass ein Großteil der Lehrpersonen nicht gezielt auf unterschätzende Schülerinnen und Schüler eingeht.

(2) Welche der untersuchten Unterrichtsprozesse und -prozessmerkmale stehen in Zusammenhang mit einer positiven Entwicklung bei dieser Schülergruppe?

Es wurde angenommen, dass eine aktive Einbindung der Schülerinnen bzw. Schüler in die Interaktionen, sowie bestimmte Unterrichtsprozessmerkmale, wie etwa eine positive Fehlerkultur, in Verbindung mit positiven Entwicklungen der Schüler steht.

### 3. Methode

Stichprobe: Der vorliegende Beitrag basiert auf Daten einer Re-analyse der IPN Videostudie (Seidel et al., 2003). Die Stichprobe umfasst N = 82 Gymnasialklassen der 9. Jahrgangsstufe mit insgesamt 954 Schülerinnen und Schülern (44% Mädchen, 56% Jungen) aus Deutschland und der Schweiz.

Design: Die Schülerinnen und Schüler wurden bezüglich ihrer kognitiven Fähigkeiten (KFT) zu Beginn des Schuljahres, zu ihrem Vorwissen (16 TIMSS-Items) und Interesse am Fach sowie zu ihrem Selbstkonzept zu Beginn und zum Ende des Schuljahres in einer Unterrichtsstunde (45 Minuten) getestet und befragt. Während des Schuljahrs wurde eine Doppelstunde des Physikunterrichts videografiert.

Statistische Analysen: Zur Bestimmung der Schülerprofile wurden latente Klassenanalysen (LCA) mit Hilfe der Software Mplus durchgeführt. Die Videos wurden mit der Software Videograph kodiert und über Einheiten der Sprecher ausgewertet.

### 4. Ergebnisse

Während sich die Anzahl der Schülerinnen und Schüler in einem als unterschätzend klassifizierten Profil über das Schuljahr verringert, ist die Entwicklung der unterschätzenden Schüler insgesamt nicht als positiv zu bewerten. Schüler im unterschätzenden Profil zu Schuljahresende haben ein noch niedrigeres Selbstkonzept, als jene zu Schuljahresbeginn.

Bezüglich der Interaktionen im Klassenzimmer bestätigen die Analysen die Annahme, dass in einem Großteil der Klassen keine besondere Förderung der unterschätzenden Schüler durch Anregungen oder Rückmeldungen zu beobachten ist. Positive Fehlerkultur, zum Nachdenken anregende Lehrerfragen sowie vermehrtes Feedback stehen jedoch in Zusammenhang mit einer positiven Entwicklung. Die Studie legt nahe, dass Lehrer eventuell mehr Informationen über ihre Schüler benötigen, um Unterrichtsprozesse individuell unterstützend gestalten zu können.

Becker, G. E. (2007). Unterricht planen (Teil 1). Weinheim: Beltz.

Mercer, N. (1996). The Quality of Talk in Children's collaborative activity in the classroom. *Learning and Instruction*, 6, 359–377.

Santagata, R. (2005). Practices and beliefs in mistake handling activities: A video study of Italian and US mathematics lessons. *Teaching and Teacher Education*, 21, 491–508.

Seidel, T. (2006). The role of student characteristics in studying micro teaching-learning environments. *Learning Environments Research*, 9, 253–271.

Shavelson, R. & Seidel, T. (2006). Approaches in measuring learning environments. *Learning Environment Research Journal*, 9, 195–197.

Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. Calfee (Hrsg.), *Handbook of educational psychology*, 726–764. New York: Macmillan.

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79, 1–28.





## Die Bedeutung von Unterrichtsqualität für den Einsatz metakognitiver Strategien

### 1. Theoretischer Hintergrund

Die Wirkung der drei Dimensionen der Unterrichtsqualität (kognitive Aktivierung, unterstützendes Klima und Classroom Management; Klieme, Pauli, & Reusser, 2009) auf die Leistung der Schülerinnen und Schüler wird theoretisch durch Annahmen des Angebots-Nutzungs-Modells (Helmke, 2010) erklärt. Es wird davon ausgegangen, dass eine hohe Unterrichtsqualität positiv auf die Lernaktivitäten der Schülerinnen und Schüler wirkt, die dann zu besseren Leistungen führen (Klieme et al., 2009). Eine solche Lernaktivität ist der Einsatz metakognitiver Strategien (z.B. Zimmerman & Schunk, 2001). Unter metakognitiven Strategien werden Aktivitäten von Lernenden verstanden, die dazu dienen den Lernprozess zu planen, zu überwachen und zu evaluieren (Hasselhorn, 2006).

Für ein unterstützendes Klima wird angenommen, dass es sich insbesondere indirekt auf Leistungen auswirkt, indem es die Lernmotivation steigert (Klieme et al., 2009). Dazu konnte gezeigt werden, dass eine hohe intrinsische Lernmotivation mit einem vermehrten Einsatz metakognitiver Strategien einhergeht (z.B. Pintrich, 1999).

Diese Annahmen zum Zusammenhang zwischen Lernaktivitäten der Schülerinnen und Schüler und den drei Dimensionen der Unterrichtsqualität wurden jedoch bisher kaum überprüft.

### 2. Fragestellung

Der vorliegende Beitrag beschäftigt sich mit der Frage, ob die drei Dimensionen der Unterrichtsqualität mit dem Einsatz metakognitiver Strategien durch die Lernenden verknüpft sind. Wir erwarten, dass alle drei Dimensionen der Unterrichtsqualität positiv mit dem Strategieeinsatz zusammenhängen. Weiterhin erwarten wir, dass der Zusammenhang zwischen einem unterstützenden Klima und dem Strategieeinsatz durch die intrinsische Motivation der Schülerinnen und Schüler mediiert wird.

### 3. Methode

Die Analysen beziehen sich auf 36 dritte Klassen mit 702 Schülerinnen und Schülern. Mittels eines Fragebogens wurden die Kinder zu ihrem Einsatz metakognitiver Strategien jeweils vor und während einer vorgegebenen Unterrichtseinheit zu einem naturwissenschaftlichen Thema (10 Items; Cronbachs  $\alpha > .82$ ) befragt. Weiterhin wurde vor der Einheit die naturwissenschaftliche Kompetenz (12 Items; EAP/PV Reliabilität = .70) und nach der Einheit die intrinsische Motivation der Kinder während des Unterrichts (6 Items; Cronbachs  $\alpha = .92$ ) erhoben. Die Unterrichtsqualität wurde mittels eines hochinferenten Videoratings (ein Item pro Dimension) während der Unterrichtseinheit erfasst (ICC  $> .67$ ). Zur Überprüfung der ersten Hypothese wurden Mehrebenenregressionsanalysen berechnet, während die zweite mittels eines Mehrebenenmediationsmodells geprüft wurde. In allen Modellen wurde für den Strategieeinsatz vor der Unterrichtseinheit und die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler kontrolliert.

### 4. Ergebnisse und Diskussion

Die Ergebnisse unterstützen die Annahme, dass bei einer hohen Unterrichtsqualität mehr Strategien im Unterricht eingesetzt werden. Es konnte gezeigt werden, dass jede der drei Dimensionen für sich ein signifikanter Prädiktor für den Einsatz metakognitiver Strategien ist (kognitive Aktivierung:  $\beta = .12$ ,  $SE = .05$ ,  $p < .05$ ; unterstützendes Klima:  $\beta = .15$ ,  $SE = .06$ ,  $p < .01$ ; Classroom Management:  $\beta = .11$ ,  $SE = .05$ ,  $p < .05$ ).

Besonders für die kognitive Aktivierung kann angenommen werden, dass dieser Zusammenhang zustande kommt, da sie Aspekte von Strategieinstruktion umfasst (Pressley & Harris, 2009). Weiterhin kann Classroom Management als Gestaltungsmerkmal der Lernumgebung verstanden werden, das die Bedingungen für den Strategieeinsatz optimiert (de Corte, Verschaffel, & Masui, 2004).

Bezüglich der zweiten Annahme weisen die Ergebnisse des Mehrebenenmediationsmodell auf eine vollständige Mediation des Zusammenhangs zwischen einem unterstützendem Klima und dem Strategieeinsatz durch die Motivation der Schülerinnen und Schüler ( $\beta_{\text{indirekt}} = .15$ ,  $SE = .04$ ,  $p < .01$ ;  $\beta_{\text{direkt}} = .03$ ,  $SE = .05$ ,  $p > .05$ ) hin. Dieses Ergebnis unterstützt die Annahme, dass die Motivation der Schülerinnen und Schüler ein Wirkmechanismus des unterstützenden Klimas ist.

Dieser Mechanismus könnte auch für die Zusammenhänge mit den anderen Dimensionen von Bedeutung sein, da bei mangelnder Lernmotivation davon ausgegangen werden kann, dass auch eine hohe kognitive Aktivierung und ein gutes Classroom Management nicht zum gewünschten Lernerfolg führen (Boekaerts, 1995).

Boekaerts, M. (1995). Self-regulated learning: Bridging the gap between metacognitive and metamotivation theories. *Educational Psychologist*, 30, 195–200.

de Corte, E., Verschaffel, L. & Masui, C. (2004). The CLIA-model: A framework for designing powerful learning environments for thinking and problem solving. *European Journal of Psychology of Education*, 19, 365–384.

Hasselhorn, M. (2006). Metakognition. In D. H. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie (Schlüsselbegriffe, 3. Aufl., S. 480–485)*. Weinheim: Beltz.

Helmke, A. (2010). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (Unterricht verbessern - Schule entwickeln, 3. Aufl.). Seelze: Klett.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom (S. 137–160)*. Münster: Waxmann.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31, 459–470.

Pressley, M. & Harris, K. R. (2009). Cognitive strategies instruction: From basic research to classroom instruction. *Journal of Education*, 189, 77–94.

Zimmerman, B. J. & Schunk, D. H. (Hrsg.). (2001). *Self-regulated learning and academic achievement. Theoretical perspectives*. Mahwah, NJ: Erlbaum.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G150



Benjamin Fauth (Goethe-Universität Frankfurt), Svenja Rieser (Goethe-Universität Frankfurt), Jasmin Decristan (DIPF), Eckhard Klieme (DIPF), Gerhard Büttner (Goethe-Universität Frankfurt)

## **Urteile von Schülerinnen und Schülern, Lehrpersonen und externen Beobachtern zur Unterrichtsqualität: Können die Perspektiven verbunden werden?**

1. Theoretischer Hintergrund: Um in empirischen Studien einen möglichst umfassenden Einblick in Unterrichtsprozesse zu erlangen, ist es wünschenswert, Urteile zur Unterrichtsqualität aus unterschiedlichen Perspektiven zu berücksichtigen (Clausen, 2002). Die Verbindung dieser Perspektiven zu einem gemeinsamen Indikator macht Sinn, wenn davon ausgegangen werden kann, dass aus verschiedenen Perspektiven dasselbe latente Konstrukt erfasst wird (Kunter et al., 2013). Gerade Unterrichtsqualität ist jedoch in früheren Studien als perspektivenspezifisches Konstrukt beschrieben worden (Kunter & Baumert, 2006). Urteile zur Unterrichtsqualität aus unterschiedlichen Perspektiven korrelieren nur gering miteinander (Clausen, 2002).

In dem vorliegenden Beitrag wird davon ausgegangen, dass Zusammenhänge zwischen Perspektiven auch davon abhängen, welcher Aspekt von Unterrichtsqualität im Fokus steht. Es werden Urteile zu zwei Basisdimensionen von Unterrichtsqualität (Klieme, Pauli, & Reusser, 2009) betrachtet, die in früheren Studien besonders deutlich mit dem Lernerfolg von Schülerinnen und Schülern zusammenhängen: strukturierte Klassenführung und kognitive Aktivierung (Lipowsky et al., 2009).

2. Fragestellungen: (1) Welche Zusammenhänge gibt es zwischen den Urteilen von Schülerinnen und Schülern, Lehrpersonen und Beobachtern zu den Dimensionen strukturierte Klassenführung und kognitive Aktivierung?

(2) Wessen Urteile zu welcher Dimension sagen Schülerleistungen vorher?

(3) Lassen sich latente Faktoren identifizieren, die Urteile aus allen drei Perspektiven integrieren?

3. Methode: Die Studie basiert auf Daten von 1070 Schülerinnen und Schülern aus 54 Klassen. In einem längsschnittlichen Design wurden zwei Unterrichtseinheiten in den Sachunterricht der dritten Jahrgangsstufe implementiert. Die Urteile von Schülerinnen und Schülern, Lehrpersonen und Beobachtern beziehen sich auf die Qualität des Unterrichts während dieser Einheiten. Der Schülerfragebogen umfasste fünf Items zur strukturierten Klassenführung und sieben Items zur kognitiven Aktivierung (Cronbachs  $\alpha = .97/.91$ , ICC2 =  $.86/.73$ ). Parallel dazu wurden die Konstrukte in einem Lehrerfragebogen erfasst (vier bzw. fünf Items, Cronbachs  $\alpha = .67/.86$ ). Die externen Beobachter gaben ihre Urteile auf zwei hoch-inferenten Ratingitems zur Beurteilung von strukturierter Klassenführung und kognitiver Aktivierung ab. Beide Items konnten mit hinreichender Beurteilerübereinstimmung eingeschätzt werden (ICC =  $.81/.77$ ). Das Wissen der Schülerinnen und Schüler zum Thema der Unterrichtseinheiten wurde mittels standardisierter Leistungstests erfasst. Der Prä-Test umfasste 16 Items, der Post-Test nach der zweiten Unterrichtseinheit 13 Items. Als Kovariaten wurden neben dem Prä-Test die kognitiven Grundfähigkeiten der Schülerinnen und Schüler (CFT 20-R) und die naturwissenschaftliche Kompetenz (adaptiert nach TIMSS 2007) kontrolliert. Für die Vorhersage von Schülerleistungen wurden Mehrebenen-Regressionsanalysen in Mplus 7 berechnet. Für die Analysen zu Fragestellung 3 wurden Strukturgleichungsmodelle auf der Klassenebene spezifiziert.

4. Ergebnisse: (1) Es zeigen sich deutliche Zusammenhänge zwischen den Urteilen aus allen drei Perspektiven zur strukturierten Klassenführung ( $r = .45$  bis  $.54$ ), jedoch keine signifikanten Zusammenhänge zwischen den Urteilen zur kognitiven Aktivierung.

(2) Urteile zur strukturierten Klassenführung aus allen drei Perspektiven haben prädiktive Kraft für Lernerfolg. Im Bereich kognitive Aktivierung haben hingegen nur die Urteile externer Beobachter bedeutsame Erklärungskraft.

(3) Aufgrund der Übereinstimmungen zwischen den Perspektiven im Bereich strukturierte Klassenführung lässt sich hier ein latenter Faktor mit den einzelnen Urteilen als Indikatoren bilden, der einen sehr guten Modell-Fit zeigt und ebenfalls prädiktiv für Schülerleistung ist.

5. Diskussion: Die Ergebnisse zeigen, dass die Vorhersagekraft der Unterrichtsqualitätsurteile sowohl von der gewählten Perspektive als auch von der betrachteten Basisdimension abhängt. Während die strukturierte Klassenführung für alle Urteilenden recht gut zu beobachten ist, erfordert die Einschätzung von kognitiver Aktivierung mehr Inferenzen von den Urteilenden. Dies kann die unterschiedlichen Zusammenhänge zwischen den Perspektiven erklären. In dem kombinierten Faktor zur Klassenführung sind die unterschiedlichen Erfahrungen der Urteilenden kombiniert, was ihn zu einem wertvollen Indikator zur Beschreibung von Unterricht macht.

Clausen, M. (2002): Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras study: investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), The power of video studies in investigating teaching and learning in the classroom (S. 137–160). Münster, Germany: Waxmann.

Kunter, M. & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. Learning Environments Research, 9, 231–251.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. Journal of Educational Psychology, 105, 805–820.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E. & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean theorem. Learning and Instruction, 19, 527–537.

## **Deutsch = Englisch? Zur Fachspezifität von Unterrichtsqualität**

Unterrichtsqualität gilt als eine der bedeutsamsten Variablen zur Beeinflussung der Leistung und Motivation von Schülern. Zur Beschreibung von Unterrichtsqualität wird seit einigen Jahren oftmals auf das Modell der Basisdimensionen von Unterrichtsqualität zurückgegriffen: Im Rahmen des Modells werden drei Bereiche unterschieden – Klassenführung, unterstützendes Unterrichtsklima und kognitive Aktivierung (s. Klieme, Pauli, & Reusser, 2009; Kunter & Baumert, 2006; Lipowsky, Rakoczy, Pauli, Drollinger-Vetter, Klieme, & Reusser, 2009). Es wird angenommen, dass diese Dimensionen über Schulformen, Schulstufen und Unterrichtsfächer hinweg gültig sind (z. B. Klieme, Schümer, & Knoll, 2001). Insbesondere zur Frage danach, inwiefern es sich bei den Basisdimensionen tatsächlich um fächerübergreifende Konstrukte handelt, gibt es bislang kaum empirische Untersuchungen. Dieser Forschungsfrage geht der vorliegende Vortrag in Bezug auf zwei der Basisdimensionen (Klassenführung und unterstützendes Unterrichtsklima) nach.

Zur Beantwortung der Frage wurden Daten der DESI-Studie (Klieme et al., 2008) reanalysiert. Es wurden diejenigen 25 Klassen mit ihren 548 Schülern ausgewählt, die in den Fächern Deutsch und Englisch von derselben Lehrkraft unterrichtet wurden. Zur Messung der Dimensionen Klassenführung und unterstützendes Unterrichtsklima wurden auf Klassenebene aggregierte Schülereinschätzungen eingesetzt. Für die Dimension Klassenführung wurden zwei Skalen mit insgesamt 9 Items eingesetzt. Die Dimension unterstützendes Unterrichtsklima wurde über drei Skalen mit insgesamt 12 Items erfasst. Ausgewertet wurden die Daten mit der Generalisierbarkeitstheorie (Shavelson & Webb, 1991). Zur Identifikation von Fachunterschieden wurde ein vollständig gekreuztes Zwei-Facetten-Design (Lehrer  $\times$  Subskala  $\times$  Fach) eingesetzt. Die Berechnungen wurden für die beiden Dimensionen getrennt durchgeführt. Es zeigt sich, dass sich die beiden untersuchten Dimensionen darin unterscheiden, wie stark Fachunterschiede eine Rolle bei der Einschätzung von Unterrichtsqualität durch Schüler spielen: Für die Dimension unterstützendes Unterrichtsklima finden sich nur geringe Unterschiede zwischen den Fächern Deutsch und Englisch, während sich für die Dimension Klassenführung substantielle Unterschiede zwischen den beiden Fächern feststellen lassen. Die Ergebnisse der durchgeführten Analysen werden im Hinblick auf ihre Implikationen für die Konzeptualisierung der drei Basisdimensionen als fachunabhängige Konstrukte diskutiert.

- Klieme, E., Eichler, W., Helmke, A., Lehmann R. H., Nold, G., Rolff, H.-G. et al. (Hrsg.) (2008). Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim: Beltz.
- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German classrooms. In T. Janik & T. Seidel, (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster: Waxmann.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: Aufgabenkultur und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (BMBF) (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–57). München: Medienhaus Biering.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E. & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean theorem. *Learning and Instruction*, 19, 527–537.
- Shavelson, R. & Webb, N. (1991). *Generalizability theory: A primer*. Thousand Oaks, CA: Sage.



## Lehrerziele und die Wahrnehmung effektiver Unterrichtsmethoden

Chair(s): Claudia Rüprich (Martin-Luther-Universität Halle-Wittenberg)

Lehrerziele bezeichnen mentale Repräsentationen von erwünschten berufsspezifischen Endzuständen, die Lehrkräfte noch nicht vollständig erreicht haben, jedoch in der Zukunft verbindlich anstreben. Qualitativ-induktive Studien verweisen auf drei Zielbereiche, die für Lehrkräfte Relevanz besitzen (Hagger & Malmberg, 2011; Rich & Almozlino, 1999). Zum einen werden sozio-emotionale Ziele thematisiert, die z.B. Bestrebungen nach guten Beziehungen zu Schülern, Schülermotivierung und der Berücksichtigung individueller Unterschiede beinhalten, zum anderen werden aufgabenbezogene Ziele wie das Streben nach Fachwissen oder einer adäquaten Planung und Organisation des Unterrichts angesprochen. Der dritte Zielbereich umfasst selbstbezogene Ziele wie das Streben nach Vereinbarkeit von Berufs- und Privatleben sowie der Wunsch, Selbstsicherheit zu erlangen. Weil Ziele das Verhalten steuern und organisieren, wird erwartet, dass Lehrkräfte die verstärkt sozio-emotionale, aufgabenbezogene oder selbstbezogene Ziele verfolgen unterschiedliche Unterrichtsstrategien anwenden und diese in ihrer Effektivität anders bewerten.

Hattie (2009) konnte in seiner Arbeit über die Wirkfaktoren schulischen Lernens aufzeigen, dass Unterrichtsstrategien, in denen der Lehrer eine aktive und führende Rolle einnimmt, eine höhere Effektivität besitzen, wohingegen sich klassische reformpädagogische Konzepte wie offener, ungeführter Unterricht oder schulische Rahmenbedingungen eine deutlich geringere Bedeutung für den Lernerfolg der Schüler aufweisen. Zudem wurde darauf verwiesen, dass sich die subjektiven Überzeugungen bezüglich der Determinanten schulischen Lernens z.T. deutlich von der Forschungslage unterscheiden können.

In der hier vorliegenden Studie soll untersucht werden, ob das Streben nach sozio-emotionalen Zielen mit einer erhöhten Wirksamkeitseinschätzung von offenen Unterrichtskonzepten (schülerzentriertes Lernen, induktives Lernen, entdeckendes Lernen) einhergeht, wohingegen das Streben nach aufgabenbezogenen Zielen mit einer erhöhten Wirksamkeitseinschätzung der lehrergeleiteten Unterrichtsstrategien (direktiver Unterricht, regelmäßige Tests, Hausaufgaben) kovariiert. Vermutet wird zudem, dass das Streben nach selbstbezogenen Zielen mit einer erhöhten Wirksamkeitseinschätzung von schulischen Rahmenbedingungen (Wiederholung der Jahrgangsstufe, Sommerferien) zusammenhängt.

An der Fragebogenstudie nahmen 188 Lehramtsstudierende verschiedener Schultypen und Unterrichtsfächer sowie 115 Gymnasial- und Sekundarschullehrkräfte teil. Die Lehrerziele wurden mit dem Fragebogen zur Erfassung von Lehrerzielen (FELZ, Rüprich & Urhahne, 2013), die subjektiv wahrgenommene Effektivität von Unterrichtsstrategien mit einem an die Arbeit von Köller (2013) angelehnten Hattie-o-Meter erfasst.

Die Ergebnisse zeigen, dass gemäß den Hypothesen das Streben nach sozio-emotionalen Zielen in beiden Stichproben mit einer höher eingeschätzten Effektivität von offenen Unterrichtsstrategien einhergeht, wohingegen das Streben nach aufgabenbezogenen Zielen signifikante Zusammenhänge zur eingeschätzten Effektivität der lehrergeleiteten Unterrichtsmethoden aufwies. Das Verfolgen von selbstbezogenen Zielen ging mit einer zunehmenden Effektivitätseinschätzung der schulischen Rahmenbedingungen einher, in der Lehrerstichprobe zeigten sich darüber hinaus signifikant positive Korrelationen zu den lehrergeleiteten Unterrichtsstrategien. Die Ergebnisse sollen im Hinblick auf die Bedeutung von Zielen und den damit verbundenen Möglichkeiten zur Erreichung einer realistischeren Effektivitätseinschätzung diskutiert werden.

Hagger, H. & Malmberg, L.-E. (2011). Pre-service teachers' goals and future-time extension, concerns, and well-being. *Teaching and Teacher Education*, 27, 598–608.

Hattie, J. A. C. (2009). *Visible learning. A synthesis over 800 meta-analyses relating to achievement*. Abingdon: Routledge.

Köller, O., Möller, J. & Möller, J. (2013). Was wirkt wirklich? Einschätzungen von Determinanten schulischen Lernens. *Schulmanagement Handbuch*, 145. München: Oldenbourg.

Rich, Y. & Almozlino, M. (1999). Educational goal preferences among novice and veteran teachers of sciences and humanities. *Teaching and Teacher Education*, 15, 613–629.

Rüprich, C. & Urhahne, D. (2013). Fragebogen zur Erfassung von Lehrerzielen (FELZ). In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (16. Aufl., S. 575). Bern: Huber.

Montag,

03.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 0.251



## ***Theoretisieren für die Praxis: Förderung der Anwendung wissenschaftlichen Wissens bei Lehramtsstudierenden anhand advokatorischer Fehler***

Die Erklärung komplexer schulischer Situationen auf der Grundlage wissenschaftlichen Wissens ist eine wichtige Basis effektiven Lehrerhandelns im Schulalltag (Meier, 2006; Neuweg, 2007). Ohlssons (1992) Konzept der Theorieartikulation als die „[...]Anwendung einer Theorie auf eine bestimmte Situation und eine Ausarbeitung dessen, was die Theorie implizit oder explizit über die Situation aussagt.“ stuft die Erklärung von Beobachtetem als die wichtigste Art der Theorieartikulation ein (ebd.). Lehramtsstudierende haben jedoch erhebliche Probleme mit der Anwendung wissenschaftlichen Wissens zur Erklärung schulischer Situationen (Seidel & Prenzel, 2008; Star & Strickland, 2007).

Typische Fehler, die Studierenden dabei unterlaufen, umfassen die Erklärung der Situation auf der Basis von Alltagswissen, die suboptimale Auswahl von wissenschaftlichen Theorien sowie die fehlerhafte Bezugnahme auf empirische Befunde (Stark, 2005).

Um dieser Problematik zu begegnen, wurde eine Lernumgebung auf Basis des Lernens anhand advokatorischer Fehler nach Oser (2007) entwickelt. Grundgedanke dabei ist der Aufbau negativen Wissens, welches eine wichtige Schutzfunktion gegen fehlerhaftes Handeln bietet (Oser, 2007; Oser & Spychiger, 2005). Hierzu muss ein Fehler gemacht und korrigiert werden. Dabei müssen aber die Fehler nicht selbst gemacht werden, typische Fehler können auch dargestellt und von den Lernenden nachvollzogen werden (vgl. z.B. Stark, Kopp & Fischer, 2011). Analog zum beispielbasierten Lernen müssen Lernende auch hier durch geeignete Feedbackmaßnahmen unterstützt werden (vgl. Curry, 2004; Stark, Kopp & Fischer, 2011). Möglichkeiten dazu bieten z.B. instruktionale Aufforderungen zu Abruf und Rekonstruktion des erworbenen Wissens (Prompts; Tyroller, 2005).

Um den lerntheoretischen Wirkmechanismen dieses Lehr-Lern-Ansatzes gerecht zu werden, wurden problemorientierte und instruktionsorientierte didaktische Designprinzipien in einer fallbasierten, integrierten Lernumgebung (Reinmann & Mandl, 2006) miteinander verwoben.

Zur Umsetzung problemorientierter Designprinzipien wurde ein narrativer Rahmen für die Lernumgebung gestaltet, in dem der Protagonist in vier Trainingsszenarien komplexe, problematische Schulsituationen erlebt und bei deren Erklärung substanzielle Fehler macht, die von einem fiktiven Schulpsychologen korrigiert werden. Instruktionsorientierte Designprinzipien wurden durch die Umsetzung der Szenarien als ausgearbeitete Lösungsbeispiele berücksichtigt. Aufgabe der Lernenden war es, die Lösungsbeispiele zu den Themen Bullying, Lernmotivation, Gruppenprozesse und Disziplinprobleme sorgfältig zu lesen. Dabei wurde das Ausmaß der instruktionalen Unterstützung je nach Bedingung durch explizite Nennung von Strategien zur Fehlervermeidung und Prompts dazu variiert.

Die Lernwirksamkeit des Ansatzes konnte in einer experimentellen Feldstudie an der Universität des Saarlandes nachgewiesen werden. Alle Probanden verbesserten sich hinsichtlich des Erwerbs negativen Wissens und der Qualität der von ihnen produzierten Theorieartikulationen vom Vor- zum Nachtest und waren einer Kontrollgruppe deutlich überlegen. Die Studierenden waren nach der Intervention sehr gut in der Lage, schulische Problemsituationen zu erklären. Darüber hinaus verdeutlichte sich ein Vorteil der Präsentation der Vermeidungsstrategien und der Prompts zur Rekonstruktion der Strategien. Studierende, denen diese Formen instruktionaler Unterstützung nicht zur Verfügung standen, schnitten im Nachtest schlechter ab.

Die Ergebnisse sprechen für eine Implementation in die Ausbildung der Lehramtsstudierenden. Hierzu soll die Lernumgebung in künftigen Untersuchungen weiter optimiert werden. Im kommenden Wintersemester wird dazu u.a. eine kooperative Variante getestet. Zusätzlich könnte eine qualitative Einzelfallauswertung besonders guter, aber auch besonders schlechter Theorieartikulationen weitere Erkenntnisse über die Bedingungen erbringen, unter denen die Fördermaßnahmen optimal wirken.

Curry, L. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem solving performance. In K. Forbus, D. Gentner, and T. Regier (Hrsg.), Proceedings of the 26th Annual Conference of the Cognitive Science Society, 1548. Mahwah, NJ: Erlbaum.

Meier, A. (2006). Theorienutzungskompetenz als Schlüsselqualifikation von Lehrkräften. Überlegungen zu einem Grundauftrag von Pädagogischen Hochschulen. In: Y.M. Nakamura, C. Böckelmann & D. Tröhler (Eds). Theorie versus Praxis? (S. 89-106). Zürich: Verlag Pestalozzianum.

Neuweg, G. H. (2007). Wie grau ist alle Theorie, wie grün des Lebens goldener Baum? LehrerInnenbildung im Spannungsfeld von Theorie und Praxis. ÖFEB Newsletter 5(1), 5–15. Zugriff am 20. 01. 2013 von [http://www.bwpat.de/ausgabe12/neuweg\\_bwpat12.shtml](http://www.bwpat.de/ausgabe12/neuweg_bwpat12.shtml)

Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education. Science & Education, 1, 181-192.

Oser, F. (2007). Aus Fehlern lernen. In: M. Göhlich, Ch. Wulf & J. Zirfas (Hrsg.): Pädagogische Theorien des Lernens (S. 203-212). Weinheim und Basel: Beltz Verlag.

Oser, F. & Spychiger, M. (2005). Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur. Weinheim: Beltz

Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), Pädagogische Psychologie (S. 613-658). Weinheim: Beltz.

Seidel, T. & Prenzel, M. (2008). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), Kompetenzdiagnostik (S. 201-216). Wiesbaden: VS Verlag für Sozialwissenschaften.

Star, J.R., & Strickland, S.K. (2007). Learning to observe: Using video to improve preservice teachers' ability to notice. Journal of Mathematics Teacher Education, 11, 107-125.

Stark, R. (2005). Constructing arguments in educational discourses. In H. Gruber, C. Harteis, R. Mulder & M. Rehr (Hrsg.), Bridging individual, organisational, and cultural aspects of professional learning (S. 64–71). Regensburg: S. Roderer.

Stark, R., Kopp, V. & Fischer, M. R. (2011). Cased-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. Learning and Instruction, 21(1), 22-33





## **Auf die Lehrerbildung kommt es an! Zur Entwicklung epistemologischer Überzeugungen in der Ausbildung von Mathematiklehrpersonen**

Die Aufgabe von Lehrerausbildung besteht nicht nur in der Vermittlung fachlichen Wissens, sondern insbesondere im Aufbau berufsrelevanten Könnens. Dabei ist mit Handlungskompetenz ein situationsadäquates Zusammenfügen von kognitiven Leistungsdispositionen sowie motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten gemeint (z.B. Baumert & Kunter, 2006, 2011; Weinert, 1999, 2001). Unter dem Aspekt der kognitiven Leistungsdispositionen wird nicht nur Wissen verstanden, sondern es werden auch Überzeugungen und Werthaltungen, denen handlungsleitende Funktion zugesprochen wird, mitgedacht (z.B. Calderhead, 1996; Goldin, Rösken & Törner, 2009; Reusser, Pauli & Elmer, 2011). Trotz des Wissens um die Bedeutsamkeit berufsbezogener Überzeugungen besteht bis anhin erst wenig gesichertes Wissen, wie gut es im Rahmen der Lehrerbildung gelingt, Überzeugungen angehender Lehrkräfte in die intendierte Richtung weiterzuentwickeln. Hauptsächlich qualitativ orientierte Forschungsergebnisse aus dem amerikanischen Raum stellen bei Lehramtsstudierenden eine hohe Veränderungsresistenz bei den berufsbezogenen Überzeugungen fest (z.B. Wideen, Mayer-Smith & Moon, 1998; Kane, Sandretto & Heath, 2002). Demnach verfestigten angehende Lehrpersonen während des Studiums mehrheitlich ihre bestehenden Perspektiven durch gezielte Informationsselektion. Befunde aus quantitativen Studien und aus dem deutschsprachigen Raum, wo die Lehrerausbildung vergleichsweise lange dauert, deuten dagegen darauf hin, dass sich epistemologische Überzeugungen doch verändern lassen. So konnte beispielsweise im Rahmen von MT21 festgestellt werden, dass sich Studierende des ersten und letzten Semesters in berufsrelevanten Überzeugungen unterscheiden (Blömeke, Müller, Felbrich & Kaiser, 2008). Bestätigung erfuhr dieses Ergebnis in der Deutschschweiz anhand einer quasi-längsschnittlichen Zusatzstudie zu TEDS-M (Teacher Education and Development Study in Mathematics; Tatto et al., 2012), in welcher sich zeigte, dass angehende Primarlehrpersonen am Ende der Lehrerausbildung über eine höhere Konstruktionsorientierung und niedrigere Transmissionsorientierung verfügen als Studierende bei Ausbildungsbeginn (Biedermann, Brühwiler & Steinmann, 2012). Dieses Ergebnis scheint insofern plausibel, da in den letzten Jahren verstärkt konstruktionsorientierte Lehr-Lern-Methoden für erfolgreiche Lernentwicklungen verantwortlich gemacht wurden, worauf auch in der Lehrerausbildung gezielt Lehrkonzepte entwickelt wurden, die den Lehr-Lern-Prozess als subjektiven Konstruktionsprozess betonen und entsprechende Überzeugungen bei Studierenden verstärken sollen (Reusser et al., 2011). In diesem Beitrag interessieren die folgenden beiden Fragestellungen aus dem Schnittbereich der Pädagogischen Psychologie, der Erziehungswissenschaften und der Fachdidaktik Mathematik: (1) Lässt sich bei angehenden Grundschullehrpersonen im Verlaufe der Lehrerausbildung eine Veränderung der Überzeugungen zur Struktur sowie zum Lehren und Lernen von Mathematik feststellen? (2) Lässt sich eine Veränderung in den Überzeugungen auf unterschiedliche mathematikdidaktische Lerngelegenheiten während der Ausbildung zurückführen? Die beiden Fragen werden anhand zweier Datensätze untersucht, die aus schweizerischen Zusatzerhebungen zu TEDS-M (Oser, Biedermann, Brühwiler & Steinmann, in Vorbereitung) stammen. Die erste Frage zielt darauf ab, die oben berichteten Ergebnisse aus dem Quasi-Längsschnitt zu validieren. Zu diesem Zweck wurden bei einer Teilstichprobe (n=108) ergänzend echte Längsschnittdaten erfasst. Zur Beantwortung der zweiten Frage nach der Bedeutung mathematikdidaktischer Lerngelegenheiten wurden die in einer deutschschweizerischen Vollerhebung bei je rund 1000 Primarlehrpersonen des ersten und letzten Studiensemesters erhobenen Daten aggregiert, so dass auf Ebene der 20 Institutionen bzw. Ausbildungsgänge ein längsschnittlicher Datensatz zur Verfügung steht. Die Ergebnisse aus der longitudinalen Untersuchungsanlage erhärten die Annahme, dass angehende Primarlehrpersonen im Verlaufe der Ausbildung zunehmend dynamischere und konstruktionsorientiertere und weniger statische und transmissionsorientierte Überzeugungen einnehmen. Mittels multipler Regressionen lässt sich zeigen, dass die Veränderungen in den Überzeugungen zur Struktur von Mathematik und zum Erwerb mathematischen Wissens mit Unterschieden in den mathematikdidaktischen Lerngelegenheiten an den verschiedenen Ausbildungsstätten zusammenhängen. Damit wird die Annahme gestützt, dass problemorientierte und authentische Lerngelegenheiten mitverantwortlich sind, um Entwicklungen von Überzeugungen bei künftigen Lehrpersonen anzustossen. Die Ergebnisse werden mit Blick auf Optimierungsmöglichkeiten in der Lehrerausbildung diskutiert.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaften, 9 (4), 469-520.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften, (S. 29-53). Münster: Waxmann.

Biedermann, H., Brühwiler, C. & Steinmann, S. (2012). Making the Impossible Possible? Establishing Beliefs about Teaching and Learning during Teacher Training Courses. In J. König (Hrsg.), Teachers' Pedagogical Beliefs Definition and Operationalisation – Connections to Knowledge and Performance – Development and Change, (S. 37-52). Münster: Waxmann.

Blömeke, S., Müller, C., Felbrich, A. & Kaiser, G. (2008). Epistemologische Überzeugungen zur Mathematik. In S. Blömeke, G. Kaiser & R. H. Lehmann (Hrsg.), Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung, (S. 219-246). Münster: Waxmann.

Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Hrsg.), Handbook of educational psychology, (S. 709-725). New York: MacMillan.

Goldin, G., Rösken, B. & Törner, G. (2009). Beliefs – No Longer a Hidden Variable in Mathematical Teaching and Learning Processes. In J. Maass & W. Schlöglmann (Hrsg.), Beliefs and Attitudes in Mathematics Education: New Research Results, (S. 1-18). Rotterdam: Sense Publishers.

Kane, R., Sandretto, S. & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. Review of Educational Research, 72 (2), 177-228.

Oser, F., Biedermann, H., Brühwiler, C. & Steinmann, S., (Hrsg.). (in Vorbereitung). Zum Start bereit? Kritische Ergebnisse aus TEDS-M zur schweizerischen Lehrerbildung im internationalen Vergleich. Opladen: Budrich.

Reusser, K., Pauli, C. & Elmer, A. (2011). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), Handbuch der Forschung zum Lehrerberuf, (S. 478-496). Münster: Waxmann.

Tatto, M.T., Schwille, J., Senk, S.L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodrigues, M. & Reckase, M. (2012). Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M). Amsterdam, International Association for the Evaluation of Educational Achievement (IEA).

Weinert, F. E. (1999). Konzepte der Kompetenz. Gutachten zum OECD-Projekt „Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)“. Neuchâtel: Bundesamt für Statistik.

Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In L. H. Salganik (Hrsg.), Defining and Selecting Key Competencies, (S. 45-65). Seattle: Hogrefe & Huber.

Wideen, M., Mayer-Smith, J. & Moon, B. (1998). A critical analysis of the research on learning to teach: Making the case for an ecological perspective on inquiry. Review of Educational Research, 68 (2), 130-178.



## **Zu einer evidenzbasierten Rechtschreibdidaktik: Modellierung von Lehrerwissenskomponenten – Effekt des Fachwissens auf die Rechtschreibleistung**

Zu den Auswirkungen des fachspezifischen Wissens auf die Unterrichtsqualität und der individuellen Lernleistung gibt es noch wenige Untersuchungen. Für den Mathematikunterricht zeigen die Ergebnisse von COACTIV, dass ein hohes fachdidaktisches Wissen der Lehrkraft zu einem substantiellen Lernzuwachs bei den Schülern beiträgt. Das Fachwissen seinerseits zeigt keinen Effekt auf die Schülerleistung, es stellt aber die Grundlage für das fachdidaktische Repertoire der Lehrkraft im Unterricht dar (Kunter et al., 2011). Für den Rechtschreibunterricht hat die EVES-Studie (Roos & Schöler, 2008) den Einfluss der Unterrichtsexpertise (Weinert, Schrader & Helmke, 1990) auf die Rechtschreibleistung von Viertklässlern untersucht und findet positive Effekte in der Rechtschreibentwicklung in Abhängigkeit vom diagnostischen Wissen der Lehrkraft. Dagegen konnte sie keine Effekte des Sachwissens sichtbar machen. Die EVES-Studie erfragt das Sachwissen über das Lehrerurteil zu exemplarischen Übungsaufgaben und schließt dabei nicht-bereichsspezifische Kriterien (wie Aufbau und Gestaltung des Arbeitsblatts) in der Auswertung der Lehrerantworten ein.

Die vorliegende Studie (Corvacho del Toro, 2013) geht der Frage nach dem Einfluss des Lehrerwissens auf Rechtschreibleistung nach. Angelehnt an die COACTIV-Studie (Kunter et al., 2011) spezifiziert sie theoriegeleitet das Fachwissen und das fachdidaktische Wissen von Lehrkräften für die Bereiche Orthographie, Rechtschreiberwerb und Rechtschreibdidaktik und überprüft Effekte dieses Wissens auf die Schülerleistung. Sie möchte zur Weiterentwicklung einer evidenzbasierten Rechtschreibdidaktik beitragen. In der Modellierung der Lehrerkompetenzen und in der theoretischen Herleitung des Sachwissens unterscheidet sich die vorliegende Studie von der EVES-Studie (Roos & Schöler, 2008). Sie operationalisiert das Fachwissen über die Fähigkeit, Aufgaben zur Phonetik, Phonologie, Graphematik und Morphologie des Deutschen zu lösen. Denn es sind diese Bereiche der Sprachwissenschaft, die als grundlegend für ein tiefes Verständnis des Sprach- und Schriftsystems betrachtet werden. Darüber hinaus wurde bei der Konstruktion des Wissenstests darauf geachtet, dass die Aufgaben sowohl prozedurales als auch explizites deklaratives Wissen erfragen. Die Rechtschreibleistung wurde anhand von frei geschriebenen Texten zu einem vorgegebenen Schreibenanlass am Ende der zweiten Klasse mit der OLFA 1-2 (Thomé & Thomé, 2009) erhoben. Die Intelligenz der Schüler wurde bei der Einschulung mit dem CFT 1 erhoben (Cattel, Weiß & Osterland, 1997). Die Stichprobe setzt sich aus 29 Klassen (421 Grundschulern) zusammen.

Als Hauptprädiktor für die Rechtschreibleistung wurde in einer Mehrebenenanalyse die Intelligenz des Schülers aufgenommen. Der Einfluss des Lehrerwissens auf die Rechtschreibleistung der Schüler wurde als Haupteffekt und als Interaktionseffekt (*cross-level-interaction*) modelliert. Der Interaktionseffekt begründet sich damit, dass der Zusammenhang zwischen der Vermittlung des Lerngegenstandes und den Leistungen der Schüler sich in Abhängigkeit von u.a. individuellen Bedingungen entwickelt. Es wird ein positiver Effekt des Fachwissens der Lehrkraft auf die Rechtschreibleistung angenommen.

Das erfasste Fachwissen zeigt einen positiven nicht signifikanten Effekt auf die mittlere Rechtschreibleistung der Klasse. Der Interaktionseffekt ist signifikant auf 5% Niveau ( $G_{11} = -.10$ ). Die Richtung des Interaktionseffekts zeigt, dass die Rechtschreibleistung der Grundschüler am Ende der zweiten Klasse umso stärker von den kognitiven Fähigkeiten beeinflusst wird, je niedriger das Fachwissen der Lehrkraft ist. Diese Ergebnisse stützen Befunde aus der empirischen Unterrichtsqualitätsforschung, die zeigen, dass qualitativer Unterricht umso bedeutsamer für das Erreichen von Kompetenzen ist, je schlechter die Lernvoraussetzungen der Schüler sind (Helmke et al., 2008).

Die Fokussierung der grundlegenden Inhalte des Rechtschreiberwerbs ist vermutlich der Grund, weshalb der Interaktionseffekt des Fachwissens sichtbar gemacht werden konnte. Demnach wäre eine starke Fokussierung der spezifischen Inhalte innerhalb jeden Faches, im Gegensatz zu Erhebung einer allgemeineren Fachexpertise, für die Modellierung und bei der Operationalisierung zu empfehlen. Die Ergebnisse können als Hinweise für die Validität der postulierten Wissensinhalte interpretiert werden. Zur Professionalisierung der Lehrerausbildung sollten demnach relevante Inhalte aus der Phonetik, Phonologie, Graphematik und Morphologie des Deutschen in die Lehrerausbildung integriert werden.

Cattel, R. B.; Weiß, R. & Osterland, J. (1997). Grundintelligenztest Skala 1 - CFT 1. Handanweisung (5., revidierte Aufl.). Göttingen: Hogrefe.

Corvacho del Toro, I. (2013). Fachwissen von Grundschullehrkräften. Effekt auf die Rechtschreibleistung von Grundschulern. Bamberg: University of Bamberg Press [i. D.].

Helmke, A., Helmke, T., Schrader, F.W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In: DESI-Konsortium (Hrsg.), Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim u.a.: Beltz, S. 345-363.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster u. a.: Waxmann.

Roos, J. & Schöler, H. (Hrsg.). (2009). Entwicklung des Schriftspracherwerbs in der Grundschule. Längsschnittanalyse zweier Kohorten über die Grundschulzeit. Wiesbaden: VS.

Thomé, D. & Thomé, G. (2009). OLFA 1 - 2. Oldenburger Fehleranalyse für die Klassen 1 und 2; Handbuch und Instrument (mit Kopiervorlagen). Oldenburg: isb.

Weinert, F. E., Schrader, F. W. & Helmke, A. (1990). Unterrichtsexpertise – ein Konzept zur Verringerung der Kluft zwischen zwei theoretischen Paradigmen. In: L. M. Alisch, J. Baumert und K. Beck (Hrsg.), Professionswissen und Professionalisierung. Braunschweig: Technische Universität, Seminar für Soziologie und Sozialwissenschaften, 173 – 206.



## **Validierung eines Fragebogens zur Erfassung der Lehrkompetenz von Hochschullehrenden (LeKo)**

Zur fächerübergreifenden Erfassung der didaktischen Kompetenz von Lehrenden an Hochschulen wurde an der Freien Universität Berlin ein Fragebogen neu entwickelt. Dieser basiert auf einem Modell der Unterrichtsqualität (Slavin, 1996; Ditton, 2006; Helmke, 2003; Kunter & Voss, 2011), das folgende Dimensionen unterscheidet: 1. Vermittlung von Wissen und Unterstützung von Verstehen, 2. Motivierung und 3. Steuerung der Interaktion in der Lerngruppe. Es wurden 16 Skalen entwickelt, die relevante Teilkompetenzen differenziert abbilden.

Mit dem Ziel der faktoriellen und kriterialen Validierung des Fragebogens wurden an der Freien Universität Berlin drei Teilstudien durchgeführt. In Teilstudie 1 (n= 1132 Studierende in 52 LV) wurde explorativ die faktorielle Struktur des Fragebogens ermittelt. Diese entspricht weitgehend den theoretisch entwickelten Skalen (Thiel, Blüthmann, & Watermann, 2012). In Teilstudie 2 (n= 1547 Studierende in 74 LV) wurde die Faktorstruktur konfirmatorisch bestätigt und es konnte gezeigt werden, dass die Faktorstruktur über Subgruppen von/aus Lehrveranstaltungsform (Vorlesungen und Seminare) und Fächerkultur (naturwissenschaftliche und nicht-naturwissenschaftliche LV) invariant ist.

Teilstudie 3 (n= 2145 Studierende in 105 LV) diente der Konstrukt- bzw. Kriteriumsvalidierung des Fragebogens und stellt den Schwerpunkt des Beitrags dar. Es wurde der Zusammenhang der Lehrkompetenzeinschätzungen (LeKo) mit dem Fachkompetenzzuwachs der Studierenden (erfasst mit dem bereits validierten Fragebogen BEvaKomp (Braun, Gusy, Leidner, & Hannover, 2008): Berliner Evaluationsinstrument für selbsteingeschätzte studentische Kompetenzen) untersucht. Unter Kontrolle individueller Voraussetzungen (Interesse, Vorwissen, Geschlecht und Fachsemester) der Studierenden und ihren Einschätzungen der Lehrkompetenz (Level 1) wurden auf Lehrveranstaltungsebene der Einfluss der aggregierten Einschätzungen der Lehrkompetenz (Level 2) sowie der LV-Größe auf den Fachkompetenzzuwachs analysiert.

Die Ergebnisse aus Teilstudie 3 zeigen auf Ebene der Studierenden (Level 1) einen signifikanten Einfluss von Interesse und Vorwissen sowie der Einschätzungen der Lehrkompetenzen „Verständlich erklären“, „Monotonie vermeiden“, „zur Selbsttätigkeit anregen“ und „Relevanz verdeutlichen“ auf den Fachkompetenzzuwachs der Studierenden. Auf der Ebene von Lehrveranstaltungen zeigt unter Kontrolle der Lehrkompetenzen „Zeitmanagement“ und „Umgang mit Störungen“ die Teilkompetenz „Erklären“ einen signifikanten positiven Kontexteffekt. Dies bedeutet, dass die aggregierte Einschätzung der Kompetenz des Lehrenden, Sachverhalte zu erklären, einen eigenständigen Effekt auf den Fachkompetenzzuwachs hat. Geschlecht und Fachsemester der Studierenden (Level 1) haben – ebenso wie die Größe der LV (Level 2) – keinen zusätzlichen signifikanten Einfluss auf den Fachkompetenzzuwachs. Weiterhin konnte mit den Daten aus Teilstudie 3 gezeigt werden, dass ein Modell mit den drei Dimensionen als Faktoren 2. Ordnung eine gute Anpassung auf die Daten aufweist. Die Ergebnisse der drei Teilstudien sprechen für die Validität und fächerübergreifende Einsetzbarkeit des für die Hochschullehre basierend auf Befunden der Unterrichtsforschung neu entwickelten Fragebogens.

Braun, E., Gusy, B., Leidner, B. & Hannover, B. (2008). Kompetenzorientierte Lehrevaluation – Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). *Diagnostica*, 54 (1), 30-42.

Ditton, H. (2006). Unterrichtsqualität. In: Arnold, K.-H., Sandfuchs, U. & Wiechmann, J. (Hrsg.): *Handbuch Unterricht*. Bad Heilbrunn: Klinkhardt, S. 235-243.

Helmke, A. (2003). Unterrichtsevaluation: Verfahren und Instrumente. *Schulmanagement*, Heft 1, S. 8-11.

Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In: Kunter, M. et al. (Hrsg.): *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster et al.: Waxmann, S. 85-114.

Slavin, R. E. (1996). *Education for all*. Lisse [u.a.]: Swets & Zeitlinger.

Thiel, F., Blüthmann, I. & Watermann, R. (2012). Konstruktion eines Fragebogens zur Erfassung der Lehrkompetenz (LeKo). In: B. Berendt, Voss, H.-P. & J. Wildt (Hrsg.): *Neues Handbuch Hochschullehre, Loseblattsammlung*. 55. Ergänzungslieferung, Beitrag I 1.13, 27 S., Stuttgart: Raabe Verlag.

Braun, E., Gusy, B., Leidner, B. & Hannover, B. (2008). Kompetenzorientierte Lehrevaluation – Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). *Diagnostica*, 54 (1), 30-42.

Ditton, H. (2006). Unterrichtsqualität. In: Arnold, K.-H., Sandfuchs, U. & Wiechmann, J. (Hrsg.): *Handbuch Unterricht*. Bad Heilbrunn: Klinkhardt, S. 235-243.

Helmke, A. (2003). Unterrichtsevaluation: Verfahren und Instrumente. *Schulmanagement*, Heft 1, S. 8-11.

Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In: Kunter, M. et al. (Hrsg.): *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster et al.: Waxmann, S. 85-114.

Slavin, R. E. (1996). *Education for all*. Lisse [u.a.]: Swets & Zeitlinger.

Thiel, F., Blüthmann, I. & Watermann, R. (2012). Konstruktion eines Fragebogens zur Erfassung der Lehrkompetenz (LeKo). In: B. Berendt, Voss, H.-P. & J. Wildt (Hrsg.): *Neues Handbuch Hochschullehre, Loseblattsammlung*. 55. Ergänzungslieferung, Beitrag I 1.13, 27 S., Stuttgart: Raabe Verlag.



Chair(s): Korbinian Moeller (Leibniz-Institut für Wissensmedien)

Diskutant/in: Laura Martignon (Pädagogische Hochschule Ludwigsburg)

Zahlen sind ein integraler und nicht mehr wegzudenkender Bestandteil unseres Lebens zu Beginn des 21. Jahrhunderts. Moderne Computer, das Internet, aber auch Ingenieurwesen und Medizin sind ohne Zahlen und deren kognitive Verarbeitung kaum vorstellbar. Die Fähigkeit Zahlen und die mit ihnen übertragenen Informationen zu verstehen und adäquat mit diesen umzugehen wird somit zu einem Schlüsselkapital in unserer post-industriellen Informationsgesellschaft (z.B. Beddington et al., 2008). Dementsprechend negativ wirken sich auch unzureichende numerische Fähigkeiten aus und dies sowohl auf individueller wie gesellschaftlicher Ebene. Sieht man sich die Lebenswege der betroffenen Individuen genauer an, so zeigt sich, dass diese nicht nur überdurchschnittlich häufig schulische Probleme haben, sondern in der Folge meist auch schlechtere Aussichten auf dem Arbeitsmarkt bzw. größere soziale Probleme haben (Bynner & Parsons, 2006; Gross, Hudson, & Price, 2009). Auf gesellschaftlicher Ebene ergeben sich daraus immense Belastungen. Für Großbritannien veranschlagen erste Schätzungen die jährlich entstehenden Kosten, die durch unzureichende numerische Fähigkeiten entstehen auf ca. 2,4 Milliarden Pfund (Gross et al., 2009).

Aus dem vorangegangenen wird klar, dass die Frage numerisch-mathematischer Fähigkeiten gerade in Zeiten wirtschaftlicher Krisen in Europa nicht nur auf individueller Ebene relevant ist, sondern als sogenanntes „mentales Kapital“ gerade auch für unsere post-industriellen Wissensgesellschaften als Ganzes von immenser sozialer aber auch ökonomischer Bedeutung ist (Beddington et al., 2008). Im vorliegenden Symposium soll an ausgewählten Beispielen die wissenschaftliche Breite dargestellt werden, in welcher sich die empirische Bildungsforschung mit dem Thema numerisch-mathematischer Fähigkeiten auseinandersetzt um damit die Vielschichtigkeit des Themas abzubilden und eine Disziplin übergreifende Diskussion von Möglichkeiten und ggf. Grenzen evidenzbasierter Bildungsforschung anzustoßen.

Der erste Beitrag (Klein et al.) stellt neue neurowissenschaftliche Erkenntnisse darüber vor, wie strukturelle bzw. funktionale Aspekte des Gehirns interindividuelle Unterschiede in numerischen Kompetenzen von frühgeborenen Kindern beeinflussen - eine Risikogruppe für die Entwicklung von Rechenschwächen.

Im zweiten Beitrag (Pixner) wird eine neue Idee zum Training von Multiplikationsfakten vorgestellt, die auf eine systematische Assoziation von Zahlen und Farben aufbaut.

Der dritte Beitrag (Wasner et al.) sucht die Frage inwieweit Fingerzählen bzw. -rechnen die numerische Entwicklung während der ersten Klasse beeinflusst in einer Längsschnittinterventionsstudie zu beantworten.

Im vierten Beitrag (Lambert & Flunger) werden Vor- und Nachteile verschiedener Kriterien zur möglichen Identifizierung rechenschwacher Kinder in Large-Scale-Untersuchungen hinsichtlich ihrer Konsistenz und Prognose verglichen.

Abschließend werden mögliche Implikationen und Perspektiven dieser empirischen Ergebnisse aus mathematik-didaktischer Sicht bewertet und diskutiert (Martignon).

Beddington, J., Cooper, C.L., Field, J., Goswami, U., Huppert, F.A., Jenkins, R., Jones, H.S., Kirkwood, T.B., Sahakian, B.J., & Thomas, S.M. (2008). The mental wealth of nations. *Nature*. 455, 1057-1060.

Gross, J., Hudson, C. and Price, D. (2009). *The Long Term Costs of Numeracy Difficulties*. Every Child a Chance Trust and KPMG: London.

Parsons, S. & Bynner, J. (2005). Does numeracy matter more? NRDC (National Research and Development Centre for adult literacy and numeracy).

Montag,

03.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 0.254

Elise Klein (Leibniz-Institut für Wissensmedien, Tübingen), Ursula Kiechl-Kohlendorfer (Medizinische Universität Innsbruck, Österreich), Christian Kremser (Medizinische Universität Innsbruck, Österreich), Michael Schocke (Medizinische Universität Innsbruck, Österreich), Korbinian Moeller (Leibniz-Institut für Wissensmedien, Tübingen), Liane Kaufmann (Tiroler Landeskrankenhaus, Hall in Tirol, Österreich)

## **Neuronale Korrelate der Verarbeitung physikalischer und numerischer Größe in frühgeborenen Kindern**

In der vorliegenden Studie wurden die Zusammenhänge zwischen strukturellen und funktionalen Merkmalen des Gehirns und numerischen Fähigkeiten in einer Gruppe von 6-7-jährigen frühgeborenen Kindern - einer Hochrisikopopulation für die Entwicklung von numerischen Lernstörungen - untersucht. Mögliche Zusammenhänge zwischen Hirnstruktur bzw. -funktion und der Fähigkeit zur Verarbeitung von physikalischer und numerischer Größe wurden mit verschiedenen Methoden in zwei Experimenten beobachtet. Dabei wurden klassische Verhaltensdaten wie Reaktionszeiten sowohl mit Daten zur Gehirnstruktur (voxelbasierte Morphometrie, VBM, Starke et al., 2013) als auch mit funktionaler Aktivierung korreliert (funktionelle Magnetresonanztomografie, fMRT).

Für die VBM-Analyse wurden die Dichtewerte von grauer sowie weißer Substanz aus Gehirnregionen extrahiert, die in einer Meta-analyse als relevant für die numerische Entwicklung identifiziert wurden (Kaufmann et al., 2011). Im Folgenden wurden diese Werte mit klassischen Verhaltensdaten korreliert: zum einen Reaktionszeiten der Erfassung von semantischem Zahlenverständnis (ermittelt als numerischer Distanzeffekt in einer Zahlenvergleichsaufgabe) sowie andererseits allgemeineren mathematischen Fertigkeiten (erhoben mittels eines standardisierten Rechentestes, TEDI-MATH, Kaufmann et al., 2009).

Ein in den Verhaltensdaten signifikanter numerischer Distanzeffekt (d.h., schnellere Reaktionen beim Vergleich von Zahlenpaaren mit großer Distanz. z.B. 1\_9 vs. 5\_6) zeigte eine generell intakte Repräsentation numerischer Größe von einstelligen Zahlen in der untersuchten Population. Darüber hinaus beobachteten wir eine positive Korrelation zwischen den Dichtewerten grauer Substanz aus der VBM-Analyse und der Größe des Distanzeffektes bilateral in parietalen Arealen (inklusive des rechten anterioren inferioren sowie des linken superioren Parietallappens) sowie dem rechten superioren temporalen Gyrus. Dagegen waren die Dichtewerte der weißen Substanz im rechten anterioren inferioren Parietallappen und dem inferior frontalen Gyrus negativ mit dem Distanzeffekt korreliert. Damit zeigen die VBM-Daten eine Assoziation individueller Unterschiede sowohl in der grauen als auch der weißen Substanz mit numerischen Fähigkeiten in frühgeborenen Kindern.

In der fMRT-Studie wurden dieselben Kinder in Aufgaben zur Verarbeitung von numerischer und physikalischer Größe (physikalische Stroop-Aufgabe) untersucht. Wieder zeigte sich in den Verhaltensdaten (sowohl Reaktionszeiten als auch Fehlerdaten) ein signifikanter numerischer Distanzeffekt. Für die Zahlenvergleichsaufgabe fanden wir darüber hinaus eine negative Korrelation des Distanzeffektes mit Aktivierung im rechten intraparietalen Sulcus, während die Aktivierung im linken Gyrus angularis positiv mit dem Distanzeffekt korrelierte. Weiterhin war Alter in beiden Aufgaben signifikant und positiv mit der beobachteten Aktivierung im linken Gyrus angularis korreliert. Für die physikalische Größenvergleichsaufgabe zeigte sich lediglich eine negative Korrelation von physikalischer Größe mit Aktivierung in superior frontalen Arealen. Zusammenfassend läßt sich sagen, dass die fMRT-Daten in der Risikogruppe der Frühgeborenen gemeinsame fronto-parietale Aktivierungen für die Verarbeitung von physikalischer und numerischer Größe zeigen. Interessanterweise gibt es aber intraparietal individuelle Signalunterschiede abhängig von numerischen Fähigkeiten und Alter.

Die vorliegenden Ergebnisse der beiden Studien sind insoweit von hohem Interesse, als dass sie zeigen, dass bereits in 6-7 Jahre alten frühgeborenen Kindern individuelle Unterschiede sowohl in der grauen als auch der weißen Substanz mit interindividuellen Unterschieden in numerischen Fähigkeiten assoziiert sind. Auch in den beobachteten Aktivierungsmustern zeigten sich Unterschiede in parietalen Arealen, die sowohl mit numerischen Fähigkeiten als auch dem Alter zusammen hingen. Hierbei ist allerdings wichtig zu betonen, dass der in der vorliegenden Studie beobachtete Zusammenhang zwischen Gehirnstruktur und Verhaltensdaten ausschließlich in den Experimentalaufgaben zu beobachten war, in denen semantisches Größenverständnis (numerisch und physikalisch) untersucht wurde, wohingegen generelle mathematische Fähigkeiten nicht mit individuellen Unterschieden in der Gehirnstruktur assoziiert zu sein schienen. Insgesamt unterstreichen die beiden Studien noch einmal die besondere Bedeutung des rechten intraparietalen Sulcus für die numerische Entwicklung.

Kaufmann, L., Nuerk, H.-C., Graf, M., Krinzinger, H., Delazer, M., Willmes, K. (2009). TEDIMATH: Test zur Erfassung numerisch-rechnerischer Fertigkeiten für 4-8-Jährige. Bern: Verlag Hans Huber.

Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analysis of

developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, 36, 763-787.

Starke, M., Kiechl-Kohlendorfer, U., Kucian, K., Pupp Peglow, U., Kremser, C., Schocke, M., Kaufmann, L. (2013). Brain structure, number magnitude processing, and math proficiency in 6- to 7-year-old children born prematurely: a voxel-based morphometry study. *Neuroreport*, 24, 419-424.





## **Multiplikationsfakten lernen trotz Rechenschwäche - Eine Interventionsstudie**

Mathematische Fertigkeiten sind ein elementarer Bestandteil unserer modernen Gesellschaft. Dabei spielen sie im Lernkontext und im alltäglichen Leben eine große Rolle. Niedrigere Leistungen im Schulfach Mathematik sind somit häufig mit einem schlechteren Ausbildungsniveau assoziiert (siehe auch NRDC, 2008). Deshalb ist eine möglichst frühe und effektive Förderung bei Rechenproblemen für Betroffene wesentlich und von lebenslanger Bedeutung. Da mathematische Fähigkeiten einer hierarchischen Struktur unterliegen ist ein Verständnis der Basis unabdingbar - haben Kinder z.B. den Prozess des Multiplizierens nicht verinnerlicht, dann können sie einzelne Aufgaben nicht dauerhaft im Langzeitgedächtnis behalten, wodurch es für sie in der Folge oft schwieriger wird die Division und im weiteren das Rechnen mit Brüchen, zu verstehen (de Brauwer & Fias, 2011).

Ziel der vorliegenden Studie war es daher festzustellen, mit welchen Strategien man Kindern mit spezifischen Problemen bei der Abspeicherung von Multiplikationsfakten am besten helfen kann. Insbesondere wurde der Frage nachgegangen, inwieweit systematische Assoziationen zwischen Zahlen und Farben Einfluss auf Abruf und Speicherung arithmetischer Fakten haben. Eine systematische Assoziation zwischen Farben und Zahlen ist dabei nichts Neues. Bereits 1880 beschrieb Galton die ersten Personen, die eine natürliche Assoziation zwischen Zahlen und Farben aufweisen. Bei dieser besonderen Form der Synästhesie werden häufig kleinere Zahlen mit helleren Farben und größeren Zahlen mit dunkleren Zahlen assoziiert (Smilek, Dixon, Cudahy, & Merikle, 2002). In Verbindung mit prozeduralen und konzeptuellen Aspekten scheint diese Assoziation zwischen Farben und Zahlen somit einen geeigneten Ausgangspunkt für die Abspeicherung und das Abrufen von Multiplikationen darzustellen (siehe auch Domahs, Lochy, Eibl, & Delazer, 2004 für erfolgreiche Ergebnisse bei hirngeschädigten Patienten).

Für die vorliegende Studie wurden Kinder der 4. Schulstufe mit deutlichen Problemen beim Multiplizieren ausgewählt und zufällig der Experimentalgruppe oder einer von zwei Kontrollgruppen zugeteilt. Die Experimentalgruppe lernte 8 Wochen lang (eine Schulstunde pro Woche im Einzelsetting) Multiplikationsfakten, wobei dies durch systematische Assoziation zwischen einzelnen Zahlen und bestimmten Farben (z.B. 0= grau, 1= gelb,..., 8= braun, 9= schwarz) unterstützt wurde. Die erste Kontrollgruppe absolvierte dasselbe Faktentraining jedoch ohne die spezifische Assoziation zwischen Zahlen und Farben. Dieses Faktentraining baute dabei auf prozeduralem und konzeptuellem Wissen der Kinder über Multiplikation auf, das in jeder Einheit mit adressiert wurde. Um das prozedurale Verständnis zu stärken lernten die Kinder die Multiplikation über das wiederholte Addieren verstehen (z.B.  $3 \times 4$  kann auch als  $4 + 4 + 4$  gerechnet werden). Bezüglich des konzeptuellen Wissens wurden Regelmäßigkeiten (z.B. Multiplizieren mit 0, 1 und 10; Komplementäraufgaben  $2 \times 3 = 3 \times 2$ , usw.) vermittelt. Das Faktentraining diente dem Erwerb und der Automatisierung der einzelnen Aufgaben um deren assoziative Stärke im Faktennetzwerk zu erhöhen (Shrager & Siegler, 1998). Eine zusätzliche Kontrollgruppe, die ein Rechtschreibtraining im gleichen Umfang absolvierte wurde in das Design aufgenommen um multiplikationsspezifische Veränderungen erfassen zu können.

Wie erwartet, zeigte die Experimentalgruppe die deutlichsten Verbesserungen im Vergleich zu beiden Kontrollgruppen: Die Multiplikationsleistung stieg in der Experimentalgruppe am stärksten an. Jedoch nicht nur die stärksten Verbesserungen in den direkten Trainingseffekten, sondern vor allem auch beobachtete Transfereffekte auf nicht direkt trainierte Aufgaben sprechen für die Validität einer systematischen Assoziation von Zahlen und Farben im Rahmen eines Trainingsprogramms für Kinder mit Defiziten beim Multiplizieren. Theoretisch implizieren diese Ergebnisse, dass eine systematische Assoziation zwischen Zahlen und Farben funktionale Relevanz in der Abspeicherung von Multiplikationsfakten erlangen könnte.

De Brauwer, J., & Fias, W. (2011). The Representation of Multiplication and Division Facts in Memory. *Experimental Psychology*, 58, 312-323.

Domahs, F., Lochy, A., Eibl, G., & Delazer, M. (2004). Adding colour to multiplication: Rehabilitation of arithmetic fact retrieval in a case of traumatic brain injury. *Neuropsychological Rehabilitation*, 14, 303-328.

Galton, F. (1880). Visualised numerals. *Nature*, 21, 252-256.

NRDC (2008). Research Briefing Numeracy. London: NRDC

Shrager, J. & Siegler, R.S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, 9, 405-410.

Smilek, D., Dixon, M.J., Cudahy, C., & Merikle, P.M. (2002). Concept driven color experiences in digit-color synesthesia. *Brain and Cognition*, 48, 570-573.



## ***Fingerrechnen - hilfreich oder hinderlich beim Erlernen numerischer Fähigkeiten in der 1. Klasse?***

Fingerzählen bzw. -rechnen und dessen mentale Repräsentation beeinflussen den Umgang mit arabischen Zahlen grundlegend. In einem bestimmten Zeitraum ihrer numerischen Entwicklung benutzen (fast) alle Kinder in irgendeiner Form ihre Finger zum Zählen und/oder initialen Rechnen. Ob und wie lange Fingerrechnen hilfreich für das Erlernen numerischer Fähigkeiten ist, wird allerdings kontrovers diskutiert (vgl. Moeller, Martignon, Wessolowski, Engel, & Nuerk, 2011): folgt man der neuropsychologischen Literatur sind Kinder mit guter Fingeragnosie (d.h. der Fähigkeit Finger zu unterscheiden bzw. zuzuordnen) tendenziell die besseren Rechner. Außerdem wird nach dem Ansatz der sog. Embodied Cognition davon ausgegangen, dass Fingerrechnen eine spezifische Verknüpfung von körperlichem Erleben und abstrakten numerischen Repräsentationen darstellt. Daraus wird geschlossen, dass Fingerzählen/-rechnen früh didaktisch vermittelt werden sollte. In der mathematisch-fachdidaktischen Literatur wird hingegen eher betont, dass fingerzählendes Rechnen von Schulanfängern möglichst bald durch fingerlose Rechenstrategien abgelöst werden sollte. Im Zusammenhang dieser Kontroverse werden drei Hypothesen diskutiert (Moeller & Nuerk, 2012): i) Vorläuferhypothese (Fingerzählen/-rechnen ist eine Vorläuferfähigkeit, die mit der abstrakten Repräsentation einer Zahl überdauernd verknüpft ist), ii) Übergangshypothese (Fingerzählen/-rechnen stellt den Übergang zu abstrakter Zahlenverarbeitung dar, der im Laufe der Entwicklung überwunden wird) und iii) Sackgassenhypothese (Fingerzählen/-rechnen bietet eine zu einfache Hilfe für zählendes Rechnen und erschwert dadurch den Übergang zu abstrakteren Repräsentationen).

Bisher gibt es jedoch keine empirische Längsschnittstudie, die untersucht hat, ob, wie lange und für welche Kinder Fingerzählen und -rechnen hilfreich oder hinderlich ist. Um diese Lücke zu schließen und die empirische Basis dafür schaffen, ab wann und bis wann Fingerrechnen im schulischen Alltag gefördert bzw. unterbunden werden sollte, wurde in der vorliegenden Studie die Entwicklung numerischer Fähigkeiten in insgesamt 21 Klassen (n = 389, Alter M = 6.45 Jahre, SD = 0.51) über das 1. Schuljahr hinweg untersucht. In 11 Interventionsklassen wurde dabei ein systematisches Fingerzähl- und Fingerrechenstraining in den Mathematikunterricht integriert. In drei inhaltlich verschiedenen Modulen wurde die jeweilige Entwicklung der Kinder mit der von Kindern aus 10 Kontrollklassen, die den Regelunterricht ohne die systematische Verstärkung des Fingerrechnens besuchten, verglichen. Jeder Baustein enthielt eine Prä-/Post-Messung und sechs Interventionseinheiten à 25 Minuten pro Woche. Inhaltlich lag der Fokus im 1. Baustein auf den Themen Fingeragnosie, Mengenwissen und Fingerrepräsentationen, im 2. Baustein wurde schwerpunktmäßig Zehner-Ergänzung, Addition und Subtraktion bis 10 mit den Fingern geübt und im 3. Baustein wurde der Übertrag über 10 mit den Fingern trainiert sowie der Zahlenraum bis 20 erweitert.

Die Ergebnisse des ersten und zweiten Moduls zeigen – bei vergleichbarem Ausgangsniveau vor Beginn des ersten Moduls – spezifische Verbesserungen der Kinder in den Interventionsklassen in Additions- und Subtraktionsaufgaben, sowohl bei leichteren als auch bei schweren Aufgaben. Diese stärkere Additions- bzw. Subtraktionsleistung der Kinder mit Fingerzähl-/Fingerrechenstraining nach dem ersten Baustein war darüber hinaus stabil auch über das zweite Modul hinweg zu beobachten. In anderen mathematisch-numerischen Aufgaben (z.B. Vorläuferfähigkeiten, Zahlenstrahl) gab es keine signifikanten Gruppenunterschiede.

Die Ergebnisse sprechen dafür, dass systematisches Trainieren von Fingerzählen und -rechnen das Erlernen von arithmetischen Aufgaben, wie Addition und Subtraktion unterstützen kann. Die förderlichen Effekte scheinen dabei spezifisch für Addition und Subtraktion zu sein. In anderen Maßen schneiden die Kinder, die das Fingerrechen-Training erhalten haben, aber keineswegs schlechter ab als die Kinder der Kontrollklassen, sodass die Kinder der Interventionsklassen entweder gleich gute oder sogar bessere Ergebnisse als die Kinder der Kontrollklassen erzielten. Ausgehend von diesen Ergebnissen kann geschlussfolgert werden, dass von Beginn bis mindestens Mitte der ersten Klasse das Verwenden der Finger zum Erlernen numerischer Fähigkeiten nicht hinderlich, sondern sogar eher förderlich zu sein scheint, was der Sackgassenhypothese widerspricht. Geschlechtsunterschiede zugunsten der Jungen und inwieweit vor allem rechenschwache oder –starke Kinder spezifisch von dem Fingerrechen-Training profitiert haben, werden diskutiert.

Moeller, K., Martignon, L., Wessolowski, S., Engel, J., & Nuerk, H.-C. (2011). Effects of finger counting on numerical development: A review of neuro-cognitive arguments and those from mathematics education. *Frontiers in Cognition*, 2:328.

Moeller, K., & Nuerk, H.-C. (2012). Zählen und Rechnen mit den Fingern: Hilfe, Sackgasse oder bloßer Übergang auf dem Weg zu komplexen arithmetischen Kompetenzen? *Lernen und Lernstörungen*, 1, 33-53.



## **Vor- und Nachteile verschiedener Kategorisierungssysteme zur Identifizierung rechenschwacher Kinder in Large-Scale-Untersuchungen**

Studien deuten darauf hin, dass Defizite im Umgang mit Zahlen und Rechenoperationen beträchtliche Auswirkungen auf die spätere gesellschaftliche Teilhabe, wie Arbeitsmarktchancen, das Gehaltsniveau, aber auch die psychische Gesundheit haben können (z.B. Beddington et al., 2008; Parsons & Bynner, 2005). Vor diesem Hintergrund kommt der Identifizierung rechenschwacher Kinder und Jugendlicher nicht nur im frühen Schulalter, sondern gerade auch nach dem Übergang in die weiterführende Schule eine erhebliche Bedeutung zu.

Ein zentrales Problem in Forschung und Praxis ist dabei die Frage, welches Kriterium für die Diagnosestellung am geeignetsten ist und welches den weiteren Entwicklungsverlauf am besten vorhersagen kann. Die am häufigsten verwendeten Kriterien, das Diskrepanzkriterium und die Verwendung von Grenzwerten, wurden vielfach kritisiert und in Frage gestellt, da beispielsweise die Vorhersage auf Basis des Diskrepanzkriteriums wenig stabil zu sein scheint (z.B. Shalev, Manor, & Gross-Tsur, 2005; Silver et al., 1999). So konnte gezeigt werden, dass die Verwendung unterschiedlicher Kriterien zu teilweise erheblichen Unterschieden in Untersuchungsergebnissen führen können (z.B. Mazzocco & Myers, 2003).

Studien, in denen an einer Stichprobe unterschiedliche Kategorisierungen vergleichend untersucht werden, beziehen sich jedoch meist ausschließlich auf Kinder mit schwachen Rechenleistungen, wodurch die Stichprobengrößen und damit die Zellohäufigkeiten gering waren. Large-Scale Daten wurden bisher für diese Fragestellung nicht verwendet. Darüber hinaus existieren insbesondere im deutschen Sprachraum kaum Untersuchungen, die die Leistungsentwicklung von Kindern und Jugendlichen mit Schwierigkeiten in Mathematik in der weiterführenden Schule untersuchten.

Mit Bezugnahme auf diese Forschungslücken wurde die vorliegende Studie durchgeführt. An N = 2008 Schülerinnen und Schüler der TRAIN-Stichprobe wurde überprüft, wie stark die mit divergierenden Kriterien erzielten Kategorisierungen der Mathematikleistungen von Kindern und in Haupt-, Real- und Mittelschule im Längsschnitt voneinander abweichen (5. und 8. Jahrgangsstufe).

Einbezogene Schülermerkmale waren neben der Mathematikleistung, die Leseleistung, Schulnoten und kognitive Grundfähigkeiten. Die Gruppenbildung wurde dabei systematisch variiert. Zunächst wurden auf Basis des Diskrepanzkriteriums vier Gruppen gebildet. Als „rechenschwach“ (RS) oder „leseschwach“ (LS) wurden alle Kinder klassifiziert, bei denen zwischen IQ und Mathematikleistung eine Diskrepanz von  $\geq 1,5$  SD lag. Die dritte Gruppe bestand aus Kindern, die sowohl in Mathematik wie auch beim Lesen eine Diskrepanz aufwiesen („kombinierte Störungen“, KS). Alle anderen Kinder wurden als „unauffällig“ (UA) eingestuft. In einem zweiten Schritt wurden fünf Kategorien, basierend auf Cut-Off-Kriterien, gebildet: UA (Lese- und Mathematikleistung über PR 25), starke RS (PR < 11), leichte RS (11 < PR < 25), LS (PR < 11) und KS (RS & LS: PR < 11). Darüber hinaus wurde die latente Profilanalyse (Muthén, 2001) eingesetzt, um zu prüfen, ob anhand der Profile in den einbezogenen Merkmalen Typen von Schülerinnen und Schülern auch auf Basis eines rein statistischen Verfahrens unterschieden werden können. Das Grundprinzip dieses personenorientierten Verfahrens ist es, Schülerinnen und Schüler, die ein ähnliches Profil aufweisen, zu einem „Typ“ zu gruppieren. Auf diese Weise lassen sich bei Schülerinnen und Schülern Gruppen bestimmen, die nicht direkt messbar sind. Die mit der latenten Profilanalyse erzielten Lösungen zeigten, dass inhaltlich ein Fünf-Gruppenmodell am besten mit den Daten übereinstimmte. Allerdings bildeten die so gefundenen Gruppen weder die mittels des Diskrepanzkriteriums, noch mittels Cut-Off-Kriterien gebildeten Kategorien ab. Mithilfe der Klassenanalyse konnte keine eindeutig „rechenschwache“ Gruppe identifizieren werden, jedoch eine Gruppe mit Leseschwierigkeiten und eine Gruppe, die als KS-Gruppe interpretierbar ist. In der längsschnittlichen Analyse vier Jahre später zeigte sich, dass die Gruppenzusammensetzungen in allen Klassifikationsvarianten zwischen dem ersten und letzten Messzeitpunkt divergierten. Die Befunde zeigen damit, dass die Mathematikleistungen sich über die Zeit verändern. Zudem liefern die Befunde Hinweise, dass die gängige Diagnosepraxis problematisch sein könnte. Um dies besser einschätzen zu können, werden im nächsten Schritt neben weiteren längsschnittlichen Analysen die Kriterien anhand von Outcome-Variablen evaluiert. Mögliche Kriterien zur Bestimmung des „besten“ Kriteriums sowie Vor- und Nachteile der einzelnen Kriterien für Forschung und Praxis werden diskutiert.

Beddington, J., Cooper, C.L., Field, J., Goswami, U., Huppert, F.A., Jenkins, R., Jones, H.S., Kirkwood, T.B., Sahakian, B.J., & Thomas, S.M. (2008). The mental wealth of nations. *Nature*, 455, 1057-1060.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disabilities in the primary school-age years. *Annals of dyslexia*, 53, 218-253.

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class-latent growth modeling. Washington, DC: American Psychological Associations.

Parsons, S. & Bynner, J. (2005). Does numeracy matter more? NRDC (National Research and Development Centre for adult literacy and numeracy).

Shalev, R. S., Manor, O. & Gross-Tsur, V. (2005). Developmental dyscalculia: a prospective six-year follow-up. *Developmental Medicine and Child Neurology*, 47, 121-125.

Silver, C.H., Pernet, B.D., Black, J.L., Fair, G.W. & Balise, R.R. (1999). Stability of arithmetic disability subtypes. *Journal of Learning Disabilities*, 32, 108-119.



Chair(s): Päivi Taskinen (Friedrich-Schiller Universität Jena)

Diskutant/in: Bärbel Kracke (Friedrich-Schiller Universität Jena)

Ein passendes Berufsziel zu entwickeln, ist eine wichtige Aufgabe im Jugendalter. Dies ist als ein längerfristiger und anforderungsreicher Entwicklungsprozess zu verstehen, der in der aktuellen Diskussion als Berufsorientierung bezeichnet wird. Obgleich sich schon Kinder mit der Arbeitswelt beschäftigen und mit möglichen Berufen für sich selbst gedanklich experimentieren, wird die Frage in der Sekundarstufe besonders bedeutsam und bis zum Schulabgang drängender. Auch wenn die Berufsorientierungsphase in eine wichtige individuelle Bildungsentscheidung mündet und jede Schule in Deutschland angehalten ist, ein Konzept für die Berufsorientierung zu entwickeln, wird das Thema Berufsorientierung in Deutschland weder auf institutioneller noch auf individueller Ebene systematisch empirisch erforscht. Es herrschen programmatische Aussagen vor, die die Notwendigkeit von Berufsorientierungsmaßnahmen herausstellen, ohne sie im Einzelnen systematisch pädagogisch oder psychologisch zu begründen. Darüber hinaus existiert eine Vielzahl von tatsächlich durchgeführten Maßnahmen, deren Wirkungen bislang kaum erforscht sind. Das Symposium hat das Ziel, unterschiedliche Forschungsansätze, die zum Teil auch aus unterschiedlichen Disziplinen stammen, im Bereich der Berufsorientierung aufzuzeigen und zu verbinden.

Das Symposium stellt zunächst zwei unterschiedliche theoretische Zugänge zur Betrachtung von Berufsorientierung der Jugendlichen dar. Der erste Beitrag befasst sich aus entwicklungspsychologischer Perspektive mit der Frage, wie angemessenes Engagement beim Übergang von der Schule zur nachschulischen Bildung aussehen kann. Der zweite Beitrag fokussiert die Aktivitäten der Jugendlichen im Berufswahlprozess aus der Perspektive der sozial-kognitiven Laufbahntheorie von Lent, Brown & Hackett. Es werden jeweils Ergebnisse aus Längsschnittstudien vorgestellt. Der dritte Beitrag fokussiert die Frage, inwieweit Berufsorientierung für weibliche und männliche Jugendliche systematisch unterschiedlich verläuft. Es werden anhand einer Längsschnittstudie Erkenntnisse über stereotype Selbstzuschreibungen gewonnen, die für die Berufswahl relevant sind. Es werden anschließend Ergebnisse einer Interventionsstudie vorgestellt, die die Verminderung solcher Selbstzuschreibungen zum Ziel hatte. Der vierte Beitrag stellt vor, wie im Rahmen eines entwicklungs- und pädagogisch psychologisch begründeten Modells der Berufswahlkompetenz eine Interventionsmaßnahme systematisch entworfen und in ihrer Wirksamkeit empirisch überprüft werden kann.

Insgesamt vereint das Symposium Erkenntnisse aus der Berufspädagogik, Erziehungswissenschaft und Psychologie und kann zu einer künftigen verstärkten wechselseitigen Bezugnahme der Disziplinen und damit zu einem fundierteren Verständnis der Berufswahlprozesse im Jugendalter beitragen.

Bandura, A. (1997). *Self-Efficacy. The exercise of control*. New York: Freeman.

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Dietrich, J., Parker, P., Salmela-Aro, K. (2012). Phase-adequate engagement at the post-school transition. *Developmental Psychology*, 48, 1575-1593.

Driesel-Lange, K.; Hany, E.; Kracke, B. & Schindler, N. (2010). Berufs- und

Studienorientierung. Erfolgreich zur Berufswahl. Ein Orientierungs- und Handlungsmodell für Thüringer Schulen. In: Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien (Hrsg.): Materialien Nr. 165. Bad Berka.

Hany, E. & Driesel-Lange, K. (2006). Berufswahl als pädagogische Herausforderung: Schulische Orientierungsmaßnahmen im Urteil von Abiturienten. *Diskurs Kindheits- und Jugendforschung*, 4, 517-531.

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D. & Besuch, G. (2010). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. Hannover: HIS-Hochschul-Informationssystem GmbH.

Lent, R. W., Brown, S. D. & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79-122.

Lent, R. W., Brown, S. D. & Hackett, G. (2002). Social cognitive career theory. In D. Brown and Associates (Hrsg.), *Career choice and development* (4. Aufl.) (S. 255-311). San Francisco, CA: Jossey-Bass.

Prager, U. & Wieland, C. (2005). Jugend und Beruf. Repräsentativbefragung zur Selbstwahrnehmung der Jugend in Deutschland. Gütersloh: Bertelsmann-Stiftung.

Ratschinski, G. (2008). Berufswahlkompetenz. In M. Koch & P. Straßer (Hrsg.), *In der Tat kompetent. Zum Verständnis von Kompetenz und Tätigkeit in der beruflichen Benachteiligtenförderung* (S. 73-90). Bielefeld: wbv.

Ratschinski, G. (2012). Berufswahlkompetenz. Versuch einer zeitgemäßen Operationalisierung der Berufswahlreife. In G. Ratschinski & A. Steuber (Hrsg.), *Ausbildungsreife. Kontroversen, Alternativen und Förderansätze* (S. 135-156). Wiesbaden: Springer VS.

Schmidt-Koddenberg, A. & Zorn, S. (2010). *Zukunft gesucht! Berufs- und Studienorientierung in der Sek. II*. Opladen: Budrich.

Montag,

03.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 1.411



## ***Angemessenes Engagement beim Übergang von der Schule zur nachschulischen Ausbildung***

Sich für einen Beruf zu entscheiden ist eine Entwicklungsaufgabe, die am Ende der Schulzeit an Dringlichkeit gewinnt. Mit Blick auf den Übergang vom Gymnasium in die nachschulische Ausbildung befasst sich die Studie mit phasenadäquatem (angemessenem) Engagement, d.h. zielgerichteten psychologischen Prozessen, mit denen Jugendliche den Übergang zu meistern versuchen (Dietrich, Parker, & Salmela-Aro, 2012). Verschiedene Theorien der Entwicklungsregulation und der beruflichen Entwicklung haben sich mit dem Konstrukt des phasenadäquaten Engagements befasst. Es umfasst u.a. jenes Verhalten, das von Jugendlichen unternommen wird, um Ziele im Zusammenhang mit bevorstehenden Übergängen und Entwicklungsaufgaben zu erreichen. Dazu gehören Ziele, wie etwa „herausfinden, welche Berufe zu meinen Stärken und Schwächen passen“ oder konkrete berufliche Ziele wie „Medizin studieren“. Wie individuell unterschiedlich vorgehen, um solche Ziele zu erreichen, ist bisher nur in Ansätzen erforscht.

Die vorliegende Studie fokussiert die Erfolgserwartungen, die Jugendliche hinsichtlich ihrer beruflichen Ziele haben. Die Erwartung Erfolg zu haben ist eine wichtige Determinante einer adaptiven Zielverfolgung. Die Studie untersucht unterschiedliche Entwicklungsverläufe zielbezogener Erfolgserwartungen über einen Zeitraum von sieben Jahren und fragt nach deren Ursachen und Konsequenzen.

Finnische Jugendliche (N = 850, 48% weiblich) wurden im Rahmen einer Längsschnittstudie zu fünf Zeitpunkten befragt, drei Mal während der Schulzeit im Alter von 16 (T1), 17 (T2) und 18 Jahren (T3), und zwei Mal nach dem Schulabschluss im Alter von 20 (T4) und 23 Jahren (T5). Zur Erfassung persönlicher Ziele nannten die Jugendlichen ein berufs(wahl)bezogenes Ziel und bewerteten auf einer 7-stufigen Likertskala, inwieweit sie erwarteten, ihr Ziel erreichen zu können.

Global gesehen stiegen die Erfolgserwartungen zwischen 16 und 23 Jahren an, jedoch gab es Jugendliche, die von diesem Entwicklungsverlauf deutlich abwichen. Die Ergebnisse von Growth Mixture Modeling zeigten vier Typen von Entwicklungsverläufen: Hohe und ansteigende Erfolgserwartungen (78%), von geringem Ausgangsniveau ansteigende Erfolgserwartungen (9%), absinkende Erfolgserwartungen (6%) und einen U-förmigen Entwicklungsverlauf (7%). Jugendliche mit hohen und ansteigenden Erfolgserwartungen waren mit höherer Wahrscheinlichkeit männlich, kamen aus einer Familie mit hohem sozioökonomischem Status, bewerteten berufliche Ziele als besonders wichtig und hatten hohe Unterstützung durch ihre Eltern. In Bezug auf Konsequenzen von berufszielbezogenen Erfolgserwartungen zeigte sich, dass Jugendliche mit absinkenden Erfolgserwartungen im Alter von 23 Jahren in einer ungünstigeren beruflichen Situation waren als andere Jugendliche.

Die Ergebnisse der Studie deuten auf die Heterogenität in der Entwicklung angemessenen Engagements in der Berufswahl hin. Individuell unterschiedliche Entwicklungsverläufe, ihre Ursachen und Konsequenzen in die Tiefe zu verstehen, bietet wertvolle Ansatzpunkte für individualisierte Interventionen zur Berufswahl.



## ***Berufswahlverhalten Jugendlicher – Bedeutung von Selbstwirksamkeit und Ergebniserwartung für die Aktivitäten im Berufswahlprozess***

Selbstwirksamkeit beschreibt nach Bandura (1997) die Einstellung einer Person gegenüber ihren Fähigkeiten, eine bestimmte Handlung auszuführen. Selbstwirksamkeit beeinflusst dabei das Handeln und die Motivation sowie die Anstrengung und Ausdauer bei einer Problembewältigung. Die sozial-kognitive Laufbahtheorie (SCCT) von Lent, Brown und Hackett (1994, 2002) besagt eine Beeinflussung von Selbstwirksamkeit und Ergebniserwartung auf das Verhalten bei der Berufswahlentscheidung. In zwei Längsschnittstudien wurde demgemäß geprüft werden, inwieweit eine Wirkung von Selbstwirksamkeit und Ergebniserwartung auf die Bereitschaft und Aktivität Jugendlicher im Berufswahlprozess (z.B. Informations- und Explorationsbereitschaft) erfolgt. In der Auswertung werden zwei Datensätze verwendet; befragt wurden Jugendlichen aus der Sekundarstufe I (n=697) an drei Messzeitpunkten und Jugendliche des Übergangssystems (n=188) an zwei Messzeitpunkten.

In beiden Untersuchungen wurde ein Fragebogen zur Berufswahlkompetenz nach dem Konzept von Ratschinski (2008, 2012) eingesetzt.

Berufswahlkompetenz beschreibt danach die eigenverantwortliche, selbstbestimmte und eigenaktive Fähigkeit zur Berufswahl, inklusive der empfundenen und erklärten Zuständigkeit zu den berufswahlbezogenen Entwicklungsaufgaben und deren Bewältigung. Eine Berufswahl ist umso erfolgreicher (bzw. erfolgsversprechender), wenn sie eigenverantwortlich und von sich aus gestaltet wird.

Die Zusammenhänge zwischen den Variablen wurden anhand von Regressionen und Pfadmodellen analysiert. Die Pfadanalysen zeigen den Einfluss von Selbstwirksamkeit und Ergebniserwartung auf die Aktivitäten im Berufswahlprozess. Pfadanalysen bieten den Vorteil, die Stärke der standardisierten Regressionskoeffizienten sowie der Anteil erklärter Varianz der abhängigen Variablen anzugeben und geben zusätzlich Informationen zur Modell-Güte. Durch die vorliegenden Längsschnittdaten wurden im Pfadmodell auch reziproke Beziehungen zwischen den Wirksamkeitsüberzeugungen und den Berufswahlaktivitäten überprüft. Gleichzeitig wurden Erkenntnisse gewonnen, über welchen Zeitraum die Effekte bestehen und inwieweit sich Schulform-, Klassen und Geschlechterunterschiede abbilden.

Die Berufswahlaktivitäten führen zur Erhöhung des Wissens über Berufe und zur Festigung der Entschiedenheit der eigenen Berufswahl.

## ***Berufswahlprozesse von Mädchen und Jungen in der gymnasialen Mittelstufe – Theoretische Überlegungen, empirische Befunde sowie Perspektiven für Forschung und Intervention***

Berufsorientierung spielte am Gymnasium lange Zeit eine marginale Rolle in der Organisation schulischer Bildungsprozesse. Dafür gibt zum einen historische Gründe, die in der Institution bzw. auf der strukturellen Ebene zu finden sind. Zum anderen wurde unterstellt, dass Schülerinnen und Schüler des Gymnasiums als scheinbar privilegierte Gruppe von Berufswählern keine berufliche Orientierung benötigen. Studien haben jedoch gezeigt, dass insbesondere Jugendliche mit Hochschulzugangsberechtigung am wenigsten planvoll ihren Berufswahlprozess gestalten (vgl. z.B. Prager & Wieland, 2005). Die Vielfalt von Optionen und die späte Auseinandersetzung mit der Berufswahl führen häufig in berufliche Entscheidungswege, die später in Frage gestellt werden. Davon zeugen nicht zuletzt die hohen Studienabbruchquoten (vgl. z.B. Heublein et al., 2010).

Studien weisen darauf hin, dass es einer individuellen Ausrichtung der Berufs- und Studienorientierung bedarf, die eine kontinuierliche Begleitung aller Mädchen und Jungen, unabhängig vom angestrebten Bildungsabschluss, ermöglicht. Die Angebote zur Berufs- und Studienorientierung sind zwar mittlerweile auch am Gymnasium vorhanden, gehen jedoch häufig an den Bedürfnissen der Adressaten vorbei (vgl. z.B. Schmidt-Koddenberg & Zorn, 2011; Hany & Driesel-Lange, 2006).

Die Vielfalt, die sich insgesamt in der Praxis der Berufsorientierung zeigt, kann jedoch für eine theoretisch-wissenschaftliche Reflexion nicht gelten. Sowohl in der theoretischen Begründung als auch in der empirischen Prüfung der Wirksamkeit pädagogischen Handelns zeigen sich Entwicklungspotentiale, werden die wenigen vorhandenen Studien dieser Einschätzung zugrunde gelegt.

Die Bereitstellung individueller pädagogischer Unterstützung setzt Erkenntnisse zu Entwicklungsständen und Entwicklungsverläufen berufswahlbezogener Einstellungen und Kompetenzen von Mädchen und Jungen voraus. Diese Befunde waren Ziel einer längsschnittlichen Untersuchung zur beruflichen Entwicklung von Mädchen und Jungen der gymnasialen Mittelstufe. Mittels Fragebogen wurden in einer ersten Teilstudie 174 Schülerinnen und Schüler befragt. Die Ergebnisse zeigen die Veränderungen über zwei bzw. drei Messzeitpunkte, die sich im Rahmen der schulischen Berufsorientierung ohne gezielte Interventionen ergeben. Neben der Betrachtung der Niveauperänderungen von berufswahlbezogenem Wissen und Einstellungen werden auch Ergebnisse der Veränderung beruflicher Interessen und Fähigkeitszuschreibungen differenziert nach Geschlecht dargestellt. Die Auswertungen zeigen stark geschlechtsstereotype Zuordnungen von Jungen und Mädchen, sowohl im Bereich der Interessen als auch der Fähigkeiten, die dem eigenen und dem anderen Geschlecht zugeschrieben werden. Die Befunde geben Anlass zu einer vertieften Diskussion über mögliche Interventionsansätze und auch –zeitpunkte im Kontext theoretischer Reflexion, um scheinbar unveränderliche Stereotypen zu hinterfragen und aufzubrechen.

Die theoretische und empirische Begründung berufsorientierender Aktivitäten ist eine wichtige aber nicht hinreichende Maßnahme im Kontext der pädagogischen Begleitung beruflicher Entwicklungsprozesse. Zudem müssen bestehende Interventionen in ihrer Wirksamkeit überprüft werden. In einer zweiten Teilstudie stand die Überprüfung der Effekte der Maßnahmen, die über zwei Schuljahre an fünf Schulen in der gymnasialen Mittelstufe durchgeführt wurden, im Fokus der Untersuchung. In einem Experimental-Kontrollgruppendesign wurde 309 Schülerinnen und Schüler zu drei Messzeitpunkte ein Fragebogen zu berufswahlbezogenen Einstellungen und beruflichen Interessen und Fähigkeitseinschätzungen vorgelegt. Das zentrale Ergebnis der Evaluation ist, dass sich Berufswahlkompetenz nicht mit punktuellen Angeboten fördern lässt, die losgelöst vom Unterricht und sonstigen berufsorientierenden Aktivitäten sind. Besonders die Veränderung beruflicher Interessen, die nicht stereotypengeleitet sind, erfordern eine frühzeitige und kontinuierliche pädagogische Arbeit, die nicht nur Schülerinnen und Schüler adressiert, sondern auch Lehrkräfte und Eltern gleichermaßen in die Arbeit einbezieht.

Die dokumentierten Interventionen zur Förderung der Berufswahlkompetenz bzw. von einzelnen Komponenten dieser sowie die beiden vorliegenden Studien zeigen jedoch wichtige Ansatzpunkte, die stärker in die wissenschaftliche und praktische Auseinandersetzung mit Berufsorientierung einfließen sollte.



## **Thüringer Berufsorientierungsmodell (ThüBOM): Befunde zur Wirksamkeit einer modellbasierten Interventionsmaßnahme**

### Theoretischer Hintergrund:

Der Übergang von der Schule in den Beruf ist eine zentrale Entwicklungsaufgabe im Jugendalter, die von der Schule durch gezielte pädagogische Angebote zur beruflichen Orientierung unterstützt werden soll. Um die Heranwachsenden im beruflichen Explorationsprozess individualisiert begleiten und fördern zu können, bedarf es geeigneter Instrumente, um Aussagen zum Entwicklungsstand des Einzelnen im Berufswahlprozess ableiten zu können. Hierfür wurde im ThüBOM-Forschungsprojekt (FSU Jena und Universität Erfurt) ein Kompetenzmodell zur Berufswahl entwickelt, das als Rahmenmodell für die Konstruktion von diagnostischen Instrumenten dient.

Das Berufswahlkompetenzmodell beschreibt vier Phasen, die Heranwachsende durchlaufen, um am Ende der Schulzeit eine begründete Berufswahlentscheidung treffen zu können. Drei Kompetenzdimensionen liegen dem Modell zugrunde - Wissen, Motivation und Handlung. Ein Ziel des Forschungsprojektes ist die Förderung dieser Berufswahlkompetenzen. Mit dem eingereichten Beitrag wird eine neu konzipierte Interventionsmaßnahme zur Förderung der Motivation im Berufswahlprozess vorgestellt und im Hinblick auf deren Wirksamkeit diskutiert. Nach dem Thüringer Berufswahlkompetenzmodell werden innerhalb der Dimension Motivation verschiedene motivationale Facetten unterschieden: Betroffenheit, Eigenverantwortung, Offenheit und Zuversicht.

Die Motivation, sich mit den eigenen beruflichen Plänen auseinanderzusetzen, gilt als essenzielle Voraussetzung für die schulische Berufsorientierung, stellt jedoch für die Heranwachsenden eine große Herausforderung dar (Driesel-Lange et al., 2010). Obwohl eine Vielzahl an schulischen Aktivitäten zur Berufsorientierung angeboten wird, ist die Wirkung vieler Maßnahmen unklar und die Motivation der Schülerinnen und Schüler sich mit dem Thema Berufswahl auseinanderzusetzen, vergleichsweise gering. Der Beitrag greift das Thema Motivation im Berufswahlprozess auf und erläutert mit Bezug auf die Selbstbestimmungstheorie von Deci & Ryan (2002), wie berufsorientierende Maßnahmen in der Schule motivationsförderlich gestaltet werden können.

### Fragestellung:

Im Rahmen des Beitrags wird der Frage nachgegangen, ob die vorgestellte Interventionsmaßnahme die Motivation der Schülerinnen und Schüler im Berufswahlprozess erhöht.

### Methode:

Um die Wirkung der Interventionsmaßnahme zu überprüfen, wurde ein quasi-experimentelles Design mit Experimental- und Kontrollgruppe gewählt. 50 Schüler aus zwei siebten Klassen einer integrierten Gesamtschule wurden zu drei Messzeitpunkten (Prätest – Posttest – Follow up) zu ihren motivationalen Einstellungen befragt.

### Ergebnisse:

Statistische Analysen der ersten beiden Messzeitpunkte zeigen systematisch positive Veränderungen auf nahezu allen Motivationsfacetten. Die Effekte liegen im mittleren Bereich. Die Analyse der mittelfristigen Effekte (Messzeitpunkt 3) weist tendenziell auf die zeitlich stabile Wirkung der Interventionsmaßnahme hin. Die Ergebnisse korrespondieren mit den subjektiven Evaluationsbefunden, bei denen die Mehrheit der befragten Schüler angegeben hat, dass die Intervention vor allem Betroffenheit hergestellt hat, sich mit der eigenen Berufswahl zu beschäftigen. Die Ergebnisse zeigen, dass eine wirksame Interventionsmaßnahme zur Förderung der Berufswahlkompetenz entwickelt wurde. Dies ist insofern von Bedeutung, da im Bereich Berufsorientierung zwar eine Vielzahl von Fördermaßnahmen existiert, aber nur sehr wenige systematisch evaluiert wurden.



## **Studiengangsspezifische Aufgaben und Anforderungen von Studierenden und Absolvent(inn)en im Spannungsfeld von Hochschule und Arbeitsmarkt**

Chair(s): Edith Braun, Julia-Carolin Brachem (beide: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

Diskutant/in: Johannes Hartig (Deutsches Institut für Internationale Pädagogische Forschung)

Lernergebnisse der Hochschule, in Form erworbener Kompetenzen, erhalten nicht zuletzt aufgrund des Bologna-Prozesses vermehrt Aufmerksamkeit. Die Förderung von Kompetenzen, die in der späteren Erwerbstätigkeit einsetzbar sind, wird zum Qualitätsmerkmal hochschulischer Lehre. Um dem gerecht werden zu können, ist es bedeutsam, die Anforderungen von Curricula und Arbeitsmarkt zu kennen und Verfahren zur Kompetenzerfassung zu entwickeln. Allerdings steht die Entwicklung von Kompetenztests in der tertiären Bildung erst am Anfang. Daher werden in diesem Symposium beispielhaft aus den Bereichen des Informatik-Lehramts, der Wirtschaftswissenschaften und der Erziehungswissenschaften Projekte vorgestellt, in denen Modelle und Verfahren zur Erfassung von Kompetenzen entwickelt werden.

Der Beitrag von Bender et al. konzeptionalisiert ein fachdidaktisches Kompetenzmodell für Lehrkräfte des Informatikunterrichts. Zunächst wurde aufgrund vorliegender Arbeiten ein Kompetenzstrukturmodell entwickelt. Die dort enthaltenen Kategorien wurden in einem zweiten Schritt mit universitären Curricula der Informatiklehrmatsausbildung verglichen, um festzustellen, inwieweit sich die identifizierten Bereiche in der hochschulischen Lehre wiederfinden lassen. In Interviews wurden Hinweise zu Aufgaben und Anforderungen von Informatiklehrkräften gesammelt. Aufbauend auf dieser Arbeit wurden schließlich Messitems entwickelt, um fachdidaktische Kompetenzen erfassen zu können.

Die Arbeitsgruppe um Zlatkin-Troitschanskaia entwickelt einen Test zur Erfassung von wirtschaftswissenschaftlicher Fachkompetenz. In diesem Beitrag widmet sie sich der grundlegenden Frage, ob sich die zu fördernden Kompetenzen aus dem Beschäftigungssystem oder aus den hochschulischen Curricula ableiten sollen. Dies ist eine Frage der Inhaltsvalidität, deren Ergebnisse und Konsequenzen für den Test im Symposium vorgestellt werden sollen.

In den beiden letzten Beiträgen handelt es sich um das Kooperationsprojekt KomPaed, das zum Ziel hat, für erziehungswissenschaftliche Hochschulabsolvent(inn)en überfachliche Aufgaben, Tätigkeiten und Anforderungen auf dem Arbeitsmarkt zu identifizieren, um daraus perspektivisch generische Kompetenzmodelle zu entwickeln. Das Hamburger Projekt hat dazu qualitative Interviews mit Arbeitgeber(inne)n geführt. Anhand von Valenzanalysen werden die Aussagen dahingehend ausgewertet, welche Anforderungen an Absolvent(inn)en aus den Erziehungswissenschaften auf dem Arbeitsmarkt gestellt werden und wie wichtig bestimmte Fähigkeiten eingeschätzt werden. Für Prinz et al. steht die Berücksichtigung von Qualitätskriterien qualitativer Forschung mit im Mittelpunkt. Im Hannoveraner Projekt wurden vor dem Hintergrund der Ergebnisse eben genannter Arbeitgeberinterviews und umfangreicher internationaler Literaturrecherche acht berufliche Anforderungsbereiche erarbeitet, zu denen jeweils verhaltensnahe Items entwickelt wurden, um perspektivisch generische Kompetenzen zu erfassen.

Johannes Hartig wird in dem Symposium die Bedeutung von zunächst rein konzeptionellen Kompetenzmodellen übergreifend diskutieren. Auch Synergieeffekte zwischen qualitativen und quantitativen Vorgehensweisen werden offensichtlich. Darüber hinaus ist der Bezugsrahmen (hochschulische Curricula versus Arbeitsmarkt) eine generell zu klärende Frage, wenn es um die Validität von Instrumenten zur Erfassung von Kompetenzen geht.

Elena Bender (Universität Paderborn), Niclas Schaper (Universität Paderborn), Peter Hubwieser (TU München), Johannes Magenheimer (Universität Paderborn), Sigrid Schubert (Universität Siegen), Marc Berges (TU München), Melanie Margaritis (Universität Paderborn), Laura Ohrndorf (Universität Siegen)

## **Modellierung fachdidaktischer Kompetenzen von Lehrkräften für den Informatikunterricht**

Zur Implementierung eines inhaltlich und didaktisch anspruchsvollen sowie lernförderlichen und schülerzentrierten Informatikunterrichts werden Lehrkräfte benötigt, die nicht nur fachwissenschaftlich, sondern auch fachdidaktisch und pädagogisch wirkungsvoll ausgebildet werden. Die Ausbildungscurricula für Lehramtsstudierende zum Unterrichtsfach Informatik weisen allerdings nicht nur fachwissenschaftlich, sondern auch fachdidaktisch sehr unterschiedliche Ausrichtungen auf. Hier existieren eine Vielzahl möglicher didaktischer Ansätze, Zielsetzungen, Inhaltskanons und Implementierungsformen für den Informatikunterricht an der Schule. Dieser unübersichtlichen und uneinheitlichen Situation entsprechend gibt es bisher nur sehr wenige allgemeine und übertragbare Forschungsergebnisse zum Wissen bzw. zu den Kompetenzen, die für die Gestaltung eines erfolgreichen Informatikunterrichts notwendig sind. Mühling et al. (2010) fanden im Rahmen einer umfassenden Umfrage unter den bayerischen Informatiklehrkräften, dass das Fachwissen der Lehrkräfte gemäß deren Selbsteinschätzung insbesondere in den Bereichen Softwareentwicklung und Projektmanagement Defizite aufweist. Zudem ergab eine Clusteranalyse, dass Lehrkräfte mit besonderen Defiziten in diesen Bereichen mit der eigenen Unterrichtsplanung und ihrem eigenen Unterrichtserfolg nicht zufrieden sind. Andererseits zeigen Studien zum Einfluss von Kompetenzen der Lehrkräfte im Mathematikunterricht, dass fachdidaktisches Wissen eine entscheidende Rolle in Bezug auf die Qualität des Unterrichts spielt (Krauss et al., 2004). Vor allem Lehrerinnen und Lehrer mit einem hohen fachdidaktischen Wissen gestalten auch qualitativ volleren Unterricht. Das heißt, sie legen mehr Wert auf Verständnisprozesse und geben Schülerinnen und Schülern mehr persönliche Lernunterstützung. Vor diesem Hintergrund wurde ein Ansatz zur Modellierung und Erfassung insbesondere der fachdidaktischen und einstellungsbezogenen Kompetenzen von Lehrkräften für den Informatikunterricht entwickelt. Auf der Grundlage eines Kompetenzverständnisses, dass sich einerseits auf die verbreitete Auffassung im Rahmen der empirischen Bildungsforschung nach Weinert (2001) und andererseits auf die Auffassung der kognitionspsychologisch orientierten Lehrerbildungsforschung (vgl. Blömeke et al., 2010) bezieht, wurden folgende Schritte bei der Konzeption eines fachdidaktischen Kompetenzmodells vollzogen: Zunächst wurde ein Rahmenmodell zur Identifizierung relevanter Dimensionen und Kategorien eines fachdidaktischen Kompetenzstrukturmodells für Informatiklehrkräfte auf der Basis von Analysen einschlägiger Vorarbeiten (z. B. Blömeke et al., 2010, Hubwieser et al., 2011 oder Schaper, 2009) erarbeitet. Hierbei wurden drei Aufgabenfelder (Unterrichtsplanung, Unterrichtsdurchführung, Unterrichtsevaluation) und 15 Facetten des fachdidaktischen Wissens (z. B. Einsatz von Lehrmethoden, Berücksichtigung von Schülerkognitionen) identifiziert und beschrieben (Hubwieser et al., 2013). In einem zweiten Schritt wurde überprüft, inwieweit die Kategorien des Rahmenmodells in Curricula der Informatiklehrer- und Lehrerbildung abgebildet sind. Damit wurde überprüft, ob an deutschen Universitäten entsprechende fachdidaktische Kompetenz- und Wissensbereiche Gegenstand der universitären Ausbildung sind. Es zeigte sich, dass die identifizierten Kompetenzkategorien in vielen Bereichen nur teilweise von den universitären Ausbildungscurricula abgedeckt werden. In einem dritten Schritt wurden Experteninterviews auf der Grundlage der Critical Incident Technik geführt. Bei dieser Interviewmethode werden die Befragten zu typischen und besonders anspruchsvollen Aufgaben und Anforderungen des informatischen Unterrichtens (auf der Basis des Modells und Expertenvorbefragungen konzipierten authentischen Unterrichtsszenarien) befragt, um die erforderlichen Denk- und Handlungsweisen sowie weitere Kompetenzanforderungen zur Aufgabenbewältigung in den entsprechenden Anforderungssituationen zu ermitteln. Auf dieser Grundlage konnte das auf der Basis von theoretisch-konzeptionellen und normativen Aspekten konzipierte Rahmenmodell anhand empirischer, erfahrungsbasierter Daten weiter ausdifferenziert und in einem ersten Schritt inhaltlich validiert werden. Das anhand dieses Vorgehens entwickelte Modell dient als Grundlage zur Konstruktion eines Kompetenzmessinstruments zur Erfassung fachdidaktischer Lehrkompetenzen bei Informatiklehrkräften. Es wird gezeigt, wie auf der Basis des Modells die Messitems entwickelt werden.

- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010). TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich. Münster: Waxmann.
- Hubwieser, P., Armoni, M., Brinda, T., Dagiene, V., Diethelm, I., Giannakos, M. N., Knobelsdorf, M., Magenheimer, J., Mittermeier, R. T., and Schubert, S. E. (2011). Computer science/informatics in secondary education. In Proceedings of the 16th annual conference reports on Innovation and technology in computer science education - working group reports. ITCSE-WGR '11. ACM, New York, NY, USA, 19-38.
- Hubwieser, P. et al. (2013). Pedagogical Content Knowledge for Computer Science in German Teacher Education Curricula. WIPsCE 2013.
- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M., Jordan, A. & Löwen, K. (2004). COACTIV. Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In Prenzel, M. & Doll, J. (Eds.), Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Waxmann.
- Mühling, A., Hubwieser, P., Brinda, T. (2010). Exploring Teachers' Attitudes Towards Object Oriented Modelling and Programming in Secondary Schools. In ACM (Ed.), ICER '10: Proceedings of the Sixth international workshop on Computing education research, pp. 59–68. ACM, New York, NY, USA.
- Schaper, N. (2009). Aufgabenfelder und Perspektiven bei der Kompetenzmodellierung und -messung in der Lehrerbildung. Lehrerbildung auf dem Prüfstand, 2 (1), 166-199.
- Weinert, F. E. (2001). Concept of Competence: A conceptual clarification. In Defining and Selecting Key Competencies, D. S. Rychen and L. Salganik, Eds. Hogrefe and Huber, Seattle.





Sebastian Brückner (Johannes Gutenberg-Universität Mainz), Manuel Förster (Johannes Gutenberg-Universität Mainz), Oliver Lauterbach (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung), Olga Zlatkin-Troitschanskaia (Johannes Gutenberg-Universität Mainz)

## ***Im Spannungsfeld von Anforderungen des Arbeitsmarktes und curricularer Validität – inhaltliche Validierung eines Instruments zur Messung betriebswirtschaftlichen Wissens***

Die wirtschaftswissenschaftlichen Studienbereiche zählen zu den meistgewählten Studienfächern in Deutschland. Trotz dieser Bedeutung liegt bislang für diese Studiendomäne kein deutschsprachiges Instrument vor, das den messmethodischen Anforderungen an die Erfassung wirtschaftswissenschaftlicher (Fach)Kompetenz auf akademischem Niveau gerecht wird (Kuhn & Zlatkin-Troitschanskaia, 2011). Zwar finden sich insbesondere für den Sek II Bereich einige Testinstrumente wie den Wirtschaftslehretest (Krumm & Seidel, 1970) und den Wirtschaftskundlichen Bildungstest (Beck et al., 1998), welche mitunter auch bereits einer umfassenden Evaluierung ihrer Einsatzmöglichkeiten im Hochschulbereich unterzogen wurden (Förster et al., 2012). Jedoch zeigt sich, dass unter Berücksichtigung der Vielfalt wirtschaftswissenschaftlicher Studienfacetten und den beruflichen Anforderungen in z. T. sehr heterogenen wirtschaftswissenschaftlichen Studiengebieten (Lauterbach, 2012) umfassendere Testinstrumente benötigt werden. Das vom BMBF geförderte internationale Projekt WiwiKom knüpft an diesem Defizit an und verfolgt iterativ zwei Ziele. (1) Entwicklung eines inhaltlich und curricular validen Modells einer kognitiv orientierten wirtschaftswissenschaftlichen Kompetenz; (2) Adaption zweier internationaler Testinstrumente (Test of Understanding in College Economics (TUCE), Walstad et al., 2007; Examen general para el Egreso de la Licenciatura (EGEL), CENEVAL, 2011a) und Konstruktion eines deutschen Instruments, sodass die theoretisch modellierten Strukturen und Niveaus der Fachkompetenz messbar werden. Um diese Ziele zu erreichen, ist sicherzustellen, dass die eingesetzten Testinstrumente inhaltlich valide sind, da nur hierdurch valide Testwertinterpretationen auf die zugrundeliegenden Wissensstrukturen der Studierenden nachvollziehbar und umfassend begründbar werden (Kane, 2013; Geisinger et al., 2012). Vor dem Hintergrund des Spannungsfeldes zwischen universitären, curricularen Anforderungen und berufsbezogenen Erwartungen potentieller Arbeitgeber muss das Kompetenzkonstrukt inhaltlich genau beschrieben und der Inhalt der Testaufgaben daraufhin geprüft werden, ob mit ihnen konstruktrelevante und repräsentative Inhalte abgebildet werden. Als notwendige Voraussetzung für eine Beurteilung der weiteren Verwendbarkeit der Aufgaben rückt daher die inhaltliche und curriculare Validität in den Fokus. Der EGEL wurde mit der Zielstellung entwickelt, die beruflichen Anforderungssituationen von Studierenden der Wirtschaftswissenschaften möglichst repräsentativ und vielfältig abzubilden, wobei sowohl Unternehmen als auch Wirtschaftswissenschaftler bei der Entwicklung mitgewirkt haben (CENEVAL, 2011b). Der EGEL umfasst in seiner Originalversion Items aus den Bereichen Marketing, Personal, Finanzierung und Unternehmensführung/Organisation, Buchführung, Rechnungslegung und Kostenrechnung. Die Items sind i. d. R. in handlungsorientierte Szenarien eingebettet und zeichnen sich durch unterschiedliche Aufgabenformate aus. Wegen der Fokussierung auf die beruflichen Anforderungssituationen wurden die hochschulcurricularen Aspekte dieses Tests bislang noch keiner weiteren Analyse unterzogen, sodass noch offen ist, inwieweit die erfassten Inhalte repräsentativ für die deutschen Hochschulcurricula sind. Im Vortrag werden daher die Ergebnisse aus verschiedenen Facetten der Inhaltsvalidierung der aus dem EGEL entwickelten BWL-Items in WiwiKom präsentiert, miteinander verknüpft und der daraus resultierende Itemauswahl- und Adaptationsprozess begründet dargestellt. In Kooperation mit Etappe 7 des Nationalen Bildungspanels („Hochschulstudium und Übergang in den Beruf“) wurden Analysen der Modulhandbücher wirtschaftswissenschaftlicher Studiengänge (N = 96) zur Beschreibung deren fachinhaltlicher Struktur durchgeführt. Weiterhin wurden die aus dem TUCE und EGEL entwickelten Aufgaben von Hochschuldozierenden (N = 85) hinsichtlich ihrer Eignung bewertet und die Bedeutung einzelner wirtschaftswissenschaftlicher Inhaltsbereiche bundesweit erfasst. Im Rahmen der Analysen der Modulhandbücher wurde betrachtet, ob die Testinhalte des EGEL mit den intendierten Curricula der deutschen Hochschullandschaft übereinstimmen. Mit Hilfe des Online-Tools konnten die Experten mittels geschlossener Ratingitems den curricularen Bezug und die Schwierigkeit der Aufgaben bewerten sowie ein Globalurteil zu jedem Item abgeben. In einem offenen Feld konnten die Dozenten ein Feedback zu den Aufgaben geben. Ziel der Analysen war es, die curriculare Übereinstimmung von Aufgaben- und Studieninhalten zu betrachten und hinsichtlich der curricularen Validität zu interpretieren sowie eine inhaltliche Validierung durch Expertenrückmeldungen vorzunehmen. Bei kritischen Items wurden dann im Anschluss immer wieder Expertengespräche zur Verbesserung der Aufgaben geführt. Auf Basis dieser methodenintegrativen Verfahrenswahl konnten sowohl qualitative als auch quantitative Daten gewonnen werden, die der Itemauswahl für den Pretest und der ersten Erhebung im WS 2012/2013 zu Grunde lagen.

- Beck, K., Krumm, V. & Dubs, R. (1998). Wirtschaftskundlicher Bildungs-Test (WBT). Handanweisung. Göttingen: Hogrefe.
- CENEVAL (2011a). Exámenes Generales de Egreso de la Licenciatura, EGEL®, Administración and Contaduría. Centro Nacional de Evaluación para la Educación Superior (CENEVAL), Mexico, D.F.
- CENEVAL (2011b). EGEL Administration. Präsentation zum Workshop "Kompetenzmodellierung in der Betriebs- und Volkswirtschaftslehre", Mexico City.
- Förster, M., Happ, R. & Zlatkin-Troitschanskaia, O. (2012). Valide Erfassung des volkswirtschaftlichen Fachwissens von Studierenden der Wirtschaftswissenschaften und der Wirtschaftspädagogik – eine Untersuchung der diagnostischen Eignung des Wirtschaftskundlichen Bildungstests (WBT). Zeitschrift bwp@ Berufs- und Wirtschaftspädagogik - online, Ausgabe 22, 1-21. Verfügbar unter: [http://www.bwpat.de/ausgabe22/foerster\\_etal\\_bwpat22.pdf](http://www.bwpat.de/ausgabe22/foerster_etal_bwpat22.pdf) [23.08.2013].
- Geisinger, K., Shaw, L. H. & McConnick, C. (2013). The validation of tests in higher education. In C. Secolsky & D. B. Denison (Eds.), Handbook on measurement, assessment and evaluation in higher education (pp. 194 – 207). Marceline: Walsworth.
- Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. Journal of Educational Measurement, 50(1), 1–73.
- Krumm, V. & Seidel, G. (1970): Wirtschaftslehretest BWL. Weinheim & Basel.
- Kuhn, C. & Zlatkin-Troitschanskaia, O. (2011). Assessment of Competencies among University Students and Graduates – Analyzing the State of Research and Perspectives. Johannes Gutenberg University Mainz. Arbeitspapiere Wirtschaftspädagogik, 59.
- Lauterbach, O. (2012). Curriculare Validität internationaler Testinstrumente zur Erfassung wirtschaftswissenschaftlicher Fachkompetenz. Vortrag am 11.09.2012 auf der Tagung der Arbeitsgruppe Empirische Pädagogische Forschung (AEPF), Bielefeld.
- Walstad, W. B., Watts, M. & Rebeck, K. (2007). Test of understanding in college economics: Examiner's manual (4th ed.). New York NY: National Council on Economic Education.



## **Valenzanalysen im BMBF Projekt KomPaed: Ein Beitrag zur Diskussion über Gütekriterien in qualitativen Untersuchungen**

Eine universitäre Ausbildung ist im Wesentlichen durch ihren Wissenschaftsbezug gekennzeichnet. Im Zuge des Bologna-Prozesses und der Einführung der gestuften Studienstruktur wurde jedoch die seit längerem geführte Debatte über die Berufs- und Praxisrelevanz eines Studiums intensiviert. Von den Hochschulen wird erwartet, „Employability“ – häufig übersetzt als Beschäftigungsfähigkeit – zu fördern. Es ist Konsens, dass generische Kompetenzen für Beschäftigungsfähigkeit zentral sind und sich ihre Bedeutung u. a. aus dem notwendig flexiblen Verhältnis von Studium und Beruf begründen lässt. Diesem hohen Stellenwert entsprechend fokussiert das Projekt KomPaed (Tätigkeitsbezogene Kompetenzen in pädagogischen Handlungsfeldern) auf generische Kompetenzen, die für erfolgreiches berufliches Handeln hoch qualifizierter Fachkräfte in Verbindung mit Fachwissen und fachlichen Kompetenzen erforderlich sind. Das besondere Augenmerk liegt dabei auf Absolventinnen und Absolventen erziehungswissenschaftlicher Studiengänge. Für das vom BMBF finanzierte Kooperationsprojekt der Universität Hamburg mit dem Deutschen Zentrum für Hochschul- und Wissenschaftsforschung – DZHW (vormals HIS-HF) leiten sich daraus zwei Kernfragestellungen ab: (1) Lassen sich für akademische Berufe disziplinübergreifende nichtfachliche Aufgaben und Anforderungen identifizieren? Und lassen sich aus diesen (2) disziplinübergreifende überfachliche Kompetenzen und Kompetenzmodelle ableiten? Zur Beantwortung der ersten Fragestellung wurde eine explorative Befragung von relevanten Arbeitsplatzanbietern für Absolventinnen und Absolventen erziehungswissenschaftlicher Studiengänge, im Sinne einer Arbeitnehmerbefragung, durchgeführt. Mit Hilfe von Leitfadenterviews wurden Personen befragt, die mit der Auswahl von Personal in pädagogischen Arbeitsfeldern betraut sind. In den 18 qualitativen Interviews ging es vor allem um die Aufgaben und Anforderungen, die Absolventinnen und Absolventen erziehungswissenschaftlicher Studiengänge in ihren Berufen erfüllen müssen und um die erforderlichen Kompetenzen, die daraus abgeleitet werden können. Die Auswahl einschlägiger sozialer Einrichtungen und damit des Samples der Interviewpartnerinnen und Interviewpartner erfolgte nach theoretischen (vgl. Fuchs, 2003) und empirischen Kriterien in Zusammenarbeit mit dem Kooperationspartner DZHW. Die aus den Expertinnen- bzw. Experteninterviews gewonnenen Daten wurden in einer ersten Auswertungsphase einer quantitativen Valenzanalyse unterzogen. Dieses inhaltsanalytische Verfahren (vgl. Mayring, 2010) dient der Bestimmung der Häufigkeiten von Textelementen und deren Bewertung aufgrund des Kontextes, die mit dem Untersuchungsgegenstand verbunden ist (vgl. Mayring, 2010, Schnell et al., 2011). Folglich wird sich im Ergebnis eine erste Tendenz feststellen lassen, welche Tätigkeiten und Anforderungen in verschiedenen pädagogischen Handlungsfeldern von Arbeitgeberinnen und Arbeitgebern als besonders bedeutsam erachtet werden. Der Beitrag fokussiert den Bericht dieser Ergebnisse. Im Anschluss an diese erste Systematisierung des qualitativen Materials erfolgen, in Anlehnung an die Grundzüge der Grounded Theory, vertiefende Analysen zur Typisierung unterschiedlicher Anforderungsprofile, die in ihrer Anlage abschließend skizziert werden sollen. Im Hinblick auf die aktuelle Diskussion um Kriterien der Qualitätssicherung in qualitativen Studien erfolgt vor dem Hintergrund der Anlage der Untersuchung eine Auseinandersetzung mit unterschiedlichen Validierungsstrategien. Die Frage der Validitätssicherung spielt bereits bei der Erhebung und Dokumentation der Experteninterviews, aber auch bei der Auswertung und Interpretation eine zentrale Rolle. Diesem wird im Rahmen des Kooperationsprojektes Rechnung getragen, indem u. a. Ergebnisse quantitativer Sekundäranalysen der DZHW (vormals HIS-HF) Absolventendaten für die Valenzanalysen herangezogen werden.

Fuchs, K. (2003). Wissen und Tun. Tätigkeitsprofile und berufsrelevante Kenntnisse. In H.-H. Krüger, T. Rauschenbach, K. Fuchs, C. Grunert, A. Huber, B. Kleifgen, P. Rostampour, C. Seeling & I. Züchner (Hrsg.), *Diplom-Pädagogen in Deutschland. Survey 2001* (S. 185-204). Weinheim: Juventa.  
Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. (11. Aufl.). Weinheim: Beltz, S. 15f.  
Schnell, R., Hill, P.B. & Esser, E. (2011). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

## Berufliche Tätigkeiten und Anforderungen erziehungswissenschaftlicher Hochschulabsolvent(inn)en

Mit der Aufnahme eines Studiums von etwa fünfzig Prozent eines Jahrgangs übernehmen Hochschulen einen bedeutenden gesellschaftlichen Auftrag hinsichtlich der Ausbildung für den Arbeitsmarkt. Fachhochschulen und Universitäten müssen, neben der Vermittlung von Fachwissen, vermehrt auch fachliche und überfachliche Kompetenzen fördern, die in der späteren Erwerbstätigkeit einsetzbar sind (Employability) (Kultusministerkonferenz, 2005). Damit eng verbunden ist, dass der neu eingeführte Bachelor-Abschluss berufsqualifizierend sein soll. Vor diesem Hintergrund erscheint es bedeutsam, die Anforderungen des Arbeitsmarktes an Hochschulabsolvent(inn)en zu kennen. Derartige Anforderungen an Erwerbstätige werden in diversen, größtenteils internationalen Studien vor dem Hintergrund des Job Requirements Approach (JRA) erfasst. Mit diesem arbeitspsychologischen Ansatz können Kompetenzen indirekt, abgeleitet von verhaltensnah abgefragten beruflichen Tätigkeiten und Anforderungen, identifiziert werden. Studien, die in diesem Zusammenhang hervorzuheben sind, sind das amerikanische Generalized Work Activities Questionnaire (GWA) (Jeanneret et al., 2002; O\*NET 2012; Peterson et al., 2001), der UK Skills Survey (BMRB, 2006; Felstead et al., 2007), das Programme for the International Assessment of Adult Competencies (PIAAC) der OECD (OECD, 2011; OECD, 2010; OECD, 2009) sowie eine niederländische Anpassung des GWA für Absolvent(inn)en von Berufsfachschulen (Toolsema, 2003). Im deutschen Kontext werden berufsübergreifende Anforderungen des Arbeitsmarktes an Erwerbstätige unter anderem in der BIBB/BAuA Erwerbstätigenbefragung (BIBB/BAuA, 2011; Rohrbach-Schmidt, 2009) und im Nationalen Bildungspanel (Matthes & Christoph, 2011) abgefragt. In Hochschulabsolvent(inn)enbefragungen werden außerdem wichtige Kenntnisse und Fähigkeiten für die berufliche Tätigkeit von Hochschulabsolvent(inn)en erhoben (Hochschul Informations System, 2011, 2010, 2006, 2002; INCHER, 2012; Rehn et al., 2011; Schaeper & Briedis, 2004). In diesem Beitrag soll den Fragen nachgegangen werden, (1) welche beruflichen Tätigkeiten Hochschulabsolvent(inn)en ausüben, (2) welche Anforderungen vom Arbeitsmarkt an sie gestellt werden, (3) welche Kompetenzbereiche für Hochschulabsolvent(inn)en in der beruflichen Praxis bedeutsam sind und, (4) ob es ggf. Unterschiede zwischen den Studienfachrichtungen gibt. Um für deutsche, insbesondere erziehungswissenschaftliche, Hochschulabsolvent(inn)en entsprechende Aussagen treffen zu können, wurden im Rahmen des vom BMBF geförderten Projekts KomPaed (Tätigkeitsbezogene Kompetenzen in pädagogischen Handlungsfeldern) mehrere Arbeitsschritte vollzogen. In einem ersten Schritt wurden vorliegende Fragebögen (GWA, UK Skills Survey, PIAAC-JRA-Module, BIBB/BAuA-Erwerbstätigenbefragung, NEPS Großpilot E8, HIS-Befragungen von Hochschulabsolvent(inn)en) auf enthaltene Items zur Abfrage beruflicher Tätigkeiten und Anforderungen und die dahinterliegenden, empirisch bestätigten Dimensionen überprüft. In einem zweiten Schritt wurde eine Literaturrecherche zum Thema Arbeitsanalyse und pädagogische Handlungsfelder durchgeführt. In einem dritten Schritt wurden von der Universität Hamburg (Kooperationspartner: Schwippert, Prinz, Pfeiffer) 18 qualitative Interviews mit potentiellen Arbeitgebern erziehungswissenschaftlicher Hochschulabsolvent(inn)en durchgeführt. Aufbauend auf dieser Arbeit wurden für Hochschulabsolvent(inn)en acht berufliche Anforderungsbereiche identifiziert (bspw. Aufarbeitung von Informationen; Planung, Entwicklung und Organisation; Soziale Interaktion). Für jede Dimension wurden dann auf der konkreten Handlungsebene und auf Basis des JRA Items entwickelt (bspw. Bei meiner derzeitigen Erwerbstätigkeit erstelle ich Fachtexte in einer anderen Sprache außer Deutsch; Bei meiner derzeitigen Erwerbstätigkeit bewerte ich die Leistung oder Qualität von Prozessen, Objekten oder Personen; Bei meiner derzeitigen Erwerbstätigkeit fördere und motiviere ich andere Personen). Das Befragungsinstrument umfasst 48 Items zu Tätigkeiten und Anforderungen im Berufsleben und wird in einer deutschlandweit repräsentativen Stichprobe von Hochschulabsolvent(inn)en eingesetzt. Dazu kann auf drei Kohorten (Prüfungsjahrgänge 2001, 2005, 2009) des DZHW-Absolventenpanels (vormals HIS) zurückgegriffen werden, die in einer Online-Befragung im Winter 2013/14 befragt werden. Es werden ca. 18.500 Personen zu der Studie eingeladen. In diesem Beitrag sollen die konzeptionellen Dimensionen des Befragungsinstruments vorgestellt werden, die aufgrund der umfassenden Instrumenten- und Literaturrecherche sowie den durchgeführten Interviews erarbeitet wurden. Zudem erwarten wir erste empirische Ergebnisse berichten zu können, ob sich die konzeptionellen Dimensionen empirisch bestätigen lassen und welche Erwartungen und Anforderungen an Hochschulabsolvent(inn)en, ggf. differenziert nach Studienfachrichtung, auf dem Arbeitsmarkt bestehen.

BIBB/BAuA, 2011. BIBB/BAuA-Erwerbstätigenbefragung 2011/2012. Arbeit und Beruf im Wandel, Erwerb und Verwertung beruflicher Qualifikationen. Erhebungsinstrument Fragebogenmaster für die CATI-Programmierung inkl. Variablenerkennung. [http://www.bibb.de/dokumente/pdf/a22\\_ETB\\_2012\\_Fragebogen\\_Endversion.pdf](http://www.bibb.de/dokumente/pdf/a22_ETB_2012_Fragebogen_Endversion.pdf) (06.12.2012). / BMRB Social Research, 2006. 2006 Skills Survey. Technical Report. <http://www.esds.ac.uk/doc/6004/mrdoc/pdf/6004userguide.pdf> (09.10.2013). / Felstead, A., Gallie, D., Green, F. & Zhou, Y., 2007. Skills at Work, 1986-2006. Oxford: Skope. / Hochschul Informations System, 2011. Zehn Jahre nach Studienabschluss. Dritte Befragung der Hochschulabsolventinnen und Hochschulabsolventen des Jahrgangs 2000/2001. Hannover: HIS. / Hochschul Informations System, 2010. Gut gestartet und zufrieden? Zweite Befragung der Hochschulabsolventinnen und Hochschulabsolventen des Jahrgangs 2004/05. Hannover: HIS. / Hochschul Informations System, 2010. Hochqualifiziert und auf dem Weg. Eine Befragung von Hochschulabsolventinnen und Hochschulabsolventen des Prüfungsjahrgangs 2008/2009. Hannover: HIS. / Hochschul Informations System, 2010. Hochqualifiziert und auf dem Weg. Eine Befragung von Masterabsolventinnen und Masterabsolventen des Prüfungsjahrgangs 2008/2009. Hannover: HIS. / Hochschul Informations System, 2006. Zwischen Hochschule und Arbeitsmarkt. Zweite Befragung der Hochschulabsolventinnen und Hochschulabsolventen des Jahrgangs 2000/01. Hannover: HIS. / Hochschul Informations System, 2006. Zwischen Hochschule und Arbeitsmarkt. Eine Befragung von Hochschulabsolventinnen und Hochschulabsolventen des Jahrgangs 2004/2005. Hannover: HIS. / Hochschul Informations System, 2002. Zwischen Hochschule und Arbeitsmarkt. Eine Befragung von Hochschulabsolventinnen und Hochschulabsolventen des Prüfungsjahrgangs 2000/2001. Hannover: HIS. / INCHER-Kassel, 2012. Fragebogen der KOAB-Absolventenbefragung 2012. Erste Befragung des Abschlussjahrgangs 2010 (ca. 1,5 Jahre nach Studienabschluss). Masterfragebogen/Allgemeine Variante. Jahrgang 2010. Mit Variablennamen. [http://koab.uni-kassel.de/images/download/jg10w1\\_fb\\_allgemein.pdf](http://koab.uni-kassel.de/images/download/jg10w1_fb_allgemein.pdf) (14.10.2013). / Jeanneret, P. R., Borman, W. C., Kubisiak, U. C. & Hanson, M. A., 2002. Generalized work activities. In: N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret & E. A. Fleishman (Hrsg.). An occupational information system for the 21st century: The development of O\*NET. Washington, DC: American Psychological Association. / Kultusministerkonferenz, 2005. Qualifikationsrahmen für Deutsche Hochschulabschlüsse. Bonn: KMK. / Matthes, B. & Christoph, B., 2011. Nationales Bildungspanel. Großpilot E8 Feldversion. Version 1.02. / OECD, 2011. PIAAC Conceptual Framework of the Background Questionnaire Main Survey. Paris: OECD. / OECD, 2010. PIAAC Background questionnaire MS version 2.1 d.d. 15-12-2010. Paris: OECD. / OECD, 2009. Outcomes of the JRA Pilot Study: Annex 1. Paris: OECD. / O\*NET, 2012. O\*NET Generalized Work Activities Questionnaire. / [Htt //://www.onetcenter.org/dl\\_files/MS\\_Word/Generalized\\_Work\\_Activities.pdf](http://www.onetcenter.org/dl_files/MS_Word/Generalized_Work_Activities.pdf) (04.12.2012). / Peterson, N. G., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., Campion, M. A., Mayfield, M. S., Morgeson, F. P., Pearlman, K., Gowing, M. K., Silver, M. B. & Dye, D. M., 2001. Understanding work using the Occupational Information Network (O\*NET): Implications for practice and research. Personnel Psychology, 54, 451-492. / - Rehn, T., Brandt, G., Fabian, G. & Briedis, K., 2011. Hochschulabschlüsse im Umbruch. Studium und Übergang von Absolventinnen und Absolventen reformierter und traditioneller Studiengänge des Jahrgangs 2009. Hannover: HIS. / Rohrbach-Schmidt, D., 2009. The BIBB/IAB- and BIBB-BAuA Surveys of the Working Population on Qualification and Working Conditions in Germany. BIBB-FDZ Daten- und Methodenberichte, 1, 1-39. / Schaeper, H. & Briedis, K., 2004. Kompetenzen von Hochschulabsolventinnen und Hochschulabsolventen, berufliche Anforderungen und Folgerungen für die Hochschulreform. Hannover: HIS. / Toolsema, B., 2003. Werken met competenties. Naar een instrument voor de identificatie van competenties. Enschede: Universiteit Twente.

Montag,

03.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 251



## Zusammenhänge zwischen mütterlichem bzw. väterlichem Feedback und lernrelevanten motivationalen Variablen ihrer Kinder

Chair(s): Jessica Marschner (Ruhr-Universität Bochum)

Neben genetischen Faktoren beeinflussen Eltern durch ihren Erziehungsstil, ihre Einstellungen und Erwartungen sowie ihr Verhalten die Entwicklung ihrer Kinder bezüglich Lernen und Lernmotivation. So sind für Motivations- und Leistungsvariablen der Kinder neben Lehrer- und Unterrichtsfaktoren vor allem familiäre Einflussfaktoren entscheidend (vgl. PISA; TIMSS; Helmke et al., 1991; Lorenz & Wild, 2007; Sacher, 2012). Nach Grolnick und Slowiaczek (1994) beeinflussen Eltern vor allem schulbezogene Einstellungen und Motivationen der Kinder, die wiederum auf die Schulleistung wirken. Eltern dienen hierbei als Modell, an dem Kinder lernen, und als bedeutsame Quelle externen Feedbacks. Eltern erhalten allerdings wenig Unterstützung, wie sie häusliche Lehr-Lernsituationen gestalten können (Wild & Gerber, 2009), obwohl viele Eltern daran interessiert sind (Krumm, 1996; Lund et al., 2001). Es bietet sich folglich an, Eltern entsprechend „fortzubilden“, beispielsweise wie sie in Lern-Leistungssituationen angemessenes, motivationsförderliches Feedback geben. Elaboriertes auf den Lernprozess bezogenes Feedback hat sich in anderen Lernumwelten als wirksam erwiesen (Mory, 2004; Narciss, 2006). In einer Vorstudie soll zunächst untersucht werden, inwiefern solch ein elterliches Feedback mit lernmotivationalen Variablen ihrer Kinder zusammenhängt. Dabei wird das Feedbackverhalten der Mütter und Väter separat erfasst. Die Annahme ist, dass Feedback zur Selbststeuerung des Lernens, Feedback zur Herangehensweise und ein prozessorientierter Umgang mit Hausaufgaben positiv mit schulischer Selbstwirksamkeit und Lernmotivation zusammenhängen. Zu prüfen ist, ob diese Zusammenhänge gleichermaßen für Mütter wie für Väter bestehen, zumal Mütter stärker in die Hausaufgaben und häuslichen Lernsituationen involviert sind. Untersucht wurden N=94 Sechstklässler (55 % Jungen; 45% Mädchen) eines Gymnasiums. Per Selbstbericht wurde ihre Lern-Leistungsmotivation mit dem SELMO (Spinath et al., 2002) erhoben, zudem ihre schulische Selbstwirksamkeit (Jerusalem & Satow, 1999) und schulische Hilfslosigkeit (Jerusalem & Schwarzer, 1993). Ebenfalls aus Kindersicht (da handlungsleitend) wurden die mütterlichen und väterlichen Variablen erfragt. Es wurde Feedback zur Selbststeuerung (z.B. „Wenn ich mit meiner Mutter Schulaufgaben mache oder gemeinsam lerne, gibt sie mir Hinweise, wie ich mein Lernen planen kann.“; nach Berner et al., 2011), Feedback zur Herangehensweise (z.B. „Wenn ich meiner Mutter eine Klassenarbeit zeige, bespricht meine Mutter mit mir, wo ich mein Wissen gut eingesetzt habe.“; nach Berner et al., 2011) sowie der prozessorientierte Umgang mit Hausaufgaben erfasst (z.B. „Meine Mutter interessiert sich dafür, wie ich die Hausaufgaben gelöst habe.“; adaptiert an Rakoczy et al., 2005). Alle Skalen wiesen gute Reliabilitäten auf. Zur Hypothesenprüfung wurden partielle Korrelationen berechnet, um den Zusammenhang zwischen den motivationalen Variablen der Kinder und dem Feedbackverhalten der Eltern unabhängig von der schulischen Leistung der Kinder (Zeugnisnoten) aufzuzeigen. Folgende Zusammenhänge zeigten sich für die mütterlichen Variablen: zwischen Feedback zur Selbstregulation und der Lernzielorientierung ( $r = .278$ ); zwischen Feedback zur guten Herangehensweise und der Lernzielorientierung ( $r = .380$ ), der schulischen Selbstwirksamkeit ( $r = .262$ ) und der schulischen Hilfslosigkeit ( $r = -.432$ ); ebenso zwischen dem prozessorientierten Umgang mit Hausaufgaben und der Lernzielorientierung ( $r = .209$ ), der schulischen Selbstwirksamkeit ( $r = .311$ ) und der schulischen Hilfslosigkeit ( $r = -.231$ ). Bei den väterlichen Variablen zeigen sich gleich viele Zusammenhänge, wenn auch teilweise anders verteilt: zwischen dem Feedback zur Selbstregulation und der Lernzielorientierung ( $r = .343$ ), der schulischen Selbstwirksamkeit ( $r = .310$ ) und der schulischen Hilfslosigkeit ( $r = -.369$ ); zwischen dem Feedback zur guten Herangehensweise und der schulischen Selbstwirksamkeit ( $r = .296$ ) sowie der schulischen Hilfslosigkeit ( $r = -.341$ ), ebenso zwischen dem prozessorientierten Umgang mit Hausaufgaben und der schulischen Selbstwirksamkeit ( $r = .287$ ) sowie der schulischen Hilfslosigkeit ( $r = -.379$ ). Keine Zusammenhänge zeigen sich bezüglich des Feedbacks zur schlechten Herangehensweise. Hier könnte möglicherweise das häusliche Fehlerklima eine Moderatorvariable darstellen. Anregungen für das Geben eines motivations- und somit auch lernförderlichen Feedbacks lassen sich ableiten und nachfolgend in Form von Interventionsstudien auf die Kausalität hin prüfen.

- Grolnick, W. S. & Slowiaczek, M. L. (1994). Parents involvement in children's schooling: a multidimensional conceptualization and motivational model. *Child Development*, 65, 237-252.
- Helmke, A., Schrader, F.W. & Lehneis-Klepper, G. (1991). Zur Rolle des Elternverhaltens für die Schulleistungsentwicklung ihrer Kinder. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 23 (1), 1-22.
- Krumm, V. (1996). Über die Vernachlässigung der Eltern durch Lehrer und Erziehungswissenschaft. *Zeitschrift für Pädagogik, Beiheft 34*, 119-137.
- Lund, B. Rheinberg, F. & Gladash, U. (2001). Ein Elterntraining zum motivationsförderlichen Erziehungsverhalten in Leistungskontexten. *Zeitschrift für Pädagogische Psychologie*, 15 (3-4), 130-143.
- Mory, E.H. (2004). Feedback research revisited. In D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 745-783). New York: Simon & Schuster Macmillan.
- Narciss, S. (2006). *Informatives tutorielles Feedback*. Münster: Waxmann.
- Sacher, W. (2012). Erziehungs- und Bildungspartnerschaft zwischen Schule und Elternhaus. In Dagmar Killus und Klaus-Jürgen Tillmann (Hrsg.), *Eltern ziehen Bilanz. Ein Trendbericht zu Schule und Bildungspolitik in Deutschland ; die 2. JAKO-O Bildungsstudie* (S. 193-216). Münster: Waxmann.
- Wild, E. & Gerber, J. (2009). Lernlust statt Lernfrust – Evaluation eines Elterntrainings zur Verringerung von Hausaufgabenkonflikten bei Schülern mit Lernschwierigkeiten. *Psychologie in Erziehung und Unterricht*, 56, 303-318.
- Lorenz, F. & Wild, E. (2007). Parental involvement in schooling. Results concerning its structure and impact on students' motivation. In Manfred Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (pp. 299-316). Münster: Waxmann.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G092





## **Entwicklung des mathematischen Selbstkonzepts von Mädchen und Jungen in der Grundschule – Welchen Einfluss haben affektiv-motivationale Merkmale der Eltern?**

In Anlehnung an Wigfield und Eccles' (2000) Erwartungs-Wert-Modell wird im Vortrag die Bedeutung zweier familiärer Merkmale für die Selbstkonzeptentwicklung von Grundschulkindern in der Domäne Mathematik untersucht und dabei überprüft, ob sich geschlechtsspezifische Unterschiede in der Bedeutsamkeit dieser beiden Elternmerkmale für die Selbstkonzeptentwicklung ergeben.

Das Selbstkonzept ist innerhalb der pädagogisch-psychologischen Forschung eine viel beachtete Variable, da es einen engen Zusammenhang zur Leistung aufweist (z. B. Valentine, DuBois & Cooper, 2004; Weinert & Helmke, 1997). Für die Domäne Mathematik werden mittlere bis hohe Zusammenhänge zwischen Selbstkonzept und Leistung berichtet (Kammermeyer & Martschinke, 2003, 2006; Renkl, Helmke & Schrader, 1997).

Mädchen weisen dabei unter Kontrolle ihrer Mathematikleistung typischerweise ein geringeres Selbstkonzept in der Domäne Mathematik auf als Jungen (Dickhäuser & Stiensmeier-Pelster, 2003; Frome & Eccles, 1998), was unter anderem mit geschlechtsstereotypen Begabungskonzepten von Sozialisatoren erklärt wird (Wigfield & Eccles, 2000; Mösko, 2010). Angenommen wird, dass Eltern direkt und indirekt einen Einfluss auf die Entwicklung kindlicher Kompetenzüberzeugungen haben, indem elterliche Werthaltungen und Überzeugungen sich in (Unterstützungs-)Verhalten niederschlagen und sich darüber ungünstig auf die Selbstkonzeptentwicklung von Mädchen in der Domäne Mathematik auswirken. Empirisch belegt ist beispielsweise, dass sich mathematikbezogene Einstellungen einen indirekten Effekt auf die Leistung in Mathematik haben (Wild & Remy, 2002), dass elterliche Leistungseinschätzungen für Mädchen geringer ausfallen (Eccles, Jacobs & Harold, 1990; Mösko, 2010) und dass Eltern der Meinung sind, Mädchen müssten sich in Mathematik mehr anstrengen (Eccles Parsons, Adler & Kaczala, 1982).

Mit der vorliegenden Untersuchung soll überprüft werden, inwieweit der elterliche Bezug zum Fach Mathematik sowie die elterliche Leistungseinschätzung des Kindes im Fach Mathematik die Genese des mathematischen Selbstkonzepts am Ende des dritten Schuljahrs erklären kann und ob sich differentielle Effekte in Abhängigkeit des Geschlechts des Kindes ergeben. Es werden sowohl direkte als auch indirekte Effekte modelliert. Zur Beantwortung der Fragestellung werden Daten aus der Grundschulstudie PERLE (Lipowsky, Faust & Kastens, 2013) herangezogen ( $N = 898$ ; genestet in 46 Schulklassen).

Der elterliche Bezug zum Fach Mathematik sowie die elterliche Leistungseinschätzung des Kindes im Fach Mathematik wurden fragebogenbasiert am Ende des zweiten Schuljahrs erfasst. Das mathematische Selbstkonzept wurde am Ende des zweiten und am Ende des dritten Schuljahrs erhoben. Die Mathematikleistung wurde mit einem aus existierenden Testverfahren entwickelten Instrument erfasst und längsschnittlich Rasch-skaliert (Lipowsky, & Faust Greb, 2009; Lipowsky, Faust & Karst, 2011).

Analysiert werden die Daten mittels eines Strukturgleichungsmodells in Mplus (Muthén & Muthén, 2007). Die geclusterte Datenstruktur wurde bei der Modellschätzung berücksichtigt und fehlende Werte wurden mittels FIML-Algorithmus geschätzt. Der Fit des spezifizierten Modells weist auf eine gute Modellanpassung hin ( $\chi^2/df = 1.99$ ;  $p < .01$ ; RMSEA = .03; CFI = .95; TLI = .95). Im Ergebnis zeigt sich für die Gesamtstichprobe, dass unter Kontrolle der Mathematikleistung und des vorherigen Selbstkonzepts der elterliche Bezug zum Fach Mathematik sowohl direkt als auch mediiert über die elterliche Leistungseinschätzung das Selbstkonzept am Ende des dritten Schuljahrs schwach, aber signifikant beeinflusst. Deutlichere Effekte werden sichtbar, wenn das Modell getrennt nach Geschlechtern betrachtet wird. So zeigt sich, dass der elterliche Bezug zum Fach Mathematik lediglich das Selbstkonzept der Mädchen beeinflusst ( $\beta = .22^*$ ), wohingegen die Leistungseinschätzung der Eltern sich lediglich auf das Selbstkonzept der Jungen auswirkt ( $\beta = .23^*$ ). Die elterliche Leistungseinschätzung für Mädchen hängt stärker vom elterlichen Bezug zum Fach Mathematik ab, für Jungen hingegen stärker von der vorherigen Mathematikleistung. Die Ergebnisse werden vor dem theoretischen Hintergrund diskutiert und Ansätze zur Intervention aufgezeigt.

Eccles Parsons, J., Adler, T. F. & Kaczala, C. M. (1982). Socialization of Achievement Attitudes and Beliefs: Parental Influences. *Child Development*, 53(2), 310-321.

Kammermeyer, G. & Martschinke, S. (2003). Schulleistungen und Fähigkeitsselbstbild im Anfangsunterricht – Ergebnisse aus dem KILIA-Projekt. *Empirische Pädagogik*, 17(4), 486–503.

Kammermeyer, G. & Martschinke, S. (2006). Selbstkonzept- und Leistungsentwicklung in der Grundschule – Ergebnisse aus der KILIA-Studie. *Empirische Pädagogik*, 20, 245–259.

Lipowsky, F., Faust, G. & Greb, K. (2009). Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE) - Teil 1. Materialien zur Bildungsforschung Band 23/1. Frankfurt/Main: GPF, DIPF.

Lipowsky, F., Faust, G. & Kastens, C. (2013). Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren. Münster: Waxmann.

Lipowsky, F., Faust, G. & Karst, K. (2011). Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE) - Teil 2. Materialien zur Bildungsforschung Band 23/2. Frankfurt/Main: GPF, DIPF.

Mösko, E. (2010). Elterliche Geschlechtsstereotype und deren Einfluss auf das mathematische Selbstkonzept von Grundschulkindern. Dissertation: Universität Kassel.

Muthén, L. K. & Muthén, B. O. (2007). *Mplus user's guide*. Fifth edition. Los Angeles, CA: Muthén & Muthén.

Renkl, A., Helmke, A. & Schrader, F. W. (1997). Schulleistung und Fähigkeitsselbstbild – Universelle Beziehungen oder kontextspezifische Zusammenhänge? Ergebnisse aus dem Scholastik-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter*. Weinheim: Beltz, 374–383.

Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A metaanalytic review. *Educational Psychologist*, 39, 111–133.

Weinert, F. E. & Helmke, A. (1997) (Hrsg.), *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlagsunion.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25, 68–81.

Wild, E. & Remy, K. (2002). Affektive und motivationale Folgen der Lernhilfen und lernbezogenen Einstellungen von Eltern. *Unterrichtswissenschaft*, 30, 27-51.





## ***Der Effekt von verschiedenen Prüfungsformaten auf Lernstrategien, Lernmotivation und Lernerfolg***

Laut der Theorie des Constructive Alignment neigen Lernende dazu, ungeachtet der eigentlichen Lernziele einer Lehr-Lern-Situation, ihre Lernaktivitäten an Prüfungen auszurichten, um in diesen gut abschneiden zu können. Der Umstand, dass sich Lernende an den Prüfungsformen orientieren wird auch Backwash-Effect bezeichnet. Während für den Lehrenden die Prüfung am Ende der Planungen steht, bildet sie für Lernende oftmals den Ausgangspunkt ihres Tuns. Ein Schlüssel für eine gelungene Lehr-Lern-Situation liegt also im Abgleich aller darin enthaltenen Elemente; beginnend bei den Lernzielen über die didaktische Aufbereitung der Inhalte bis hin zur Prüfungsform (constructive alignment). Assoziieren Lernende mit einem Prüfungsformat andere Lernziele als der Lehrende, verlaufen Lehr- und Lernprozesse nicht parallel.

Ziel dieser Studie war es zu überprüfen, ob Lernende ihre Lernaktivitäten an einer angekündigten Prüfungsform (Faktenwiedergabe versus Transfer) ausrichten, indem sie entsprechende Lernstrategien anwenden und in Folge dessen ein unterschiedliches Verständnisniveau erreichen. Es wurde erwartet, dass eine mit Transferwissen assoziierte Prüfungsform zu tiefenorientierten Lernstrategien und damit zu einem besseren Verständnis führt. Des Weiteren wurde der Einfluss der Prüfungsformen auf den Lernerfolg und die Motivation untersucht.

Einundachtzig Studierende im zweiten Fachsemester Bildungswissenschaft nahmen an der quasi-experimentellen Feldstudie teil. Die Studierenden verteilten sich entsprechend ihres Stundenplans auf zwei identische Seminare, die von der gleichen Dozentin, zur gleichen Uhrzeit, an unterschiedlichen Tagen angeboten wurden. Es handelte sich um ein Pflichtseminar, das alle Studierenden dieses Jahrgangs besuchen mussten. Im Rahmen der Untersuchung wurde eine Seminarsitzung konzipiert, die nach strengem Ablaufplan in beiden Seminargruppen identisch gehalten wurde. Zu Beginn der jeweiligen Seminarsitzung wurde eine Prüfung (Faktenwissenstest oder Transferwissenstest) für die darauf folgende Seminarsitzung angekündigt. Diese Prüfung wurde in der folgenden Woche auch wie angekündigt durchgeführt, ging aber nicht in die Analyse ein. Zusätzlich beschrieben die Studierenden, wie sie sich auf diesen Test vorbereitet hatten. Diese Selbstbeschreibungen wurden anhand verschiedener Lernstrategien (Elaboration, Organisation, Wiederholung) kategorisiert. Nach drei Wochen unterzogen sich alle Studierenden einem identischen und nicht angekündigten Nachtest, mit dessen Hilfe das Verständnisniveau der Studierenden beurteilt werden sollte. Zu beiden Erhebungszeitpunkten wurden Interesse, Anstrengung, Selbstwirksamkeitserwartung und Stresserleben der Studierenden erhoben.

Hinsichtlich der Lernstrategienutzung zeigten sich Gruppenunterschiede. Während keine signifikanten Unterschiede hinsichtlich der Verwendung von Organisations- und Wiederholungsstrategien gefunden werden konnten, gaben die Lernenden, die sich auf den Transferwissenstest vorbereiteten, signifikant häufiger an, Elaborationsstrategien verwendet zu haben, als die Lernenden, die sich auf die Faktenwiedergabe vorbereiteten. Dies weist auf den vermehrten Einsatz tiefenorientierter Lernstrategien bei der Vorbereitung eines Transferwissenstests hin.

Bei der Analyse des Nachtests zeigte sich, dass Studierenden, die sich drei Wochen zuvor zu diesem Thema auf einen Transferwissenstest vorbereitet hatten sowohl mehr Fakten erinnern konnten als auch bei den Anwendungsaufgaben bessere Ergebnisse erzielten.

Während sich beim ersten Messzeitpunkt bei Interesse und Anstrengung keine signifikanten Unterschiede zeigten, erlebten Studierende in der Faktengruppe ein höheres Belastungslevel und eine geringere Selbstwirksamkeitserwartung als Studierende in der Transfergruppe. Die Erwartung eines Faktenwissenstests wirkte sich also auf einzelne Motivationsvariablen negativ aus. Beim zweiten Messzeitpunkt zeigten sich keine Motivationsunterschiede.

Die Ergebnisse weisen darauf hin, dass die postulierte Ausrichtung der Lernaktivitäten an einer angekündigten Prüfungsform stattfindet und sich in der Verwendung spezifischer Lernstrategien äußert. In Erwartung eines Transferwissenstests verwenden Studierende mehr Elaborationsstrategien und erleben sich selbst als kompetent. Insgesamt sind sie besser in der Lage sich Fakten zu merken und ihr Wissen auch tatsächlich anzuwenden. Das Format einer Prüfung kann daher als wirkungsvolles didaktisches Mittel betrachtet werden und sollte so gewählt werden, dass die Lernenden zur Verwendung geeigneter Strategien zum Erreichen angedachter Lernziele angeregt werden.

## **Die Mediation von Effekten der sozialen Herkunft auf post-instruktionales physikalisches Wissen durch kognitive und motivationale Schülermerkmale**

Unter Nutzung der Daten der großen Schulleistungsstudien der letzten fünfzehn Jahre ist festgestellt worden, dass in Deutschland, trotz eines leicht gegenläufigen Trends, eine relativ enge Kopplung zwischen Kompetenzerwerb und sozialer Herkunft besteht (z.B., Ehmke & Jude, 2010; Bensen, Frey, & Bos, 2008). Vergleichsweise spärlich ist allerdings beforscht worden, welche psychologischen Merkmale von Schüler(innen) in konkreten Lernsituationen für diese Kopplung verantwortlich sind (für eine Ausnahme siehe Baumert, Watermann, & Schümer, 2003). Die vorliegende Untersuchung ergänzt den bestehenden Wissensstand, indem sie durch Übertragung von vornehmlich mit inhaltlich breit angelegten Kompetenztests gewonnenen Erkenntnissen auf eine eng umgrenzte Unterrichtsreihe psychologische Mediatoren der Beziehung von sozialer Herkunft und Schulleistung in den Blick nimmt.

Hierzu wurde auf Daten der initialen querschnittlichen Phase des PLUS-Projektes, welches im Rahmen der Forschergruppe NWU Essen durchgeführt wurde, zurückgegriffen (z.B., Möller, Kleickmann, & Lange, 2013). Es waren 60 Grundschullehrkräfte und 54 Sekundarschullehrkräfte der Hauptschule und des Gymnasiums aufgefordert worden eine Unterrichtssequenz zum Thema „Verdunstung und Kondensation“ durchzuführen. Entsprechend standen aus Befragungen der betroffenen Schüler(innen) Angaben von 1326 Viertklässlern und 1354 Sechstklässlern zu fluider Intelligenz, themenspezifischen Vorwissen, physikbezogenem Fähigkeitsselbstkonzept und individuellen Interesse an physikalischen Inhalten vor der Unterrichtsreihe als potentiellen kognitiven und motivationalen Mediatoren zur Verfügung. Die soziale Herkunft der Schüler(innen) wurde mittels des International Socio-Economic Index of occupational status operationalisiert ( Ganzeboom, de Graaf, Treiman, & de Leeuw, 1992), welcher auf Basis von Elternangaben zur aktuellen Berufstätigkeit kodiert wurde. Die abhängige Variable schließlich bildete das post-instruktionale Wissen der beteiligten Schüler(innen) über Verdunstung und Kondensation. Die genutzten Tests und Skalen wurden gemäß des einfachen Rasch-Modells (Rasch, 1960), des Partial-credit-Modells (Masters, 1982) oder des Rating-scale-Modells (Andrich, 1978) kalibriert. Zur Behandlung fehlender Werte wurden 50 multipel imputierte Datensätze mithilfe des R-Pakets mice erzeugt (van Buuren & Groothuis-Oudshoorn, 2011). Das Imputationsmodell bezog Variablen aus externen Quellen mit ein (z.B. Arbeitslosenquote auf Kreisebene). Nach Berechnung des sozialen Gradienten im post-instruktionales Wissen,  $B = 0.09$ ,  $SE = 0.02$ ,  $p < .001$ , wurden zunächst separat für jeden potentiellen Mediator Mediationsanalysen durchgeführt. Bei diesen Auswertungen wurden die Mehrebenenstruktur der Daten sowie Mittelwertsdifferenzen in den abhängigen Variablen zwischen den verschiedenen Schulformen berücksichtigt. Weiterhin wurde geprüft, inwieweit Zusammenhänge zwischen sozio-ökonomischen Status, Mediatoren und post-instruktionalem Wissen durch die Schulformzugehörigkeit der Schüler(innen) moderiert wurden. Schließlich wurden die Ergebnisse in einem finalen Mehrebenen-Pfadmodell mit allen relevanten Mediatoren zusammengefasst. Der weitaus größte Teil des Effektes der sozialen Herkunft auf post-instruktionales Wissen wurde über das themenspezifische Vorwissen der Schüler(innen) vermittelt. Obschon fluide Intelligenz sich in den separaten Mediationsanalysen als relevanter Mediator erwiesen hatte, vermittelte es über das themenspezifische Vorwissen hinaus praktisch keinen eigenständigen Anteil des Effektes der sozialen Herkunft auf das post-instruktionale Wissen. Das Fähigkeitsselbstkonzept hingegen vermittelte zusätzlich zum Vorwissen einen sehr kleinen Teil des Effektes des sozio-ökonomischen Status auf das post-instruktionale Wissen. Das individuelle Interesse an physikalischen Inhalten stand in keinerlei Zusammenhang zur sozialen Herkunft. An zwei Stellen wurde das Gefüge der Beziehungen zwischen den Variablen durch die Schulformzugehörigkeit moderiert: Ein Zusammenhang zwischen sozialer Herkunft und Fähigkeitsselbstkonzept war nur für die Grundschule und nicht für die Schulformen der Sekundarstufe zu beobachten. Außerdem fiel in der Hauptschule der Effekt des Vorwissens auf das post-instruktionale Wissen tendenziell schwächer aus als in den anderen Schulformen. Das themenspezifische Vorwissen der Schüler(innen) erwies sich als mit Abstand wichtigster psychologischer Mediator von Effekten der sozialen Herkunft auf das post-instruktionale Wissen. Dieses Ergebnis steht im Einklang mit Befunden, welche vor allem einen differenziellen Kompetenzerwerb außerhalb des Unterrichts als Ursache sozialer Disparitäten in akademischen Kompetenzen ausgemacht haben (z.B., Alexander, Entwistle, & Olson, 2001, 2007). Aus pädagogischer Sicht erfreulich, ist die Tatsache, dass in der hier untersuchten Altersgruppe Kinder aus sozial benachteiligten Familien genauso interessiert am Lernen physikalischer Inhalte waren wie ihre sozial begünstigteren Altersgenossen.

Alexander, K. L., Entwistle, D. R., & Olson, L. S. (2001). Schools, achievement and ine-quality: A seasonal perspective. *Educational Evaluation and Policy Analysis*, 23, 171–191.

Alexander, K. L., Entwistle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, 72, 167–180.

Andrich, D. A. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561–573.

Baumert, J., Watermann, R., & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. *Zeitschrift für Erziehungswissenschaft*, 6, 46–72.

Bensen, M., Frey, K. A., & Bos, W. (2008). Soziale Herkunft. In W. Bos, M. Bensen, J. Baumert, M. Prenzel, C. Selzer & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 141–156). Münster: Waxmann.

Ehmke, T., & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 231–254). Münster: Waxmann.

Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D. J., & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56.

Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.

Möller, K., Kleickmann, T., & Lange, K. (2013). Naturwissenschaftliches Lernen im Übergang von der Grundschule zur Sekundarstufe. In H. E. Fischer & E. Sumfleth (Hrsg.), *nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht* (S. 57–120). Berlin: Logos.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67.



## **Entwicklung eines Fragebogens zu Determinanten freizeithlicher Leseaktivitäten bei Grundschulkindern**

Warum sollten Kinder in ihrer Freizeit lesen? Im Kontext von Leseforschung und Forschung zu kultureller und ästhetischer Bildung wird regelmäßig darauf verwiesen, dass Lesen zum Erwerb von Lesekompetenz und zur Persönlichkeitsentwicklung beiträgt. Zudem ist Lesekompetenz eine wichtige Voraussetzung für die Partizipation am kulturellen, politischen und gesellschaftlichen Leben. Für eine optimale Förderung dieser Prozesse ist es wünschenswert, dass Kinder über die Schule hinaus auch freiwillig in ihrer Freizeit lesen. Um das freizeithliche Lesen zu fördern, bedarf es daher umfassenden Wissens zu den Determinanten des Lesens in der Freizeit bei Kindern. Zwar liegen bereits Studien und Instrumente zur Lesemotivation von Kindern wie beispielsweise der Motivation for Reading Questionnaire (MRQ) von Wigfield und Guthrie (Wigfield & Guthrie, 1997) vor. Allerdings weist dieser Fragebogen keine klare theoretische Fundierung auf. Außerdem nimmt der MRQ die Lesemotivation nur unspezifisch in den Blick: Unterschiedliche Motive für das Lesen im Freizeitkontext vs. das Lesen im Schulkontext werden nicht differenziert. Ziel der vorliegenden Studie war daher die Entwicklung eines Fragebogens, der auf die Erfassung der Determinanten des freizeithlichen Lesens bei Grundschulkindern fokussiert. Die Konstruktion des Fragebogens erfolgte basierend auf der Theorie des geplanten Verhaltens (Ajzen, 1991) mit den Prädiktoren Einstellung ( $\alpha = .84$ , 11 Items), subjektive Norm ( $\alpha = .61$ , 4 Items) und wahrgenommene Verhaltenskontrolle ( $\alpha = .76$ , 6 Items). Das Kriterium bildete die Häufigkeit des freizeithlichen Lesens ( $\alpha = .79$ , 3 Items). Zur Überprüfung der Faktorenstruktur der Prädiktoren wurde der entwickelte Fragebogen in einer Stichprobe von  $N = 1014$  Kinder (davon 50.2 % Jungen) im Alter von  $M = 8.68$  ( $SD = .60$ ) Jahren eingesetzt. Die konfirmatorische Faktorenanalyse (CFA) wies einen guten Fit auf ( $RMSEA = .031$ ,  $CFI = .97$ ,  $TLI = .96$ ). Allerdings korrelierte eine Subskala des Prädiktors Einstellung hoch mit anderen Subskalen ( $r \geq .61$ ); dies deutet auf Multikollinearität hin. Insgesamt sprechen die Ergebnisse der CFA für eine erfolgreiche Entwicklung eines Fragebogens zur Erfassung der Determinanten freizeithlichen Lesens bei Grundschulkindern. Lösungsansätze für den Umgang mit Multikollinearität werden vorgestellt und diskutiert. Implikationen für ein Rahmenmodell kultureller Partizipation werden erörtert (Kröner, im Druck).

Ajzen, I. (1991). The Theory of Planned Behavior. *Organisational Behavior And Human Decision Processes*, 50(2), 179.

Kröner, S. (im Druck). Determinanten kultureller Aktivitäten bei Jugendlichen. *Zeitschrift für Erziehungswissenschaft*.

Wigfield, A. & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89(3), 420-432.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G092

## Die Bedeutung der sozialen Einbettung für den Bildungserfolg von Migranten in Deutschland: Analysen mit dem Nationalen Bildungspanel

### Chair(s): Benjamin Schulz (Universität Mannheim)

Oftmals leben Migranten und deren Nachkommen räumlich konzentriert in ethnischen Gemeinschaften. Infolgedessen sind die Beziehungsmuster und sozialen Netzwerke vieler Migranten über Generationen hinweg ethnisch geprägt (Cheng et al., 2007; Kalmijn & van Tubergen, 2006; Li et al., 2008; Völker et al., 2008). Was aber folgt aus einer solchen Einbettung in die eigene Herkunftsgruppe für den Kompetenzerwerb und die Bildungsbeteiligung von Migranten? Trägt die Einbettung in ethnische Gemeinschaften zur Entstehung ethnischer Bildungsungleichheit bei? Können Unterschiede in der sozialen Einbettung zwischen Migrantengruppen dazu beitragen, Unterschiede in der Bildungsbeteiligung zwischen diesen Gruppen zu erklären? In der Integrationsforschung ist umstritten, ob ethnische Netzwerke den Bildungserfolg und die Kompetenzentwicklung eher befördern oder behindern. Einerseits wird argumentiert, dass ethnische Netzwerke hinderlich sind, weil sie a) weniger soziale Ressourcen bereitstellen als ethnisch heterogene Netzwerke, weil sie den Aufbau inter-ethnischer Beziehungen beeinträchtigen (Farwick, 2009), b) weniger Informationen über die Funktionsweise des Bildungssystems bieten (vgl. (Kristen, 2008) und die Gelegenheiten zum Erwerb der Sprache des Aufnahmelandes reduzieren können (Bialystok, 2009; Esser, 2006; Kaushanskaya & Marian, 2007; Leseman et al., 2009). Andererseits weisen vor allem Vertreter der Segmented Assimilation Theorie darauf hin, dass ethnische Netzwerke den Bildungserfolg positiv beeinflussen können, weil sie a) Kinder und Jugendliche mit Migrationshintergrund davor bewahren können, etwaige negative Einstellungen und Verhaltensweisen von Peers zu übernehmen (Portes & Rumbaut, 2001; Garcia-Reid, 2007), weil sie vor Beeinträchtigungen durch Diskriminierungserfahrungen und Ausgrenzung schützen können (Portes & Rumbaut, 2001) und weil sie den Bildungserfolg befördern können, wenn sie kulturelle Werte und Orientierungen vermitteln, die auf Bildungserwerb und sozialen Aufstieg ausgerichtet sind (Zhou & Bankston, 1994; Shah et al., 2010; Zhou & Xiong, 2005). Empirische Studien zu diesen komplexen Fragen, weisen bislang äußerst gemischte Ergebnisse aus (vgl. Schulz, 2013). Für Deutschland gibt es zudem vergleichsweise wenige belastbare Befunde. Dies ist auch darauf zurückzuführen, dass in vielen deutschen wie internationalen Studien ethnische Netzwerkcompositionen und soziale Ressourcen oftmals nur ungenügend erfasst werden (Helbig, 2010, S. 676; Li et al., 2008, S. 393). Insgesamt scheint eine bedingte Sichtweise angemessen: Ob ethnische Netzwerke den Bildungserfolg von Migranten befördern oder behindern, hängt von den spezifischen sozialen Ressourcen, den verfügbaren Gelegenheiten und den vermittelten Aspirationen ab (vgl. beispielsweise (Kroneberg, 2008). Die vermittelnden Mechanismen müssen also genauer spezifiziert und überprüft werden (Crosnoe, 2009; Helbig, 2010; Zhou, 2009). Ausgehend von sozialkapitaltheoretischen Überlegungen, untersucht dieser Beitrag zwei Mechanismen zur Wirkung der sozialen Einbettung auf den Kompetenzerwerb und den Bildungserfolg: a) die Vermittlung von Aspirationen durch normative Bezugsgruppen, b) die Bedeutung sozialer Ressourcen wie Informationen oder Unterstützungsleistungen. Dafür werden Daten des Nationalen Bildungspanels (NEPS) genutzt. Die NEPS-Daten ermöglichen es aufgrund umfassender Messungen ethnischer Netzwerk- und Nachbarschaftskompositionen sowie der Erfassung verschiedener Sozialkapitaldimensionen, die skizzierten Fragen wesentlich direkter zu testen, als bisherige Studien dies zumeist vermochten. Zudem können Neuntklässler/innen traditioneller und neuer Migrantengruppen untersucht werden (Spätaussiedler aus der ehemaligen Sowjetunion, Ex-Jugoslawien, Polen, Südeuropa, Türkei). Außerdem können Migranten bis in die dritte Generation identifiziert werden. Aufgrund der komplexen NEPS-Datenstruktur werden Mehrebenenmodelle angewandt: Der Effekt ethnischer Netzwerke und sozialer Kapitalien auf die Mathematik- und Sprachkompetenzen werden unter Kontrolle einer Vielzahl weiterer Merkmale mittels hierarchischer linearer Modelle geschätzt. Um die Chancen auf einen Gymnasialbesuch zu analysieren, werden logistische random-coefficient-Modelle genutzt. Die Ergebnisse weisen für fast alle untersuchten Migrantengruppen darauf hin, dass ethnisch geprägte Freundschaftsnetzwerke zunächst tatsächlich mit geringeren Kompetenzniveaus und geringerem Bildungserfolg einhergehen. Diese Ausgangseffekte werden, mit Ausnahme türkischer Schüler/innen, jedoch durch die Ressourcenausstattung und Netzwerkpositionierung vermittelt. Negative Effekte normativer Bezugsgruppen, die für eine „Aspirationsabkühlung“ sprechen könnten (Haeberlin et al., 2004), zeigen sich nicht. Insbesondere für Schüler/innen mit türkischem Migrationshintergrund scheint der Sprachgebrauch in der Familie und in Peernetzwerken zudem bedeutsam für den Kompetenzerwerb.

- Bialystok, Ellen (2009): Effects of bilingualism on cognitive and linguistic performance across the lifespan. In: I. Gogolin und U. Neumann (Hg.): Streitfall Zweisprachigkeit-The Bilingualism Controversy. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 53–67.
- Cheng, Simon; Martin, Leslie; Werum, Regina E. (2007): Adult social capital and track placement of ethnic groups in Germany. In: American Journal of Education 114 (1), S. 41–74. DOI: 10.1086/520691.
- Crosnoe, Robert (2009): Low-income students and the socioeconomic composition of public high schools. In: American Sociological Review 74 (5), S. 709–730. DOI: 10.1177/000312240907400502.
- Esser, Hartmut (2006): Sprache und Integration: die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten. Frankfurt a. M.: Campus.
- Farwick, A. (2009): Segregation und Eingliederung. Zum Einfluss der räumlichen Konzentration von Zuwanderern auf den Eingliederungsprozess: VS Verlag für Sozialwissenschaften.
- Garcia-Reid, Pauline (2007): Examining social capital as a mechanism for improving school engagement among low income Hispanic girls. In: Youth & Society 39 (2), S. 164–181. DOI: 10.1177/0044118X070303263.
- Helbig, Marcel (2010): Neighborhood does matter! In: KZfS Kölner Zeitschrift für Soziologie und Sozialpsychologie 62 (4), S. 655–679. DOI: 10.1007/s11577-010-0117-y.
- Kalmijn, Matthijs; van Tubergen, Frank (2006): Ethnic intermarriage in the Netherlands: confirmations and refutations of accepted insights. In: European Journal of Population 22 (4), S. 371–397. DOI: 10.1007/s10680-006-9105-3.
- Kaushanskaya, Margarita; Marian, Viorica (2007): Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. In: Language Learning 57 (1), S. 119–163.
- Kristen, Cornelia (2008): Primary school choice and ethnic school segregation in German elementary schools. In: European Sociological Review 24 (4), S. 495–510.
- Kroneberg, Clemens (2008): Ethnic communities and school performance among the new second generation in the United States: Testing the theory of segmented assimilation. In: The Annals of the American Academy of Political and Social Science 620 (1), S. 138–160.
- Leseman, Paul P. M.; Scheel, Anna F.; Mayo, Aziza Y.; Messer, Marielle H. (2009): Bilingual development in early childhood and the languages used at home: competition for scarce resources? In: I. Gogolin und U. Neumann (Hg.): Streitfall Zweisprachigkeit-The Bilingualism Controversy. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 289–316.
- Li, Yaojun; Savage, Mike; Warde, Alan (2008): Social mobility and social capital in contemporary Britain. In: The British journal of sociology 59 (3), S. 391–411.
- Portes, Alejandro; Rumbaut, Rubén G. (2001): Legacies: The story of the immigrant second generation. Berkeley, CA: University of California Press.
- Schulz, Benjamin (2013). Social Embedding and Educational Achievement of Immigrants. A Review and Appraisal. Mannheim: Arbeitspapier Nr. 152, Mannheimer Zentrum für Europäische Sozialforschung. <http://www.mzes.uni-mannheim.de/publications/wp-wp-152.pdf>
- Shah, Bindi; Dwyer, Claire; Modood, Tariq (2010): Explaining Educational Achievement and Career Aspirations among Young British Pakistanis: Mobilizing Ethnic Capital? In: Sociology 44 (6), S. 1109–1127.
- Völker, Beate; Pinkster, Fenne; Flap, Henk (2008): Inequality in social capital between migrants and natives in the Netherlands. In: Frank Kalter (Hg.): Migration und Integration. Kölner Zeitschrift für Soziologie und Sozialpsychologie. Special Volume. Wiesbaden: VS, S. 325–350.
- Zhou, Min (2009): How neighbourhoods matter for immigrant children: the formation of educational resources in Chinatown, Koreatown and Pico Union, Los Angeles. In: Journal of Ethnic and Migration Studies 35 (7), S. 1153–1179.
- Zhou, Min; Bankston, Carl L. (1994): Social capital and the adaptation of the second generation: The case of Vietnamese youth in New Orleans. In: International migration review 28 (4), S. 821–845.
- Zhou, Min; Xiong, Yang Sao (2005): The multifaceted American experiences of the children of Asian immigrants: Lessons for segmented assimilation. In: Ethnic and Racial Studies 28 (6), S. 1119–1152.

Montag,

03.03.2014,

15:15 - 17:00,

PEG-Gebäude

Raum 1.G102



## **Der unterschiedliche Bildungserfolg von Migranten der zweiten Generation in der Schweiz und in Deutschland: Gruppen- oder Länderunterschiede?**

In diesem Beitrag untersuchen wir den Bildungserfolg von Migranten der zweiten Generation und erklären den unterschiedlichen Erfolg einzelner Gruppen zwischen der Schweiz und Deutschland. Ausgangspunkt sind die gefundenen Unterschiede zwischen Migranten italienischer Herkunft, deren Bildungsnachteile in der Schweiz geringer zu sein scheinen als in Deutschland. Als mögliche Ursachen für den ungleichen Bildungserfolg werden zum einen sozial differenzielle Migrationsbewegungen und zum anderen institutionelle Unterschiede (Italienisch ist in Teilen der Schweiz Amtssprache, EU-rechtlich garantierte Freizügigkeit zwischen Deutschland und Italien) diskutiert. In aktuellen Studien wurden insbesondere die Bedeutung der Sprache und die der unterschiedlichen sozialen Zusammensetzung bezweifelt (Thränhardt, 2000; Dollmann, 2004).

Mit diesem Beitrag verfolgen wir drei Ziele: 1. Die Replikation der Befunde des größeren Bildungserfolg von Kindern italienischer Einwanderer in der Schweiz im Vergleich zu Deutschland mit den Daten der PISA-Studie von 2003; 2. die Untersuchung von Länderunterschieden für anderen Herkunftsgruppen; 3. die Aufklärung dieser Unterschiede mittels tiefergehender Betrachtung von Gruppenunterschieden und institutionellen Rahmenbedingungen.

Die Gründe für ethnische Bildungsgleichheiten sind vor allem sozioökonomische Unterschiede der einzelnen Herkunftsgruppen. Migranten verfügen über weniger bildungsrelevante Ressourcen, was sich negativ auf nötige Bildungsinvestitionen auswirkt (Esser, 2006). Hinzu kommen Belege für den Einfluss unterschiedlicher institutioneller Konfigurationen auf den Bildungserfolg von Migranten (Van de Werfhorst & Mijs, 2010). Diese wurden jedoch in bisherigen Studien nicht ausreichend als Erklärungsfaktor beachtet. Grundsätzlich kann davon ausgegangen werden, dass im schweizerischen und deutschen Kontext dieselben Wirkmechanismen den vorhandenen ethnischen Bildungsungleichen zugrunde liegen. Zudem weisen die beiden Bildungssysteme ähnliche Grundzüge auf. In beiden Ländern ist das Bildungssystem föderal organisiert (Bundesländer bzw. Kantone) und das Sekundarschulsystem ist stark stratifiziert und standardisiert, und am Ende der Grundschule erfolgt eine leistungsorientierte Selektion der Schüler/innen. Es gibt jedoch zwei maßgebliche Unterschiede: Zum einen ist in einem Teil der Schweiz Italienisch die Amts- und Unterrichtssprache und zum anderen dauert die Grundschule in den meisten Kantonen sechs Jahre, wohingegen in Deutschland die Grundschule mehrheitlich vier Jahre dauert. Wir argumentieren, dass die unterschiedliche Performanz von Migranten insbesondere durch die unterschiedlich lange Grundschuldauer erklären lässt. Aufgrund ihrer sozioökonomischen Benachteiligungen, ihres geringeren Wissens um das Bildungssystem des Aufnahmelandes, und sprachlicher Benachteiligung erfahren Migrantenkinder weniger häusliche Förderung und treffen andere Bildungsentscheidungen, was dazu führt, dass sie häufiger in niedrigen Bildungsgängen vertreten sind (Kristen, 2005; Kristen & Granato, 2007). Diese Bildungsgänge bieten nur wenig leistungsförderliche Lernumwelten u.a. aufgrund von nachteiliger Schülerkomposition (Baumert et al., 2006). Je früher eine Leistungsselektion stattfindet, desto stärker sollten diese Effekte sein und zu einer stärkeren Benachteiligung von Migranten in Deutschland führen als in der Schweiz. Für unsere Analyse verwenden wir die nationalen Erweiterungsstudien von PISA 2003 für Deutschland und die Schweiz ( $n(\text{gepoolt})=36.958$ ) und vergleichen die Leseleistung der fünfzehnjährigen Schüler/innen in der neunten Klasse. Fehlende Werte bei den relevanten soziodemografischen Hintergrundvariablen wurden multipel imputiert. Fallzahlbedingt werden neben den italienischen Migranten noch Türkischstämmige und Migranten aus dem ehemaligen Jugoslawien verglichen. Im ersten Schritt können wir die Ergebnisse bisheriger Studien replizieren und finden auch nach Kontrolle der familiären Sprachnutzung neben dem Kompetenzvorsprung der Italiener auch einen Vorsprung für Türken in der Schweiz gegenüber diesen Herkunftsgruppen in Deutschland. Im zweiten Schritt können wir jedoch zeigen, dass der Vorteil der Italiener lediglich durch Gruppenunterschiede zustande kommt. Im letzten Schritt wird der Kompetenzvorteil der Türken in der Schweiz schließlich durch die institutionellen Unterschiede erklärt. Unter Kontrolle der Grundschuldauer, Schulformplatzierung und Schulzusammensetzung lassen sich keine Länderunterschiede der Leseleistung in den drei genannten Herkunftsgruppen finden.

Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungssystem* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.

Dollmann, Jörg (2004). Ethnic Differences in Educational Attainment: Italian Immigrants in Germany and Switzerland (Poster Presentation). European Society or European Societies? EURESCO/ECSR-Conference on the Causes and Consequences of Low Education in Contemporary Europe, Granada, 18. bis 23. September 2004.

Kristen, Cornelia (2005). *School Choice and Ethnic School Segregation: Primary School Selection in Germany*. Münster: Waxmann.

Kristen, Cornelia; Granato, Nadia (2007). The educational attainment of the second generation in Germany: Social origins and ethnic inequality. *Ethnicities* 2007; 7; 343

Thränhardt, D. (2000). „Einwandererkulturen und soziales Kapital; Eine komparative

Analyse“, in: Thränhardt, D.; Hunger, U. (Hrsg.): *Einwanderer-Netzwerke und ihre Integrationsqualität in Deutschland und Israel*, Münster: LIT, 15-51.

Van de Werfhorst, H.G. & Mijs. J.J.B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407-428.





## **Schul(miss)erfolg von Migranten: Ein bildungs- und gesundheitssoziologischer Erklärungsversuch**

Die nachteilige Bildungspositionierung von Migranten ist gut dokumentiert, hochaktuell und weiterhin erklärungsbedürftig. Nach bisherigem Kenntnisstand stehen hinter diesem Befund in erster Linie schlechte Noten, nicht aber ein Motivations- oder Aspirationsmangel. Nach den Bestimmungsgründen primärer ethnischer Herkunftseffekte im Schulerfolg fragend, widmet sich der Beitrag theoretisch wie empirisch der frühen – bereits im Mutterleib angelegten – Prägung des Schulerfolgs. Getestet wird die für den deutschen Forschungskontext kaum geprüfte Annahme, dass sich ethnische Bildungsnachteile auf eine Kumulation bereits zum Lebensbeginn bestehender (gesundheitlicher) Nachteilen zurückführen lassen. Dafür sprechen folgende empirische Regelmäßigkeiten: Kinder mit Migrationshintergrund kommen häufiger zu früh zur Welt, nehmen seltener an Maßnahmen zur Früherkennung von Entwicklungsverzögerungen teil und verbringen weniger Zeit in lernförderlichen Kontexten, wie dem Kindergarten. Wenig überraschend weichen Migrantenkinder im Vergleich zu Einheimischen bezüglich ihres Einschulungsalters nach „oben“ ab (Zurückstellung). Gleichzeitig, und eher unerwartet, sind Migrantenkinder unter den Früheingeschulerten überrepräsentiert. Hierzu besteht soziologischer Klärungsbedarf, da Früheinschulungen eine hohe Informiertheit und damit aufnahmelandsspezifisches Wissen voraussetzen, über welches zugewanderte Eltern nur begrenzt verfügen. Treffen Migranten Früheinschulungsentscheidungen losgelöst vom kindlichen Entwicklungsstand? Denkbar ist es, stellt man die hohen (bzw. überhöhten) Bildungsaspirationen von Migranten in Rechnung. Auch können herkunftslandspezifische Einschulungstraditionen oder eine Reaktion auf institutionelle Diskriminierungsakte (welche bei der zum regulären Einschulungstermin verbindlichen Einschulungsuntersuchung möglich sind) hinter diesem Befund stehen. Da Spät- wie Früheinschulungen unter dem Verdacht schulischer Ineffizienz stehen, könnte eine selektive Einschulung einen Erklärungsbeitrag für die schulischen Nachteile von (türkischstämmigen) Migrantenkindern liefern. Den aktuellen Forschungsstand zu den (potentiell) vermittelnden Größen zwischen ethnischer Zugehörigkeit auf der einen Seite und Schulerfolg auf der anderen Seite zusammenfassend, liegt derzeit weder zu den schulischen Folgen einer komplizierten Schwangerschafts- und Geburtshistorie, noch zur Erfolgsfrage von vorzeitig oder verspätet eingeschulerten Kindern empirisch belastbares Wissen vor. Ebenfalls kaum untersucht wurde die Frage, ob sich eine ethnisch selektive Nutzung der Früherkennungsuntersuchungen für Kinder (die sogenannten U-Untersuchungen) in ein höheres schulisches Misserfolgsrisiko übersetzt. Da diese Merkmale soziokulturell überformt sind, werden mit der sozioökonomischen Ressourcenausstattung und die im Elternhaus gesprochene Sprache zwei weitere für den Schulerfolg zentrale Faktoren ins theoretische Feld geführt. Der Fokus des Beitrags liegt dabei auf dem Primar- und Vorschulbereich: Den Lebensanfang, Vorschulalter und Schulbeginn in den Fokus rückend, geht der Beitrag der Frage nach, ob sich der geringe Grundschulerefolg von Migrantenkindern auf komplizierte und riskante Schwangerschaften, geringe vorschulische Humankapitalinvestitionen und eine ethnisch (und sozial) selektive Einschulungspraxis zurückführen lassen. Für eine quantitativ-empirische Analyse der aufgeworfenen Fragen bedarf es einer ausreichend großen Datenbasis, die eine unter ethnischen und sozialen Herkunftsaspekten differenzierte Analyse des Schulerfolgs erlaubt. Zudem müssen Informationen zu den hier betrachteten unabhängigen Variablen vorliegen. Der Kinder- und Jugend-Gesundheitssurvey (2003-2006) erfüllt diese Datenanforderungen, so dass dieser für die empirischen Analysen verwendet wird. Der Schulerfolg wird über die Noten in Deutsch und Mathematik sowie das Ereignis einer Klassenwiederholung operationalisiert. Neben der einheimischen Referenzgruppe beziehen sich die Theoriediskussion und Analysen auf die zwei größten Migrantengruppen in der Bundesrepublik: Türkischstämmige und (Spät-)Aussiedlerkinder. Die Ergebnisse bestätigen einerseits die Abhängigkeit lebensanfänglicher Nachteile von der ethnischen und sozialen Herkunft. Insbesondere türkischstämmige Kinder starten häufiger zu früh ins Leben, nehmen seltener an den Vorsorgeuntersuchungen und an Vorschulbetreuung teil. Auch lassen sich die bekannten ethnischen Unterschiede in der sozioökonomischen Ressourcenkomposition replizieren, wonach Kinder von (Spät-)Aussiedlern den einheimischen Kindern ähnlicher sind als jene mit türkischen Eltern. Andererseits zeigen sich die bekannten ethnischen Schulmisserfolgsmuster. Türkische Kinder bekommen die schlechtesten Noten und haben das größte Risiko einer Nichtversetzung in die nächsthöhere Klassenstufe der Grundschule. Multiple lineare Regressionen mit den Schulnoten als Kriteriumsvariablen sowie logistische Regressionen (mit der Klassenwiederholung als abhängiger Variablen) können die ethnischen Nachteile mit dem hier berücksichtigten Modellvariablen vollständig aufklären. Dabei erweisen sich sowohl die elterliche Bildungsnähe sowie die überwiegend im Alltag gesprochene Sprache als auch die frühen gesundheitlichen Nachteile, ethnisch selektive Vorschulinvestitionen und die Einschulungssituation als erklärungskräftig.



## ***Disparitäten im Bildungserfolg von Jugendlichen mit Migrationshintergrund in Deutschland: Gibt es tatsächlich Hinweise auf segmentierte Assimilation?***

Theoretischer Hintergrund: Zur Integration von Migrantinnen und Migranten wurden seit Anfang des 20. Jahrhunderts zahlreiche soziologische Theorien entwickelt. Unter ihnen betont die Theorie der segmentierten Assimilation (SAT; Portes & Zhou, 1993) die besondere Bedeutung der Generationenfolge und der zweiten Generation für den Integrationsprozess. Sie postuliert drei verschiedene Integrationsmuster in Abhängigkeit einer Vielzahl von möglichen Einflußfaktoren, darunter die familiäre Ausstattung mit ökonomischen, kulturellem und sozialen Kapital (Portes & Rumbaut, 2001):

- aufwärtsgerichtete Assimilation, d.h. die Angleichung der Zugewanderten auf zentralen Integrationsdimensionen an die Mittelschicht der Aufnahmegesellschaft, womit bei den häufig sozial schwächeren Migranten eine soziale Aufwärtsbewegung verbunden ist,
- abwärtsgerichtete Assimilation, d.h. die Angleichung an die Unterschicht der Aufnahmegesellschaft, und
- selektive Assimilation, d.h. die Angleichung auf einigen Integrationsdimensionen unter weitgehender Beibehaltung gruppenspezifischer Merkmale auf anderen Dimensionen.

Eine der zentralen Integrationsdimensionen stellt Bildung dar, insbesondere wenn sie mit institutionalisiertem kulturellem Kapital verbunden ist. Während die SAT früher vor allem auf die Situation von Migranten und ihren Nachkommen in den USA angewandt wurde, gibt es seit einigen Jahren auch Untersuchungen, die sie auf die Situation in den europäischen Ländern und auch Deutschland beziehen und dabei die unterschiedliche Ausstattung mit verschiedenen Kapitalsorten berücksichtigen (Alba & Waters, 2011; Kalter, 2011). Aus der erziehungswissenschaftlichen Bildungsforschung ist bislang häufig ein geringerer Bildungserfolg für verschiedene Migrantengruppen in Deutschland festgestellt worden als für Personen ohne Migrationshintergrund (z. B. Stanat, Rauch & Segeritz, 2010). Dies gilt insbesondere für Migrantinnen und Migranten aus den ehemaligen Anwerbestaaten Türkei, Italien und ehemaliges Jugoslawien. Für diese Gruppen, insbesondere türkischstämmige Migranten, wurde sogar der Prozess der abwärtsgerichteten Assimilation vermutet und mit einigen empirischen Befunden gestützt, während für Jugendliche aus Aussiedler- und Spätaussiedlerfamilien ein Muster der aufwärtsgerichteten Assimilation angenommen werden kann (Segeritz, Walter & Stanat, 2010).

Fragestellung: Vor diesem Hintergrund wird im Beitrag der Frage nachgegangen, ob für Jugendliche aus den größten Migrantengruppen in Deutschland (ehemalige Sowjetunion, Türkei, Polen, ehemaliges Jugoslawien und Italien) tatsächlich unterschiedliche Assimilationsmuster zu beobachten sind und insbesondere ein Muster der abwärtsgerichteten Assimilation für die Jugendlichen, deren Familien aus den ehemaligen Anwerbestaaten stammen. Darüber hinaus wird die differenzielle Bedeutung des ökonomischen, kulturellen und sozialen Kapitals der Familien für den Bildungserfolg der Jugendlichen über die Einwanderergenerationen hinweg untersucht.

Methode: Den Fragestellungen wird anhand der Test- und Fragebogendaten aus der Erweiterungsstichprobe von PISA 2003 (N = 44.473) mit 5 linearen Regressionsmodellen nachgegangen, die den Zusammenhang der Kapitalsorten mit den Bildungsdisparitäten erfaßt über die Lesekompetenz untersuchen. Die Stichprobe ermöglicht aufgrund eines Oversamplings von Jugendlichen mit Migrationshintergrund innerhalb der Herkunftsgruppen eine Differenzierung nach erster und zweiter Generation. Fehlende Daten wurden entweder durch Indikatorvariablen indiziert oder imputiert, so dass die gesamte Stichprobengröße für die Analysen zur Verfügung stand.

Ergebnisse: Die Ergebnisse bestätigen zwar den bekannten Befund, dass türkisch-, jugoslawisch- und italienischstämmige Jugendliche unter den fünf größten Gruppen im Durchschnitt den niedrigsten Bildungserfolg haben, sie weisen aber nicht darauf hin, dass es über verschiedene Generationen hinweg keine Verbesserungen im Bildungserfolg und damit Hinweise auf abwärtsgerichtete Assimilationsprozesse gibt. Außerdem weisen die Befunde darauf hin, dass das kulturelle Kapital der Familie eine etwas größere Bedeutung als ihr ökonomisches Kapital zu haben scheint, während die verwendeten Indikatoren des sozialen Kapitals praktisch keine Bedeutung für die migrationskorrelierten Disparitäten haben. Dies bedeutet erstens, dass Prozesse der Angleichung auf einer der zentralen Integrationsdimension selbst bei schwach integrierten Gruppen stattfinden, wenn auch wohl mit geringer Geschwindigkeit und zweitens dass möglicherweise Defizite im ökonomischen Kapital, die auf die Abwertung von Bildungsabschlüssen auf dem Arbeitsmarkt des Einwanderungslandes zurückzuführen sind, durch kulturelles Kapital kompensiert werden könnten. Dagegen ist die geringe Bedeutung des sozialen Kapitals in den Analysen sehr wahrscheinlich auf die verwendeten Indikatoren zurückzuführen, die lediglich die Anzahl der Familienmitglieder erfassen, die mit den Schülerinnen und Schülern zusammen leben.

Alba, R. M. & Waters, M. C. (2011). *The next generation. Immigrant youth in a comparative perspective*. New York City, NY: New York University Press.

Kalter, F. (2011). *The second generation in the German labor market: Explaining the Turkish phenomenon*. In R. M. Alba & M. C. Waters (Eds.), *The next generation: Immigrant youth in a comparative perspective* (pp. 166-184). New York City, NY: New York University Press.

Portes, A. & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation*. Berkeley, CA: University of California Press.

Portes, A. & Zhou, M. (1993). *The new second generation: Segmented assimilation and its variants*. *Annals of the American Academy of Political and Social Science*, 530, 74-96.

Stanat, P., Rauch, D. & Segeritz, M. (2010). *Schülerinnen und Schüler mit Migrationshintergrund*. In PISA-Konsortium Deutschland (Eds.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200-230). Münster: Waxmann.

Segeritz, M., Walter, O. & Stanat, P. (2010). *Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation?* *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62 (1), 113-138.



## **Besonderheiten in der Beratung von Eltern mit Migrationshintergrund im Schulalltag**

Theoretischer Hintergrund: Beratungsgespräche zwischen Eltern und Lehrkräften nehmen im Schulalltag einen wichtigen Stellenwert ein. Untersuchungen zur Beratungssituation an Schulen zeigen, dass Eltern sich Beratung durch Lehrpersonen wünschen (vgl. Krumm, 1996). Lehrpersonen fühlen sich jedoch oft nicht ausreichend auf die Beratungsarbeit im Schulalltag vorbereitet (vgl. Wild, 2003; Hertel, 2009). Darüber hinaus stimmen das Beratungsangebot der Lehrpersonen und der Beratungsbedarf der Eltern nicht immer überein (vgl. Hertel & Schmitz, 2010). Aktuelle Befunde aus PISA 2009 weisen darauf hin, dass Eltern mit Migrationshintergrund durch die bestehenden Beratungsangebote an Schulen in Deutschland nicht gut erreicht werden (Hertel et al., 2011).

Fragestellung: Der Beitrag zielt darauf ab, die Beratungssituation an Grundschulen aus der Perspektive von Lehrern und Eltern zu beschreiben und einen binationalen Vergleich der Befunde für Schulen in Deutschland (Ballungsraum Rhein-Main) und der Schweiz (Kanton Zürich) vorzunehmen. Im Fokus stehen Vergleiche zwischen Eltern mit und ohne Migrationshintergrund zu folgenden Fragestellungen: (1) Wie gut fühlen sich Lehrpersonen durch ihre Ausbildung auf die Elternberatung vorbereitet? (2) Wie nutzen Eltern mit Migrationshintergrund das Beratungsangebot der Schule? (3) Welche Besonderheiten ergeben sich in Bezug auf den Beratungsbedarf von Eltern mit Migrationshintergrund? (4) Wie unterscheidet sich der Beratungsbedarf der Eltern vom Beratungsangebot der Lehrer?

Methode: Die Datenerhebungen für das Projekt elbe fanden parallel im Großraum Frankfurt und im Kanton Zürich statt. Grundlage der Analysen bildet eine Befragung von Eltern und Lehrkräften an Grundschulen. Für Deutschland liegen insgesamt Daten von N=2005 Eltern und N=149 Lehrpersonen von 20 Grundschulen vor. Der Anteil von Eltern mit Migrationshintergrund (mind. ein Elternteil im Ausland geboren) liegt bei 37,1 %. In der Schweiz fand eine Eltern-Lehrerbefragung an 5 Primarschulen im Kanton Zürich statt. Es liegen Daten von N=303 Eltern und N=35 Lehrpersonen vor. Der Anteil von Eltern mit Migrationshintergrund (mind. ein Elternteil im Ausland geboren) liegt bei 27,1 %.

Ergebnisse und Diskussion: Die Analysen zu Fragestellung (1) zeigen, dass sich Lehrpersonen in der Schweiz durch ihre Ausbildung signifikant besser auf die Beratung von Eltern vorbereitet fühlen, als Lehrpersonen in Deutschland. Für Fragestellung (2) ergab sich, dass in Deutschland das Beratungsangebot der Schule häufiger von Eltern ohne Migrationshintergrund in Anspruch genommen wird als von Eltern mit Migrationshintergrund. In der Schweiz hingegen nutzen Eltern mit Migrationshintergrund signifikant häufiger weitere Beratungsangebote. Die Analysen zu Fragestellung (3) zeigten, dass in beiden Ländern der Wunsch nach mehr Beratungsangeboten durch die Klassenlehrperson bei Eltern mit Migrationshintergrund höher ausfällt, als bei Eltern ohne Migrationshintergrund. Hinsichtlich des Ländervergleichs ergab sich, dass der Beratungsbedarf insgesamt betrachtet bei Eltern aus Deutschland mit 42,5% höher liegt als bei Eltern aus der Schweiz (17,0%). Eltern aus Deutschland sind demnach weniger zufrieden mit der Beratungssituation an der Schule ihres Kindes und wünschen sich mehr Angebote durch die Klassenlehrperson.

Die Ergebnisse haben wichtige Implikationen für die Gestaltung der Lehreraus- und Weiterbildung sowie für Schulentwicklungsprozesse.

Hertel, S. (2009): Beratungskompetenz von Lehrern. Kompetenzdiagnostik, Kompetenzförderung und Kompetenzmodellierung. Münster: Waxmann.

Hertel, S./Schmitz, B. (2010): Der Lehrer als Berater in Schule und Unterricht. Stuttgart: Kohlhammer.

Hertel, S./Bruder, S./Jude, N./Steinert, B. (2013): Elternberatung an Schulen im Sekundarbereich: Schulische Rahmenbedingungen, Beratungsangebote der Lehrkräfte und Nutzung von Beratung durch die Eltern. In: Zeitschrift für Pädagogik 59, S. 40-63.

Krumm, V. (1996): Über die Vernachlässigung der Eltern durch die Erziehungswissenschaft. Plädoyer für eine veränderte Rolle der Lehrer bei der Erziehung der Kinder. In: Leschinsky, A. (Hrsg.): Die Institutionalisierung von Lehren und Lernen. Beiträge zu einer Theorie der Schule. In: Zeitschrift für Pädagogik, 34. Beiheft, Weinheim: Beltz, S. 119-137.

Wild, E. (2003): Lernen lernen: Wege einer Förderung der Bereitschaft und Fähigkeit zum selbstreguliertem Lernen. In: Unterrichtswissenschaften 31, S. 2-5.

## **Immigrationsbezogene Disparitäten in den Kompetenzen von Kindergartenkindern**

Hintergrund und Fragestellung: Sowohl Ergebnisse als auch der IGLU Studie weisen auf einen substantiellen Leistungsunterschied zwischen Kindern mit und ohne Immigrationshintergrund hin. Bei der Frage nach der Erklärung dieser Unterschiede verweisen viele Bildungssoziologen auf Boudon's (1974) Unterscheidung von primären und sekundären Effekten der sozialen Herkunft, um die Beziehung zwischen sozialer Herkunft und schulischen Leistungen zu beschreiben. Primäre Effekte beziehen sich dabei auf Kompetenzunterschiede, die mit (sozialen) Unterschieden in den (ökonomischen, kulturellen und sozialen) Ressourcen der Eltern einhergehen, während sich sekundäre Effekte bei Bildungsentscheidungen an Übergängen im Bildungssystem zeigen. In den letzten Jahren wurde die Unterscheidung primärer und sekundärer Effekte sozialer Herkunft um Effekte der ethnischen Herkunft erweitert, um soziale Bildungsungleichheiten zwischen Familien mit und ohne Immigrationshintergrund differenzierter zu betrachten (van de Werfhorst & van Tubergen, 2007; Heath et al., 2008; Kristen & Dollmann, 2009, 2012). Ethnische primäre Effekte können dabei als Ressourcen von Immigranten verstanden werden, die im Aufnahmeland nicht eingesetzt werden können (z.B. Beherrschen der Herkunftssprache). Bei der Betrachtung ethnischer Effekte ist deshalb die Berücksichtigung der Aneignung von Ressourcen, die im Aufnahmeland eingesetzt werden können (z.B. Beherrschen der Verkehrssprache) von zentraler Bedeutung. Während immigrationsbezogene Bildungsungleichheiten für Schülerinnen und Schüler in Deutschland gut dokumentiert sind (z. B. Walter & Taskinen, 2007; Schwippert et al., 2007), sind vergleichbare Studien mit Kindern im Kindergartenalter sehr selten. Wir verwenden Daten des Nationalen Bildungspanels (NEPS) und präsentieren erste Ergebnisse immigrationspezifischer Ungleichheiten in sprachlichen und naturwissenschaftlichen Kompetenzen von Kindergartenkindern. Folgende Forschungsfragen sind dabei leitend: (1) Wie stark sind die Unterschiede in den sprachlichen und naturwissenschaftlichen Kompetenzen ausgeprägt und (2) welche Faktoren erklären die immigrationsbezogenen Kompetenzunterschiede? Im Anschluss an theoretische Betrachtung primärer Effekten der sozialen Herkunft und primärer Effekte ethnischer Herkunft, lauten unsere Hypothesen: (1) Leistungslücken zwischen Immigranten und Autochthonen lassen sich überwiegend durch Unterschiede des sozialen Hintergrunds erklären und (2) zumindest für bestimmte Gruppen von Immigranten, sollten ethnische primäre Effekte (z.B. elterlicher Gebrauch der Herkunftssprache) einen zusätzlichen Beitrag zur Erklärung der Kompetenzunterschiede leisten.

Daten und Methoden: Unsere Analysestichprobe umfasst N = 2.138 ca. fünf Jahre alte Kinder mit Elternangaben, die N = 279 Kindergärten besuchen. Sprachliche und naturwissenschaftliche Kompetenzen (skaliert in Prozent korrekt gelöster Aufgaben) bilden die abhängige Variable. Wir konzentrieren uns auf Immigranten der zweiten Generation und unterscheiden zwischen Immigranten türkischer und russischer Herkunft. Als Indikator der elterlichen sozialen Klasse verwenden wir das EGP-Schema, sowie für den elterlichen Bildungshintergrund die CASMIN Skala. Die Häufigkeit gemeinsamer (Bildungs-) Aktivitäten wird als Indikator für die häusliche Lernumwelt verwendet, sowie weitere Variablen (wie z. B. die zuhause gesprochene Sprache, Teilnahme an Sprachförderung im Kindergarten, Familienstruktur und die Anzahl der Geschwister) in Analysen miteinbezogen.

Unsere Analysen basieren auf OLS Regression mit auf Kindergartenebene geclusterten Standardfehlern, in deren Rahmen wir schrittweise Gruppen unabhängiger Variablen einfügen. Ziel ist dabei die Veränderung der Differenzen über die verschiedenen Regressionsmodelle darzustellen.

Ergebnisse und Diskussion: Bereits im Alter von fünf Jahren finden wir stark ausgeprägte Differenzen, sowohl in sprachlichen als auch in naturwissenschaftlichen Kompetenzen. Vor allem Kinder türkischer Herkunft liegen weit zurück. Primäre Effekte sozialer Herkunft sollten eine vergleichsweise große Bedeutung für Immigranten in Deutschland spielen, da viele Einwanderer aufgrund ihrer Geschichte als Gastarbeiter in Deutschland hinsichtlich ihrer sozialen Herkunft (im Verhältnis zum Durchschnitt des Aufnahmelandes) negativ selektiert sind. Allerdings sehen wir auch erhebliche primäre ethnische Effekte, zumindest bei Unterschieden in der Sprachkompetenz von Kindern, während Familienstruktur, die häusliche Lernumwelt und die betrachteten Kindergartenmerkmale die Kompetenzunterschiede nur geringfügig beeinflussen.

Boudon, R. (1974). Education, Opportunity, and Social Inequality. Changing Prospects in Western Society. New York: Wiley & Sons.

Heath, A., & Birnbaum, Y. (2007). Explaining Ethnic Inequalities in Educational Attainment. *Ethnicities*, 7, 291-304.

Kristen, C. & Dollmann, J. (2009). Sekundäre Effekte der ethnischen Herkunft? Kinder aus türkischen Familien am ersten Bildungsübergang. *Zeitschrift für Erziehungswissenschaft, Sonderheft 12*, 205-229.

Kristen, C. & Dollmann, J. (2012). Migration und Schulerfolg: Zur Erklärung ungleicher Bildungsmuster. In: M. Matzner (Hrsg.), *Handbuch Migration und Bildung*. Weinheim/Basel: Beltz, 102-117.

Schwippert, K.; Hornberg, S.; Freiberg, M.; & Stubbe, T. C. (2007). Lesekompetenzen von Kindern mit Migrationshintergrund im internationalen Vergleich. In: Bos, W.; Hornberg, S.; Arnold, K.-H., Faust, G.; Fried, L.;

Lankes, E.-M.; Schwippert, K. & Valtin, R. (Hrsg.), IGLU 2006. Lesekompetenzen von Grundschulkindern im internationalen Vergleich. Münster: Waxmann, 249-268.

van De Werfhorst, H.-G. & van Tubergen, F. (2007). Ethnicity, schooling, and merit in the Netherlands. *Ethnicities*, 7, 416-444.

Walter, O. & Taskinen, P. (2007). Kompetenzen und bildungsrelevante Einstellungen von Jugendlichen mit Migrationshintergrund in Deutschland: Ein Vergleich mit ausgewählten OECD-Staaten. In: PISA Konsortium Deutschland (Hrsg.), *PISA '06. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann, 337-365.



## **Soziale Selektivität beim Studieneinstieg von Abiturienten nach einer beruflichen Ausbildung**

Ein beachtlicher Anteil der Abiturienten beginnt nach Erwerb der Hochschulreife zunächst eine nicht-tertiäre berufliche Ausbildung. Von denjenigen, die eine Ausbildung beginnen, steigen jedoch später einige noch in ein Studium ein. In dem Beitrag geht es um die Frage, welche Faktoren die Neigung, nach Beginn einer Ausbildung ein Studium aufzunehmen, erhöhen. Insbesondere wird untersucht, welche Rolle die soziale Herkunft, gemessen anhand des höchsten Bildungsniveaus der Eltern, spielt. Zunächst werden die Tätigkeiten, denen Personen nach Erwerb der Hochschulreife nachgehen, beschrieben und Unterschiede nach sozialer Herkunft herausgearbeitet. Daraufhin wird der Einfluss der sozialen Herkunft auf die Studienneigung nach Beginn einer Ausbildung untersucht. Dabei wird eine Reihe weiterer Einflussfaktoren schrittweise kontrolliert und die Zeitabhängigkeit der Einstiegsneigung berücksichtigt.

Den theoretischen Rahmen bildet erstens die Lebenslaufperspektive. Auf ihrer Grundlage wird argumentiert, dass Lebensphasen mit bestimmten Rollen verbunden sind, die auf Altersdefinitionen beruhen. Außerdem wird angenommen, dass aufgrund der steigenden Unabhängigkeit von den Eltern der Herkunftseffekt mit zunehmendem Alter nachlässt (Lebenslaufhypothese). Darüber hinaus wird die Pfadabhängigkeit der Ereignisse im Lebenslauf berücksichtigt, indem untersucht wird, inwiefern sich Merkmale der Schullaufbahn auf die Studienneigung nach Beginn einer Ausbildung auswirken. Als weitere theoretische Grundlage werden soziologische Theorien der Bildungsentscheidung, die auf der Annahme der rationalen Wahl basieren, herangezogen. Hier wird konkurrierend zur Lebenslaufhypothese die Hypothese aufgestellt, dass die soziale Herkunft auch in späten Bildungsphasen von großer Bedeutung ist, insbesondere aufgrund drohender Kosten des Statusverlusts. Die Analysen bauen auf Studien zur Entwicklung des Herkunftseffekts über den Lebenslauf und zu Bildungsungleichheit in späten Phasen des Bildungssystems auf.

Zur deskriptiven Darstellung der Tätigkeitsverläufe nach Erwerb der Hochschulreife wird die Methode der Sequenzanalyse herangezogen. Diese ermöglicht, große Mengen individueller Verläufe graphisch darzustellen und Muster in den Verläufen zu erkennen. Die multivariate Analyse wird anhand einer Ereignisanalyse ausgeführt. Um die Zeitabhängigkeit der Studienneigung und ihrer Einflussfaktoren zu berücksichtigen, wird ein periodenspezifisches Exponentialmodell mit zeitabhängigen Variablen und periodenspezifischen Effekten angewendet.

Die Analysen basieren auf den Daten der Startkohorte 6 des Nationalen Bildungspanel (NEPS): „Bildung im Erwachsenenalter und lebenslanges Lernen“. Diese bieten durch die retrospektive Erfassung der Lebensverläufe die Möglichkeit, den Übergang nach Erwerb der Hochschulreife in ein Studium monatsgenau im Längsschnitt zu erfassen und beispielsweise Merkmale der Schullaufbahn zu berücksichtigen.

Die Ergebnisse zeigen, dass Personen, deren Eltern einen Hochschulabschluss haben, stärker dazu neigen, nach einer Ausbildung noch ein Studium zu beginnen. Dieser Effekt der sozialen Herkunft bleibt auch unter Kontrolle aller berücksichtigten weiteren Einflussfaktoren bestehen. Damit bestätigen die Befunde Studien, die eine anhaltende Bedeutung des Herkunftseffekts bei späten Bildungsentscheidungen finden. Weiterhin sprechen die Ergebnisse dafür, die Zeitabhängigkeit des Übergangs in ein Studium zu berücksichtigen, da sich die Studienneigung und die Bedeutung der Einflussfaktoren über die Zeit verändert. Darüber hinaus geben die Befunde Hinweise darauf, dass Unterschiede in der Schullaufbahn zusätzlich zum Schulabschluss einen weitreichenden Effekt auf spätere Bildungswege haben.

Die Studie zielt darauf ab, den Stand der Forschung zu späten Bildungsübergängen in den folgenden Punkten zu erweitern. Erstens werden, soweit anhand der Daten operationalisierbar, einschlägige Theorien der Bildungsforschung auf ihren Erklärungsgehalt für den untersuchten Übergang geprüft. Zweitens wird der Zeitaspekt des Übergangs untersucht, indem die Studienneigung und deren Einflussfaktoren im Längsschnitt betrachtet sowie vorausgehende Ereignisse des Lebenslaufs berücksichtigt werden. Drittens wird die Geltung bisheriger Befunde anhand eines neuen großen Längsschnittdatensatzes geprüft.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823





## ***Der Zusammenhang von Bildungsherkunft mit Leistungsmotivation, Studieninteresse und akademischem Selbstkonzept im Studium sozialwissenschaftlicher Fächer***

Laut Bildung in Deutschland (2010) nehmen „Studienberechtigte, von denen kein Elternteil bereits ein Hochschulstudium abgeschlossen hat,(...) auch bei gleicher Schulleistung seltener ein Studium auf.“ Die US-amerikanische Forschung zeigt, dass diese sogenannten Erstakademiker, sofern sie dennoch ein Studium aufnehmen, häufig zusätzlichen Problemen gegenüberstehen (z.B. Stephens et al., 2012). Im Gegensatz zum US-amerikanischen Forschungsraum, in dem die Frage von Erstakademikern und höherer Bildung derzeit zunehmend diskutiert wird (z.B. Aspelmeier et al., 2012; Hellman, 1996; Ishitani, 2006), gibt es für den Deutschen Forschungsraum bislang nur wenig Befunde zu Unterschieden zwischen Erstakademikern und solchen mit mindestens einem Akademikerelternteil. Ziel der vorliegenden Studie war daher die Analyse des Zusammenhangs von Bildungsherkunft mit Studieninteresse, Leistungsmotivation und akademischem Selbstkonzept an einer deutschen Stichprobe von Studierenden sozialwissenschaftlicher Fächer im Rahmen des Projekts „WiKom-Sowi: Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern“ (Zhang et al., 2013). In Bezug auf Studieninteresse und Leistungsmotivation wäre denkbar, dass Erstakademiker, die sich bewusst für ein Studium entschieden haben, eine höhere Leistungsmotivation und ein höheres Studieninteresse aber ein geringeres akademisches Selbstkonzept aufweisen als Studierende mit mindestens einem Akademikerelternteil. Daten wurden mittels Fragebögen im Rahmen der ersten Erhebung des längsschnittlich angelegten Projekts erhoben. Neben der Bildungsherkunft wurden das Leistungsmotivationsinventar in der Kurzversion (LMI-K), der Fragebogen zum Studieninteresse (FSI) und die Subskala „absolut“ der Skalen zur Erfassung des schulischen Selbstkonzepts (SESSKO-A, Studierendenversion) erhoben. Für die Auswertung wurden Daten von N = 299 Bachelor- und Masterstudierenden sozialwissenschaftlicher Fächer an zwei Standorten betrachtet (davon in Heidelberg N = 46 Studierende der Psychologie, N = 62 Studierende der Soziologie, N = 52 Studierende der Politikwissenschaft; in Saarbrücken N = 139 Studierende der Psychologie). Bildungsherkunft wurde sowohl als eine dichotome Variable gefasst [Bildungsherkunft (Eltern), mindestens ein Elternteil besitzt einen Studienabschluss vs. kein Elternteil besitzt einen Studienabschluss], als auch als zwei eigenständige dichotome Variablen [Bildungsherkunft (Mutter) bzw. Bildungsherkunft (Vater), jeweils Studienabschluss vs. kein Studienabschluss], welche den Bildungsabschluss eines Elternteils unabhängig vom Bildungsabschluss des anderen Elternteils in den Fokus der Analysen rücken. Für die Teilstichproben ergab sich in Bezug auf die Standorte und Fächer ein heterogenes Bild bezüglich der Bildungsherkunft: in der Saarbrücker Teilstichprobe (nur Psychologie) lag der Anteil von Erstakademikern bei 50%. In der Heidelberger Teilstichprobe lag der Anteil in allen Fächern bei 33%. Eine separate Betrachtung der Fächer führte zu einem Erstakademikeranteil von 52% in der Soziologie, 28% in der Psychologie und 16% in der Politikwissenschaft. Betrachtet man Bildungsherkunft als elterliche Variable, so lässt sich mittels linearer Regression keine Varianz in Bezug auf Studieninteresse, Leistungsmotivation oder akademischem Selbstkonzept dadurch aufklären. Zieht man jedoch für die Bildungsherkunft die mütterliche und die väterliche Bildung getrennt heran, ergeben sich für die Studierenden Unterschiede in Bezug auf zwei Subskalen des Studieninteresses (FSI, Subskalen „Intrinsischer Charakter“ sowie „Persönliche wertbezogene Valenzen“) sowie in Bezug auf eine Subskala der Leistungsmotivation (LMI-K, Subskala „Flow“). Die Bildungsherkunft (Vater) klärt in der Gesamtstichprobe 1,5% der Varianz der Subskala „Persönliche wertbezogene Valenzen“ auf. Die Bildungsherkunft (Mutter) klärt in der Gesamtstichprobe 1,3% der Varianz der Subskala „intrinsischer Charakter“ und 2,2% der Varianz der Subskala „Flow“ auf. Für das akademische Selbstkonzept kann keine Varianz durch die väterliche oder mütterliche Bildung aufgeklärt werden. Die vorliegende Studie gibt Anhaltspunkte für Unterschiede im Anteil der Erstakademiker an verschiedenen Standorten und in verschiedenen sozialwissenschaftlichen Studienfächern. Die Ergebnisse deuten darauf hin, dass die Bildungsherkunft auch getrennt nach väterlicher und mütterlicher Bildung erfasst werden sollte. Insgesamt konnten nur geringe Varianzanteile in der Leistungsmotivation und dem Studieninteresse durch die väterliche und mütterliche Bildung aufgeklärt werden, was dafür spricht, dass in der vorliegenden Stichprobe kaum Unterschiede hinsichtlich Leistungsmotivation, Studieninteresse und akademischem Selbstkonzept zwischen Erstakademikern und Studierenden mit mindestens einem Akademikerelternteil bestehen.

- Aspelmeier, J. E., Love, M. M., McGill, L. A., Elliott, A. N., & Pierce, T. W. (2012). Self-esteem, locus of control, college adjustment, and GPA among first- and continuing-generation students: A moderator model of generational status. *Research in Higher Education*, 53, 755-781.
- Autorengruppe Bildungsberichterstattung (2010). *Bildung in Deutschland 2010: Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demografischen Wandel*. Zugriff am 29.10.2013, von [http://www.bildungsbericht.de/daten2010/f\\_web2010.pdf](http://www.bildungsbericht.de/daten2010/f_web2010.pdf)
- Hellman, C. M. (1996). Academic self-efficacy: Highlighting the first-generation student. *Journal of Applied Research in the Community College*, 4, 69-75.
- Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education*, 77, 861-885.
- Stephens, N. M., Fryberg, S. A., Markus, H., Johnson, C. S., & Covarrubias, R. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal Of Personality And Social Psychology*, 102(6), 1178-1197.
- Zhang, Y., Klopp, E., Dietrich, H., Brünken, R., Krause, U.-M., Spinath, B., Stark, R., & Spinath, F. M. (2013). Modeling and Measuring Scientific Competencies in Social Sciences. In S. Blömeke & O. Zlatkin-Troitschanskaia (Eds.), *The German funding initiative "Modeling and Measuring Competencies in Higher Education": 23 research projects on engineering, economics and social sciences, education and generic skills of higher education students (KoKoHs Working Papers, 3)*. Berlin & Mainz: Humboldt University & Johannes Gutenberg University.



## ***Persönlichkeitseigenschaften und Bildungserfolge entgegen aller Erwartungen***

Theoretischer Hintergrund: Schülerinnen und Schüler aus Familien ohne akademischen Hintergrund sind im tertiären Bildungsbereich nach wie vor unterrepräsentiert. Dies trifft sogar in einem Bildungssystem zu, in dem Hochschulbildung frei von direkten Kosten (Studiengebühren) für Studierende ist und finanzielle Beschränkungen eine eher untergeordnete Rolle beim Übergang in die tertiäre Bildung spielen sollten.

Viele theoretische und empirische Studien analysieren Ursachen für diese Ungleichheit. Neben primären und sekundären Herkunftseffekten, kann z.B. die kulturelle Distanz der Schülerinnen und Schüler zur Hochschule und damit verbundene Hemmungen der Studienaufnahme eine Rolle spielen. Darüber hinaus können Informationsdefizite Schülerinnen und Schüler von der Aufnahme eines Studiums abhalten.

Fragestellung und Methode: Diese Erklärungsfaktoren stehen im Zusammenhang mit Persönlichkeitsmerkmalen bzw. können durch bestimmte Persönlichkeitsmerkmale ausgeglichen werden. Auf Basis der Daten des Sozio-oekonomischen Panels zeigen wir, welche Persönlichkeitsmerkmale im Zusammenhang mit der Aufnahme eines Hochschulstudiums stehen und wie sich diese zwischen Kindern aus akademischen und nicht-akademischen Haushalten unterscheiden. Während es für Schülerinnen und Schüler die z.B. sehr ängstlich sind schwieriger ist die Distanz zur Universität zu überwinden, kann es für Ihre neugierigeren Schulkameraden leichter sein diese Hürde zu nehmen. Während wir den elterlichen Hintergrund und kognitive Fähigkeiten konstant halten, zeigen wir durch die Anwendung der Propensity Score Matching Methode, welche Persönlichkeitsmerkmale einen Studienbeginn erklären. Ergebnisse und Diskussion: Erste Ergebnisse zeigen, dass „Offenheit für Erfahrung“ die Absicht von Schülerinnen und Schüler sich für ein Hochschulstudium zu entscheiden, erhöht. Dies gilt nur für Schülerinnen und Schüler aus nicht-Akademiker Familien. Für Schülerinnen und Schüler mit einem akademischen Hintergrund verringert Ängstlichkeit und Unsicherheit die Studienwahrscheinlichkeit. Die Ergebnisse implizieren, dass bildungspolitische Maßnahmen die Stärkung von nicht-kognitiven Fähigkeiten der Schülerinnen und Schüler nicht vernachlässigen sollten. Über die gezielte Förderung bestimmter Persönlichkeitsmerkmale können Schülerinnen und Schüler aus bildungsfernen Familien an die Hochschulbildung herangeführt werden.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## Genderstereotype bei österreichischen Lehrkräften

Die Frage nach der Genderfairness des Bildungssystems ist trotz vieler positiver Entwicklungen in den letzten Jahren nicht obsolet: Es finden sich noch immer zahlreiche Benachteiligungen für beide Geschlechter, die sich in nicht genutzten Chancen und eingeschränkten Handlungsspielräumen in individuellen Bildungs- und Berufskarrieren niederschlagen (z.B. Hannover & Kessels, 2011; Europäische Kommission, 2009). Die Ursachen für die Entwicklung geschlechtsspezifischer Bildungskarrieren sind vielfältig (z.B. Spiel, Schober, & Finsterwald, 2011), unstrittig ist jedoch die zentrale Rolle der Sozialisation. Das Handeln von SozialisationsagentInnen wird oft massiv durch Genderstereotype – also durch Annahmen darüber, wie Mädchen bzw. Buben sind und zu sein haben – beeinflusst. Dies führt zu negativen Kreisprozessen: Buben und Mädchen verhalten sich den stereotypen Erwartungen entsprechend, was Konsequenzen für Leistung, Motivation und Berufswahl nach sich zieht (z.B. Dresel, Heller, Schober, & Ziegler, 2001; Keller, 2001; Rustemeyer 1999). Systematische Untersuchungen dazu wie Genderstereotype bei österreichischen Lehrkräften ausgeprägt sind fehlen jedoch weitgehend. Im Fokus der vorliegenden Studie stehen daher Genderstereotype österreichischer Lehrkräfte sowie zwei damit in Zusammenhang stehende Konstrukte, die basierend auf Erkenntnissen aus anderen Forschungskontexten als mögliche Ansatzpunkte zur Reduktion von Stereotypen gesehen werden könnten: (1) Wissen über Gender sowie (2) Implizite Theorien über die Veränderbarkeit von Persönlichkeitseigenschaften. So führt – laut Studien aus dem Bereich der Rassismusforschung – erhöhtes Wissen zu einer Reduktion von Vorurteilen, weil durch den Zugang zu alternativen Denkstrukturen der Zugriff auf das automatisch aktivierte stereotype Denken unterbrochen werden kann (Jackson, 2011). Mehr Wissen über Gender sollte also zu geringer ausgeprägten Genderstereotypen führen (Gocłowska & Crisp, 2012). Unter der Impliziten Theorie über die Veränderbarkeit von Persönlichkeitseigenschaften werden Überzeugungen über die Stabilität bestimmter Merkmale verstanden, die sowohl die Motivation als auch das Handeln von Personen beeinflussen (Dweck, 2000). Personen, die Persönlichkeitseigenschaften als eher formbar und veränderbar ansehen (Veränderbarkeitstheorie) neigen weniger zu Stereotypisierung (z.B. Ziegler 1999) als Personen mit einer stabilen Sicht von Persönlichkeit (Unveränderbarkeitstheorie).

Vor diesem Hintergrund wurden folgende Fragestellungen untersucht: 1. Wie sind die Genderstereotype bei Lehrkräften in Österreich ausgeprägt? 2. Lassen sich die Befunde zu Wissen und Impliziten Theorien aus anderen Forschungskontexten auch für Genderstereotype replizieren? Konkret wurde dabei die Annahme überprüft, dass sich Lehrkräfte mit unterschiedlich ausgeprägten Genderstereotypen in ihrem Wissen über Gender und in ihren Impliziten Theorien über die Veränderbarkeit von Geschlechtsunterschieden unterscheiden. Zur Beantwortung der Fragestellungen wurde 262 österreichischen Lehrkräften (71.2% weiblich) ein Online-Fragebogen vorgegeben. Genderstereotype wurden mittels Einschätzung (a) der Berufseignung für Mädchen und Buben für verschiedene Berufe sowie (b) des Anteils an Mädchen, die in Mathematik als begabt angesehen werden, erhoben (Indexbildung angelehnt an Dresel et al., 2001). Zur Erfassung des Wissens über Gender bearbeiteten die Lehrkräfte einen Multiple-Choice Test. Die Implizite Theorie über die Veränderbarkeit von Geschlechtsunterschieden im Leistungskontext wurde über drei an Dweck, Chiu und Hong (1995) angelehnte Items erhoben. Die Ergebnisse weisen auf genderstereotypes Denken bei österreichischen Lehrkräften hin: 62,8% der Lehrkräfte zeigten stereotypes Denken. Der Wissensstand zu Gender ist eher niedrig, im Mittel wurden nur 4,35 der 10 Fragen korrekt beantwortet. Genderstereotypes Denken der Lehrkräfte korrespondiert dabei mit geringerem Wissen über Gender und einer stärker ausgeprägten Unveränderbarkeitstheorie über Geschlechtsunterschiede im Leistungskontext.

Die Ergebnisse machen den Handlungsbedarf im Kontext Schule evident: Da Genderstereotype das (Unterrichts-)Verhalten von Lehrpersonen und in weiterer Folge Motivation, Leistung und Berufswahl von SchülerInnen beeinflussen, ist es für die Herstellung der Chancengleichheit für Buben und Mädchen zentral, diese Thematik in die Lehrkräfteaus- und -weiterbildung zu transportieren. Der Aufbau von Wissen und die Reflexion der impliziten Theorien über die Veränderbarkeit von Geschlechtsunterschieden im Leistungskontext scheinen dabei sinnvolle Ansatzpunkte zu sein.

Dresel, M., Heller, K. A., Schober, B., & Ziegler, A. (2001). Geschlechtsunterschiede im mathematisch-naturwissenschaftlichen Bereich. In C. Finkbeiner & G. W. Schnaitmann (Eds.), *Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik* (S. 270–288). Donauwörth: Auer.

Dweck, C. S. (2000). Self-theories: Their role in motivation, personality, and development. Lillington: Taylor & Francis.

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry*, 6, 322–333.

Europäische Kommission. (2009). *Geschlechterunterschiede bei Bildungsergebnissen: Derzeitige Situation und aktuelle Maßnahmen in Europa*. Brüssel: Eurydice.

Gocłowska, M., & Crisp, R. (2012). On counter-stereotypes and creative cognition: When interventions for reducing prejudice can boost divergent thinking. *Thinking Skills and Creativity*, 872–79.

Hannover, B., & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25, 89–103.

Jackson, L. M. (2011). *The Psychology of Prejudice: From Attitudes to Social Action*. Washington: American Psychological Association.

Keller, C. (2001). Effect of Teachers' Stereotyping on Students' Stereotyping of Mathematics as a Male Domain. *The Journal of Social Psychology*, 141, 165–173.

Rustemeyer, R. (1999). Geschlechtstypische Erwartungen zukünftiger Lehrkräfte bezüglich des Unterrichtsfaches Mathematik und korrespondierende (Selbst-) Einschätzungen von Schülerinnen und Schülern. *Psychologie in Erziehung und Unterricht*, 46, 187–200.

Spiel, C., Schober, B., & Finsterwald, M. (2011). Frauen und Technik - Warum kommen sie so schwer zusammen? *Konstruktiv*, 282, 8–9.

Ziegler, A. (1999). Das Motivationsprozeßmodell von C. Dweck aus pädagogisch-psychologischer Perspektive. In Bettina Hannover, U. Kittler, & H. Metz-Göckel (Eds.), *Sozialkognitive Aspekte der Pädagogischen Psychologie*. Band 1: Dokumentation des 3. Dortmunder Symposions für Pädagogische Psychologie 1998 (S. 164–177). Essen: Die Blaue Eule.



## **Berufsbildungsentscheidungen: Die Berufsmaturität in der beruflichen Bildung**

In der Bildungsforschung spielen Fragen der beruflichen Bildung eine wesentliche Rolle. Die hohe Attraktivität des Berufsbildungssystems der Schweiz zeigt sich darin, dass etwa 2/3 der Jugendlichen ihren Bildungsweg in der Berufsbildung fortsetzen. Während im Mittelpunkt der beruflichen Grundbildung der Erwerb berufsspezifischer Qualifikationen steht bietet das Berufsbildungssystem zahlreiche Möglichkeiten der Weiterbildung, wobei je nach Ausbildungsmodell die Allgemeinbildung unterschiedlich stark integriert wird. Der Verknüpfung von Ausbildungs- und Beschäftigungssystem wird durch Weiterentwicklungen der Weiterbildungsmöglichkeiten Rechnung getragen. Eine zentrale Innovation war die Einführung der Berufsmaturität (Gonon, 1994), die es jungen Erwachsenen ermöglicht, nach der Berufslehre ihren Bildungsweg im tertiären Bildungsbereich fortzusetzen. Ziel war zudem, die Berufsbildung zu stärken und „schulbegabte“ Jugendliche dafür zu interessieren sowie Berufs- und Hochschulbildungssystem durchlässiger zu gestalten.

Da einerseits von den Jugendlichen eine berufliche Entscheidung getroffen werden muss, andererseits selektive Mechanismen des Arbeitsmarktes wirken, wird aus strukturell-individualistischer Perspektiv davon ausgegangen, dass unterschiedliche Bildungschancen von Jugendlichen aggregierte Ergebnisse sozial divergierender Berufsvorstellungen, herkunftsabhängiger Leistungsvoraussetzungen sowie selbstselektiver Prozesse beim Eintritt in die nachobligatorischen Ausbildung sind. Ausgangspunkt sind die handlungstheoretischen Überlegungen von Boudon (1974), Breen und Goldthorpe (1997) und Esser (1999). Demnach sind Bildungsentscheidungen als Resultat individueller Abwägungen von Alternativen zu verstehen, die unter Berücksichtigung der subjektiv erwarteten Kosten, Nutzen und Erfolgswahrscheinlichkeiten hinsichtlich der zu erwartenden Erträge eingeschätzt werden. Erwartete Bildungserträge können mittels der Theorie der sozialen Produktionsfunktion (Lindenberg, 1989) spezifiziert werden. Berufslehre und Berufsmaturität sind aus Sicht der Jugendlichen als sekundäres Zwischengut zu verstehen, das ermöglicht ein weiteres Zwischengut zu erwerben: einen bestimmten Beruf ausüben zu können bzw. der Zugang zur tertiären Bildung. Dieses Bildungspatent ist wiederum ein Zwischengut um begehrte Ziele durch einen erfolgreichen Arbeitsmarkteintritt zu realisieren: privilegierte Berufs- und Einkommenschancen sowie ökonomischer Nutzen (Einkommen), soziale Anerkennung und Vermeidung von Verlusten (Statuserhalt) (Lindenberg & Frey, 1993). Im Rahmen des Vortrags wird der Stellenwert der Berufsmaturität in der beruflichen Bildung und im Handlungsset jugendlicher Schulabgänger und Schulabgängerinnen behandelt und deren Aspiration von Abschlussoptionen im Kontext individueller Berufsbildungsentscheidungen untersucht. Der Beitrag geht der Frage nach, wie sich die Wahl einer Berufsausbildung mit begleitender oder anschließender Berufsmaturität erklären lässt und welche der Jugendlichen bereits früh den Weg über die Berufsausbildung in die tertiäre Bildung aspirieren.

Die Daten der DAB-Panelstudie Panelstudie „Berufsbildungsentscheidungen: Determinanten der Ausbildungswahl und der Berufsbildungschancen“ ermöglichen eine Beschreibung des Entscheidungsprozesses am Übergang in die nachobligatorische Ausbildung und mechanismenbasierte (Mehrebenen-)Analysen zur Ausbildungswahl wie auch zur Veränderung von Bildungsaspirationen. Das Sample entspricht einer 10%-Zufallsstichprobe nach Gemeindetyp (9-stufig), stratifiziert nach Schultyp und Migrationspopulation der Schulen. Umfangreiche Daten zum Entscheidungsprozess wurden mit Online-Befragungen Mitte der 8. Klasse sowie zu Beginn und gegen Ende der 9. Klasse von rund 3'200 Jugendlichen aus ca. 200 Klassen erhoben. Die Eltern wurden zum ersten Messzeitpunkt befragt. Es wird zunächst die Berufsmaturität aus Sicht der Jugendlichen dargestellt und deren Aspiration und Entscheidung für/gegen eine Berufsmaturität untersucht. Es werden diejenigen Merkmale identifiziert, die Entscheidungen im Handlungsspektrum *„nur“ Berufsausbildung – Berufsmaturität – gymnasiale Maturität* erklären. Dabei werden Geschlecht, soziale und ethnische Herkunft sowie schulische Leistung individuelle Präferenzstrukturen, Persönlichkeitsmerkmale und die Kosten-Nutzen-Einschätzungen berücksichtigt sowie deren Zusammenspiel analysiert. Erste Ergebnisse verweisen auf Unterschiede nach Geschlecht und Schultyp hinsichtlich der Ausbildungsentscheidung. Im Entscheidungsprozess verändern sich die Ausbildungswünsche insbesondere zwischen den ersten beiden Befragungszeitpunkten. Herkunftsbedingte Disparitäten bei der Ausbildungsentscheidung können zu einem wesentlichen Anteil mit subjektiven Kosten-Nutzen-Einschätzungen der Jugendlichen erklärt werden.

Boudon, R. (1974). Education, opportunity, and social inequality: changing prospects in western society. New York: John Wiley.

Breen, R. & Goldthorpe, J. H. (1997). Explaining Educational Differentials: Towards a Formal Rational Action Theory. *Rationality and Society*, 9, 275-305.

Esser, H. (1999): Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln. Frankfurt a. M.: Campus.

Gonon, P. (1994). Die Einführung der Berufsmatura in der Schweiz als Prüfstein einer Neuorientierung von Allgemeinbildung und Berufsbildung. *Zeitschrift für Pädagogik*, 40(3), 389-404.

Lindenberg, S. (1989). Social Production Functions, Deficits, and Social Revolutions Prerevolutionary France and Russia. *Rationality and Society*, 1(1), 51-77.

Lindenberg, S. & Frey, B. S. (1993). Alternatives, frames, and relative prices: a broader view of rational choice theory. *Acta sociologica*, 36(3), 191-205.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



Corinna Geppert (Universität Wien), Sonja Bauer-Hofmann (Universität Wien), Tamara Katschnig (Universität Wien), Michaela Kilian (Universität Wien), Mariella Knapp (Universität Wien)

## **Schul(reform) im Spannungsfeld inner- und außerschulischer Ressourcen - Ergebnisse der NOESIS-Studie zur aktuellen Hausaufgabensituation in den Neuen Niederösterreichischen Mittelschulen (Längsschnitt)**

Erfolg oder Misserfolg im Bildungswesen ist geprägt durch das dynamische Zusammenspiel inner- und außerschulischer Ressourcen, wenngleich der Forschungslage nach die außerschulischen Ressourcen oder Voraussetzungen weitgehend wirkungsmächtiger sind als innerschulische Prozesse (siehe dazu zusammenfassend Baumert, Stanat & Watermann, 2006; Becker & Lauterbach, 2006). Es kommt zu einer Verschränkung zwischen den in der Schule zur Verfügung gestellten innerschulischen und den im elterlichen Haushalt erhaltenen außerschulischen Ressourcen, wie auch zu einer Verschränkung zwischen der Institution Schule und der Institution Nachhilfe (siehe dazu auch Baker, 2006). In Österreich sowie international lässt sich ein Anstieg der außerschulischen Lernbegleitung erkennen. Österreichweit unterstützen etwa  $\frac{1}{4}$  der Eltern ihre Kinder beim Erledigen der Hausaufgaben und beim Lernen. 50% der Eltern geben an, sich in der 5. Schulstufe so gut wie täglich mit der Hausaufgabenbetreuung und dem Lernen ihrer Kinder zu beschäftigen (IFES, 2013). Die zeitlichen und persönlichen Ressourcen, die Eltern damit aufwenden sind ziemlich groß.

Die als Reaktion auf die ersten PISA Ergebnisse im Schuljahr 2008 neben der AHS eingeführte Neue Mittelschule (neue Schulform der Sekundarstufe I – Schule der 10-14-Jährigen) setzt sich zum Ziel, durch zugeschnittenen Förderunterricht die Notwendigkeit von Nachhilfe und sonstige außerschulische Betreuung zu minimieren, weshalb interessiert, wie sich das „parental involvement“ in der Neuen Mittelschule darstellt: Wie beschreiben SchülerInnen die Veränderung der Hausübungssituation über den Verlauf der NMS (5.-8. Schulstufe) mit besonderem Bezug auf die ihnen zur Verfügung stehenden Ressourcen?

Datenbasis bilden Messzeitpunkt 2 bis 4 der Kohorte 1 (Jahrgang 2010/11) der NOESIS-Transitions-Erhebungen, somit SchülerInnen der fünften bis zur siebenten Schulstufe in 16 Neuen Niederösterreichischen Mittelschulen. Ergänzt werden diese SchülerInnen- und Lehrenden-Angaben in Kohorte 1 (N = 500 SchülerInnen) durch Querschnittsvergleiche der beiden Folgekohorten (N = 680 in Kohorte 2; N = 700 in Kohorte 3).

Ergebnisse: Die Ergebnisse verdeutlichen, dass gleich zu Beginn der Neuen Niederösterreichischen Mittelschule die elterliche Beteiligung an den Hausübungen noch nicht so ausgeprägt notwendig zu sein scheint. So geben zwar beinahe 70 % der SchülerInnen an, sie würden ihre Eltern bitten, ihnen bei der Hausübung zu helfen, dieser Prozentsatz steigert sich aber im Laufe der Zeit - gerade zu Beginn der sechsten Schulstufe zu knapp 90 %, um dann Anfang des dritten Jahres wieder leicht abzusinken. Eine Panelanalyse mit Fokus auf inner- und außerschulische Ressourcen legt den Schluss nahe, dass vor allem Netzwerke bei der Erledigung von Hausaufgaben wichtig sind. Auf Basis der Ergebnisse kann man schlussfolgern, dass bei SchülerInnen, die viele außerschulische Peerkontakte haben und die eine verbesserte Beziehung zur ihren Lehrenden über die Zeit wahrnehmen, auch die elterliche Involviertheit zunimmt. Auch die persönliche Lernmotivation der SchülerInnen wirkt sich positiv auf die Hausaufgabenunterstützung aus. Je lernmotivierter die SchülerInnen über die Zeit werden, umso mehr nehmen sie auch elterliche Unterstützung in Anspruch.

Conclusio: Aus den Analysen wird ein aus der Forschung bekanntes Phänomen deutlich. Je gleichmäßiger und einheitlicher das Schulangebot ist, beziehungsweise je mehr von schulischer Seite unternommen wird soziale Unterschiede in der Schule auszublenden, umso wahrscheinlicher erklärt das Bildungskapital d.h. außerschulische, soziokulturelle und sozial-ökonomische Variablen die verbleibenden Leistungsunterschiede (siehe dazu beispielsweise Hopmann, 2012). Entgegen der Annahme, je homogener ein Schulsystem ist desto weniger Nachhilfe, zeigt die aktuelle Forschung jedoch das Gegenteil. Eltern bilden somit eine relevante Ressource im Verlauf der Mittelschule, die es SchülerInnen erlaubt, ihre Schulkarriere zu gestalten und in der Schule zu Recht zu kommen. Eigene Lernmotivation sowie die Ressource „Freunde“ sind ebenfalls relevante Einheiten für die Ausprägung der elterlichen Involviertheit in schulische Belange. Die Notwendigkeit dieser außerschulischen Ressource wird über die Zeit nicht geringer, sondern nimmt – auch aufgrund der immer größer werdenden gesellschaftlichen und Arbeitsmarktanforderungen stetig zu. Dies kann auch durch die Einführung einer Schulform, die „eine Schule für alle“ sein möchte, kaum abgeschwächt werden.

Baker, D. (2006). Institutional Change in Education: Evidence from Cross-National Comparisons. In: H.D. Meyer & B. Rowan (Hrsg.), The New Institutionalism in Education (S. 163-186). Albany: SUNY Press.

Baumert, J., Stanat, P., & Watermann, R. (2006). Herkunftsbedingte Disparitäten im Bildungswesen. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, R., & Lauterbach, W. (2007). Vom Nutzen vorschulischer Erziehung und Elementarbildung – Bessere Bildungschancen für Arbeiterkinder? In R. Becker, & W. Lauterbach, Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit (2. Ausg., S. 129-159). Wiesbaden: VS Verlag für Sozialwissenschaften.

Hopmann, S., Bauer, S. & Werkl, T. (2012). Zwischen Scylla und Charybdis? Die österreichische Matura als Maßstab öffentlicher Bildung. In: Schneider-Taylor, B., Bosse, D. & Eberle, F. (Hrsg.): Matura und Abitur in den Zeiten von Bologna (S.69-85). Beltz Juventa Verlag, Weinheim Basel.

IFES. (2013). AK-Studie: Nachhilfe in Österreich. Bundesweite Elternbefragung 2013. Studienbericht. Wien: AK-Wien, Abteilung Bildungspolitik.

GEBF TAGUNG  
2014

B P1

Postergruppe

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



Schuleffektivität und Schulentwicklung



## Reanalysen der Gesamtschulstudien von Fend zur Schul- und Unterrichtsqualität

„Schule im Wandel V“ (SchiWa) ist eine Vorstudie für eine größer angelegte mixed-method Studie zur systematischen und historischen Untersuchung der Veränderung des Schulsystems. Ausgangspunkt bilden die Gesamtschulstudien von Helmut Fend aus den 1970er Jahren und deren Auswirkungen auf die Debatte zur Entwicklung von Schulqualität in Deutschland (Fend 1986, 1998; Klieme, 2013), die um eine aktuelle quantitative Erhebung und eine zeithistorische Studie zur Einführung der Gesamtschulen ergänzt werden sollen. Die ebenfalls als mixed-method Design konzipierte Vorstudie setzt bei den Gesamtschulstudien der 1979er Jahre an (Fend, 1982). Sie besteht aus zwei Teilprojekten.

Dieses Poster zum Projekt „Schule im Wandel“ wird den quantitativen Teil dieser Vorstudie vorstellen und die Arbeitsschritte sowie erste Ergebnisse des quantitativen Teilprojekts näher beleuchten. Für den qualitativen Teil des Projekts wird ein separates Poster eingereicht.

Im Rahmen der quantitativen Teilstudie steht die Re-Analyse von Daten der Gesamtschulstudien aus den 1970er Jahren sowie die Validierung der damals eingesetzten Instrumente an einer aktuellen Stichprobe im Vordergrund, insbesondere die der beiden Erhebungen 1978 in Hessen (Haenisch, 1979b; Lukesch, 1979b) und Nordrhein-Westfalen (Krüger-Haenisch o. J.; Haenisch 1979a; Lukesch, 1979a) sowie der Erhebung 1979 in Niedersachsen (Fend, 1981), da diese auch Informationen zu den Leistungen der Schülerinnen und Schüler und zum Unterricht enthalten. In Hessen wurden Schüler der 6. (N=905) und 9. Klassen (N=1012) aus allen Gesamtschulen des Kreises Wetzlar sowie deren Lehrpersonen (N=320) befragt. In Nordrhein-Westfalen und Niedersachsen wurden neben den Gesamtschulen auch Schulen des traditionellen Schulsystems (Hauptschule, Realschule oder Gymnasium) in die Untersuchung mit einbezogen. Für Nordrhein-Westfalen wurden Daten von Schülerinnen und Schülern der 6. (N=1544) und 9. Klassen (N=1478) und von 250 Lehrpersonen erhoben. In Niedersachsen wurden keine Sechstklässler des traditionellen Schulsystems befragt, da diese größtenteils die Orientierungsstufe besuchten. Das führt in dieser Befragung zu einem relativen kleinen Anteil an Schülern der 6. Klasse (N=350) im Vergleich zur 9. Klasse (N=1082). Es wurden 202 Lehrpersonen befragt. Die Auswahl der Gesamtschulen und Schulen des traditionellen Schulsystems erfolgte in Nordrhein-Westfalen und Niedersachsen nach Schulkriterien und Standortbedingungen, die Auswahl der jeweiligen Klassen nach dem Zufallsprinzip. Die Erhebungsinstrumente waren in den drei Bundesländern nahezu identisch. Die Schülerinnen und Schüler wurden zu ihrer (Schul-)Biographie, ihren Einstellungen, affektiven Reaktionsmustern, der Unterstützung ihrer Eltern, dem Verhalten ihrer Lehrer und weiterer Unterrichtsvariablen befragt. Außerdem wurden ihre Intelligenz und ihre Schulleistungen in den Fächern Deutsch, Mathematik, Englisch und in der 9. Klasse auch in Naturwissenschaften erfasst. Die Lehrpersonen wurden nach biographischen Merkmalen, Einstellungen, methodisch-didaktischen Maßnahmen und Lerngelegenheitsbeurteilungen gefragt. Für die Re-Analyse wurden zunächst die Variablen der Schüler- und Lehrerdatsätze rekonstruiert und die dazugehörigen Test- und Fragebogeninstrumente elektronisch erfasst. Nach der Aufbereitung der Datensätze erfolgen derzeit erste Analysen auf Itemebene hinsichtlich der Verteilungseigenschaften und die Überprüfung der internen Konsistenz und Dimensionalität der Skalen. Außerdem werden die Messeigenschaften der Items und Skalen der Fragebögen nach heutigen methodischen Standards (CFA; IRT) einer Prüfung unterzogen. Neben der Darstellung der Vorstudie liegt der Schwerpunkt des Posters auf den ersten Ergebnissen zur Güte und Qualität der Daten.

Fend, H. (1981). Leistungsvergleich zwischen Gesamtschulen und Schulen des traditionellen Schulsystems. Wissenschaftliche Bestandsaufnahme der Arbeit der niedersächsischen Gesamtschulen. Eine Information des niedersächsischen Kultusministers.;

Fend, H. (1982). Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs. Weinheim: Beltz.;

Fend, H. (1986). „Gute Schulen-Schlechte Schulen“. Die einzelne Schule als pädagogische Handlungseinheit. Deutsche Schule, 78, 275–293.;

Fend, H. (1998). Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim [etc.]: Juventa.;

Haenisch, H. (1979a). Schulleistungsvergleich zwischen Gesamtschulen und Schulararten des gegliederten Schulsystems in Nordrhein-Westfalen am Ende des 6. Schuljahres. In H. Haenisch, H. Lukesch, R. Klaghofer & E.-M. Krüger-Haenisch (Hrsg.), Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen. Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik (S. 1–226). Paderborn: Schöningh.;

Haenisch, H. (1979b). Schulleistungsvergleiche zwischen Gesamtschulen in Hessen und Schulen des gegliederten Schulsystems am Ende des 6. Schuljahres. Arbeitsbericht 12. Konstanz.;

Klieme, E. (2013). Qualitätsbeurteilung von Schule und Unterricht: Möglichkeiten und Grenzen einer begriffsanalytischen Reflexion – ein Kommentar zu Helmut Heid. Zeitschrift für Erziehungswissenschaft, 16(2), 433–441.;

Krüger-Haenisch, E.-M. (o. J.). Erhebungsinstrumente der Leistungsstudie 1978 (außer Leistungstests). Arbeitsbericht. Konstanz.;

Lukesch, H. (1979a). Leistungsvergleich zwischen Gesamtschulen und herkömmlichen Schulen am Ende der Pflichtschulzeit in Nordrhein-Westfalen. In H. Haenisch, H. Lukesch, R. Klaghofer & E.-M. Krüger-Haenisch (Hrsg.), Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen. Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik (S. 227–364). Paderborn: Schöningh.;

Lukesch, H. (1979b). Schulleistungsvergleiche zwischen Gesamtschulen in Hessen und Schulen des gegliederten Schulsystems am Ende der Pflichtschulzeit. Arbeitsbericht 13. Konstanz.

Anja Felbrich (Humboldt-Universität zu Berlin), Dirk Richter (IQB, Humboldt-Universität zu Berlin), Poldi Kuhl (IQB, Humboldt-Universität zu Berlin), Petra Stanat (IQB, Humboldt-Universität zu Berlin), Hans Anand Pant (IQB, Humboldt-Universität zu Berlin)

## ***Die Jahrgangsmischung auf dem Prüfstand: Die Effekte jahrgangsübergreifenden Lernens auf die Kompetenzen und das sozio-emotionale Wohlbefinden von Schülerinnen und Schülern in der Grundschule***

Die Einführung jahrgangsgemischter Lerngruppen in der Schulanfangsphase wird kontrovers diskutiert. Einerseits werden positive Effekte der Jahrgangsmischung auf die Entwicklung der Kinder sowohl im fachlichen als auch im sozio-emotionalen Bereich erwartet, andererseits befürchtet, dass die größere Heterogenität der Schülerschaft die Lehrkräfte überfordere. Auf der Basis internationaler Metaanalysen konnten bislang keine oder nur sehr geringe Differenzen in den kognitiven und sozial-emotionalen Merkmalen zwischen Klassen gefunden werden, die jahrgangsübergreifend oder jahrgangshomogen unterrichtet wurden (z.B. Gutiérrez & Slavin, 1992; Veenman, 1995). Im deutschen Sprachraum liegen bislang nur wenig belastbare Befunde zur Jahrgangsmischung vor. Hier setzt die vorliegende Studie an und überprüft anhand der ersten Ländervergleichsstudie des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) in der Primarstufe (Stanat, Pant, Böhme & Richter, 2012) an einer großen, auf Länderebene repräsentativen Stichprobe, ob sich die Kompetenzen und sozio-emotionalen Merkmale von Schülerinnen und Schülern in jahrgangsgemischten Klassen und Regelklassen unterscheiden. Im Fokus der Untersuchung stehen dabei die Kompetenzen im Lesen und in Mathematik am Ende der Jahrgangsstufe 4. Darüber hinaus wird untersucht, ob sich mögliche Effekte der Jahrgangsmischung auch für sozio-emotionale Merkmale wie die Lernfreude der Kinder und die wahrgenommene soziale Integration in der Klasse nachweisen lassen.

Die Ergebnisse der Mehrebenenanalysen zeigen, dass keine oder nur geringe Unterschiede zwischen Schülerinnen und Schülern aus jahrgangsgemischten Klassen und Jahrgangsklassen bestehen. Während sich die Lesekompetenz und die Lernfreude unter Kontrolle zentraler Hintergrundvariablen für Kinder in jahrgangsübergreifenden und jahrgangshomogenen Klassen nicht unterschieden, wurden für Mathematik und die soziale Integration geringfügige Effekte zwischen beiden Organisationsformen ersichtlich, wenn man zusätzlich den Verbreitungsgrad jahrgangsgemischter Lerngruppen im Land berücksichtigte. So weisen die Ergebnisse darauf hin, dass jahrgangsgemischte Klassen in Ländern mit hohem Verbreitungsgrad schwächere Mathematikleistungen erreichen und sich die Schülerinnen und Schüler weniger integriert fühlen als Gleichaltrige in den Regelklassen. Die Befunde werden in die nationale und internationale Befundlage zu Effekten jahrgangsgemischten Lernens eingeordnet und auch im Hinblick auf ihre praktische Relevanz diskutiert.

Gutiérrez, R. & Slavin, R. E. (1992). Achievement effects of the nongraded elementary school: A best evidence synthesis. *Review of Educational Research*, 62, 333-376.

Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011. Münster: Waxmann.

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, 65, 319-381.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## **Können zentrale Lernstandserhebungen spezifische Lese- und Rechtschreibscreenings ersetzen?**

Eine erfolgreiche Bewältigung schulischer Anforderungen gelingt nur, wenn ein Kind die Schriftsprache ausreichend beherrscht. In höheren Klassenstufen setzen altersadäquate Lernerfolge voraus, dass ein Schüler über ausreichende Fähigkeiten im sinnentnehmenden Lesen verfügt. Seit PISA ist die Bedeutung des Schriftspracherwerbs für die Entwicklungschancen eines Kindes in das Bewusstsein einer breiten Öffentlichkeit gelangt. Seither werden erhebliche Anstrengungen unternommen, Kinder mit Risiken beim Schriftspracherwerb frühzeitig zu erkennen und zu fördern. Eine zutreffende Beurteilung des Leistungsstandes eines Schülers im schulischen Alltag ist allerdings schwierig und gelingt vielen Lehrkräften nur unzureichend. Um Lehrkräften und Schulen die Möglichkeit zu geben, den Leistungsstand einzelner Schüler und ganzer Klassen in den Kernfächern Deutsch und Mathematik besser einzuschätzen, wurden 2004 zentrale Lernstandserhebungen (VERgleichsArbeiten - VERA) eingeführt. In der vorliegenden Studie wird der Frage nachgegangen, ob Kinder mit basalen Schwächen beim Lesen und Schreiben durch Schulnoten und zentrale Vergleichsarbeiten (VERA) ausreichend zuverlässig erfasst werden oder ob dazu ergänzend ein spezifisches Lese- und Rechtschreibscreening erforderlich ist.

In die Studie wurden 761 Drittklässler (390 Mädchen und 371 Jungen) einbezogen. Die Kinder waren im Durchschnitt 8 Jahre und 7 Monate alt (SD = 0;5 Jahre;Monate). Die Untersuchungen erfolgten in der 3. bis 5. Unterrichtswoche der 3. Klasse. Die Lese- und Rechtschreibfähigkeiten wurden als Gruppentests mit dem Salzburger Lese-Screening für die Klassenstufen 1-4 (SLS 1-4) (Mayringer & Wimmer, 2003) und dem Rechtschreibteil des revidierten Salzburger Lese- und Rechtschreibtests (SLRT-II) (Moll & Landerl, 2010) erfasst. Die Beurteilung der nonverbalen Intelligenz erfolgte mit den Coloured Progressive Matrices (CPM) (Raven et al., 2002). Gegen Ende des Schuljahres wurden die Deutschnoten auf den Zeugnissen Ende der 2. und Mitte der 3. Klasse sowie die VERA 3-Ergebnisse erfragt. Bei der Lernstandserhebung VERA 3, die im Mai 2012 erfolgte, wurden für das Fach Deutsch die Kompetenzbereiche „Lesen“ und „Sprachgebrauch“ überprüft (Fischer & Wagner, 2013).

Die Ergebnisse weisen auf eine moderate Übereinstimmung zwischen basalen Lese- und Rechtschreibleistungen und den VERA-Ergebnissen bzw. den Schulnoten im Fach Deutsch hin. Am höchsten waren die Beziehungen zwischen spezifischen Noten für Lesen und Schreiben und den psychometrischen Testergebnissen. In der Stichprobe wurden aufgrund der Testergebnisse 99 Kinder (13%) als lese- bzw. rechtschreibschwach eingestuft. Bei diesen Kindern lagen die Ergebnisse entweder im Rechtschreibtest (SLRT-II-Skala: Wörter falsch) und/oder im Lesetest (SLS 1-4: Wörter richtig) mindestens 1½ Standardabweichungen unterhalb des Mittelwerts. Nur wenige von diesen Kindern mit einer Lese-Rechtschreibschwäche haben im Fach Deutsch die Note 4 oder 5 erhalten. Anhand der Schulnoten lassen sich somit Kinder mit einem Förderbedarf im Lesen und Rechtschreiben kaum identifizieren. Häufiger fallen lese-rechtschreibschwache Kinder bei der VERA-Überprüfung auf. Bei 35 % dieser Kinder wurden die Kompetenzen im Bereich „Lesen“ und bei 38 % im „Sprachgebrauch“ als „unter dem Mindeststandard“ liegend eingestuft und 47 % liegen in mindestens einem der Bereiche auf diesem untersten Kompetenzniveau. Dennoch weisen die vorliegenden Ergebnisse darauf hin, dass anhand von Schulnoten und VERA Ergebnissen mindestens die Hälfte der Kinder mit unzureichenden basalen schriftsprachlichen Fähigkeiten übersehen wird. Deshalb sollte, um diese Kinder rechtzeitig zu erkennen und zu fördern, spätestens Anfang der dritten Klasse routinemäßig ein Screening der Lese-Rechtschreibfähigkeiten eingeführt werden. Dies wäre mit den in dieser Studie eingesetzten oder ähnlichen Instrumenten ohne eine erhebliche Belastung der Lehrkräfte und ohne eine nennenswerte Beeinträchtigung des Unterrichts realisierbar.

Mayringer, H. & Wimmer, H. (2003). Salzburger Lese-Screening für die Klassenstufen 1-4 (SLS 1-4). Bern: Huber.

Moll, K. & Landerl, K. (2010). Lese- und Rechtschreibtest SLRT-II. Bern: Huber.

Raven, J. C., Raven, J. & Court, J. H.; deutsche Bearbeitung und Normierung von Bulheller, S. & Häcker, H. (2002). Coloured Progressive Matrices - CPM. Frankfurt: Pearson Assessment.

## **VERA-Kompetenzwerte sagen Leistungskriterien am Ende der Grundschulzeit in Brandenburg vorher**

Die Bildungsstandards der KMK definieren Kompetenzstufen, die von Schülerinnen und Schülern in bestimmten Jahrgangsstufen erreicht werden sollten. Mit den Kompetenzstufen werden fachliche und fachübergreifende Basisqualifikationen beschrieben, die für die weitere schulische und berufliche Ausbildung von Bedeutung sind (z.B.: KMK, 2013a; KMK, 2013b). Die Bildungsstandards dienen dabei u.a. als Grundlage für die Itementwicklung der Vergleichsarbeiten (VERA) für Schülerinnen und Schüler unterschiedlicher Jahrgangsstufen. Die Antworten der Schülerinnen und Schüler bei den Items zur Erfassung der Bildungsstandards sollen dann Hinweise liefern, inwieweit die Schülerinnen und Schüler in den Bildungsstandards definierte Leistungsniveaus erreicht haben. VERA stellt daher als ein Instrument datengestützter Unterrichts- und Schulentwicklung eine zentrale Komponente der KMK-Gesamtstrategie zum Bildungsmonitoring dar (KMK, 2012).

Für Kompetenztests, wie beispielsweise die PISA-Kompetenztests, konnte gezeigt werden, dass diese Vorhersagekraft für Klassenübergänge, Schulnoten und Abschlussprüfungen besitzen (u.a. Fischbach, Keller, Preckel, & Brunner, 2012). Bisher ist jedoch ungeklärt, inwieweit die VERA-Kompetenzwerte tatsächlich für die weitere schulische Ausbildung bedeutsam sind. So gibt es bspw. noch keine empirischen Untersuchungen, die untersucht haben, ob VERA-Kompetenzwerte die schulische Leistungsentwicklung (z.B. Kompetenzstände, Schulnoten oder Übergangsempfehlungen) vorhersagen können. Weiterhin ist unklar ob VERA-Kompetenzwerte einen nachweisbaren prognostischen Mehrwert über Schulnoten hinaus liefern können. Diese Fragen wollen wir mit der vorliegenden Studie beantworten.

Unsere Untersuchung basiert auf einer Längsschnitterhebung an Brandenburger Grundschulen (KEGS; N=1.745 Schülerinnen und Schüler aus 75 öffentlichen Grundschulen). Hierbei wurden u. a. jährliche Leistungstests der mathematischen Kompetenz von der zweiten Jahrgangsstufe und des Leseverständnisses im Fach Deutsch von der vierten Jahrgangsstufe bis zur sechsten Jahrgangsstufe durchgeführt. Die Leistungsdaten wurden durch eine Raschskalierung auf die Bildungsstandardmetrik überführt. Ausgangspunkt zur Betrachtung der Fragestellung stellen die Kompetenzmessungen der mathematischen Kompetenz in der dritten Jahrgangsstufe (Schuljahr 2007/2008) und des Leseverständnisses in Deutsch in der vierten Jahrgangsstufe (Schuljahr 2008/2009), sowie die Halbjahresnoten zur vierten Jahrgangsstufe für Mathematik und Deutsch dar. Als Vorhersagekriterien dienen die Kompetenzstände, Schulnoten und Übergangsempfehlungen der sechsten Jahrgangsstufe (Schuljahr 2010/2011).

Multivariate Regressionsmodelle zeigen Evidenz für die prognostische Validität der eingesetzten Kompetenztests auf Leistungen in Kompetenztests in (a) Mathematik und (b) Deutsch, sowie für (c) Schulnoten und (d) die Übergangsempfehlung in der sechsten Jahrgangsstufe. (e) Diese prognostische Validität konnte auch dann nachgewiesen werden, wenn in den Regressionsmodellen für die Halbjahresnoten in Mathematik und Deutsch kontrolliert wurde. Zusammenfassend kann festgehalten werden, dass (an Brandenburger Grundschulen) die VERA-Kompetenzwerte zentrale Leistungskriterien am Ende der Grundschulzeit vorhersagen können. Die Ergebnisse liefern Hinweise dafür, dass für Lehrkräfte im Rahmen von VERA, hier speziell zur dritten Jahrgangsstufe, ein geeignetes Instrument vorliegt, Schülerleistungen als Tendenzen für Risiken und Erfolge am Ende der Primarstufe auszuwerten und als zusätzliche Informationsquelle neben den Schulnoten einzubeziehen. Solche Evidenzen prognostischer Validität könnten die Akzeptanz von VERA bei Lehrkräften steigern und deren Auseinandersetzung mit den Rückmeldungen der VERA-Testungen intensivieren und somit die Implementation fördern.

Fischbach, A., Keller, U., Preckel, F., & Brunner, M. (2012). PISA proficiency scores predict educational outcomes. *Learning and Individual Differences*, 24 (2013), 63-72.

KMK Kultusministerkonferenz (2012). Vereinbarung zur Weiterentwicklung von VERA. Beschluss der Kultusministerkonferenz vom 08.03.2012.

KMK Kultusministerkonferenz (2013a). Kompetenzstufenmodell zu den Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4). Stand 11. Februar 2013.

KMK Kultusministerkonferenz (2013b). Kompetenzstufenmodell zu den Bildungsstandards für das Fach Deutsch im Kompetenzbereich „Lesen – mit Texten und Medien umgehen“ -Primarbereich -. Stand 13. Februar 2013.

## **Relative subject matter performance – Ein attraktiver Ansatz für die Schuleffektivitätsforschung?**

*Value Added* Ansätze sind in der Schuleffektivitäts- und Schulentwicklungsforschung eine weit verbreitete Methode zur Einschätzung und Überprüfung der Leistungsfähigkeit und Effektivität von Schulen (Tekwe et al., 2004). Vor allem in den USA werden diese Verfahren dazu genutzt den Erfolg einer Schule zu messen und über eventuelle Sanktionen oder Förderungen zu entscheiden. Eine Möglichkeit zur Berechnung des Value Added ist der Vergleich der schulspezifischen Ergebnisse eines standardisierten Tests einer Kohorte mit den Ergebnissen desselben Tests einer vorherigen Kohorte. Es werden also Querschnittsdaten in einem Trenddesign über mehrere Jahre und Kohorten hinweg miteinander verknüpft, um somit die Entwicklung einer Schule festzustellen. Wendet man diese Methode auf die bundesweit durchgeführten Vergleichsarbeiten am Ende der 3. Jahrgangsstufe (VERA 3) an, ergibt sich der Value Added beispielsweise wie folgt: Differenz des Mittelwertes der Schule bei VERA 3 im Schuljahr 2012/13 vom Mittelwert der Schule bei VERA 3 im Schuljahr 2011/12. Ist der Wert positiv, hat sich die Schule verbessert. Unklar ist nun aber, ob diese Verbesserung auf einen „wahren“ Schuleffekt oder auf eine unterschiedliche Zusammensetzung der Schülerkohorten hinsichtlich leistungsrelevanter Kontextmerkmale (z.B. sozioökonomischer oder herkunftsbezogener Hintergrund) zurückzuführen ist. Um also festzustellen, was eine Schule über solche nicht von der Schule zu beeinflussenden Faktoren hinaus leistet, ist es erforderlich genau diese Faktoren im Model zur Berechnung des Value Added zu berücksichtigen.

Kelly und Monczunski (2004) versuchen diesem Problem mit einem erweiterten Value Added Ansatz Rechnung zu tragen. Mit ihrem Konzept der Relative Subject Matter Performance setzen sie die schulspezifischen Ergebnisse in einem Fach in Relation zu den fachspezifischen Gesamtergebnissen (beispielsweise im Falle von VERA 3 für das Fach Deutsch, Kompetenzbereich Lesen im Land Berlin). Von dem daraus resultierenden Wert werden wiederum die mittleren Ergebnisse in allen anderen Fächern (im Falle von VERA 3 das Fach Mathematik) subtrahiert. Bei dieser Methode steht die Schule mit ihren spezifischen Bedingungen (e.g. sozialen Umfeld, Einzugsgebiet, finanzielle Ressourcen) nicht in direktem Vergleich zu anderen Schulen mit ihren wiederum spezifischen Bedingungen. Vielmehr kann hier z.B. herausgefunden werden, welche Schule es trotz schwieriger Bedingungen schafft, ihre Schülerinnen und Schüler mit den abgefragten Kompetenzen auszustatten und welche Schule möglicherweise unter besseren Bedingungen gleich gute oder sogar schlechtere Leistungen in einem bestimmten Fach erzielt.

Wir glauben, dass dieser Ansatz der Realität einer Schule näher kommt, da auch sie, die erzielten Ergebnisse stets in den Kontext des sie umgebenden Umfeldes einordnet. Davon ausgehend, möchten wir in unserem Beitrag die, in Deutschland bisher kaum verwendete, Methode der Relative Subject Matter Performance am Beispiel von 345 Berliner Grundschulen, die in den Jahren 2010 bis 2013 an VERA 3 teilgenommen haben, den bisher verwendeten Verfahren von Value Added gegenüberstellen und die jeweiligen Vor- und Nachteile diskutieren. Damit versuchen wir einen Beitrag zur Diskussion der Methoden der Schuleffektivitätsforschung zu leisten.

Kelly, S., & Monczunski, L. (2007). Overcoming the volatility in school-level gain scores: A new approach to identifying value added with cross-sectional data. *Educational Researcher*, 36(5), 279-287.  
Tekwe, C. D., Carter, R. L., Ma, C. X., Algina, J., Lucas, M. E., Roth, J., ... & Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11-36.



## **Online Diagnose, Ergebnisrückmeldung und dann? – Ein differenzierter Blick auf den Rezeptionsprozess von Lehrkräften**

Mit dem internetbasierten Diagnose- und Förderinstrument ‚Online Diagnose‘ steht den Lehrkräften an Haupt- und Werkrealschulen in Baden-Württemberg seit mehreren Jahren ein Instrument zur Verfügung, das ihnen hilft, mögliche Teildefizite ihrer Schülerinnen und Schüler im Fach Deutsch festzustellen, und das ihnen entsprechende Fördermaterialien bereitstellt. Ziel des vorliegenden Beitrags ist, anhand dieses Beispiels den Rezeptionsprozess von Lehrkräften im Umgang mit derartigen Rückmeldungen zu untersuchen. In deutschsprachigen Publikationen hat sich das Zyklenmodell der Rückmeldung (Hosenfeld & Groß Ophoff, 2007) als theoretische Grundlage etabliert (Müller 2009; vgl. auch Koch, 2011; Hosenfeld, 2010; Schneewind, 2006). Das Zyklenmodell beschreibt den Umgang mit Evaluationsergebnissen und Rückmeldungen in vier aufeinanderfolgenden Stufen: Rezeption, Reflexion, Aktion und schließlich Evaluation. Einfluss auf diesen Prozess nehmen individuelle, schulische und externe Bedingungen. Im Rahmen dieser Studie wurden zudem klassenspezifische Bedingungen als potentielle Determinanten ergänzt.

Der vorliegende Beitrag greift den Prozessschritt der Rezeption von Ergebnisrückmeldungen heraus. Eine misslungene oder erschwerte Rezeption von Ergebnisrückmeldungen gefährdet bereits in dieser frühen Prozessphase den Erfolg einer gewinnbringenden Nutzung von Ergebnisrückmeldungen (Koch, 2011). Die unterschiedlichen Bedingungsfaktoren erschweren oder erleichtern die Rezeption von Ergebnissen. Welche Aspekte der individuellen, schulischen und externen Bedingungen mit der Rezeption von Ergebnisrückmeldungen zusammenhängen, wird nachfolgend erläutert. In der Untersuchung wurden N=150 Deutschlehrkräfte zum internetbasierten Diagnose- und Förderinstrument ‚Online Diagnose‘ befragt. Die Rezeption der Ergebnisrückmeldungen wurde mittels drei Indikatoren erfasst: (1) die selbsteingeschätzte Erfolgserwartung im Umgang mit den Ergebnisrückmeldungen (z. B. „Es fällt mir leicht, die Ergebnisse des Haupttests zu interpretieren“), (2) die Klarheit und Verständlichkeit der Darstellungen (z. B. „Die Grafiken sind selbsterklärend“), sowie (3) der wahrgenommene Zeitaufwand (z. B. „Es ist mir zeitlich nicht möglich, die Ergebnisrückmeldungen von allen Schülerinnen und Schülern ausführlich zu lesen“). Die Determinanten individuelle, schulische und externe Bedingungen wurden wie folgt erfasst: (1) Individuellen Bedingungen: (a) Vorwissen (Anzahl der bisher durchgeführten Diagnose- und Förderprozesse im Rahmen der ‚Online Diagnose‘), (b) Computerbezogene Einstellung und Ängstlichkeit (z. B. „Das Arbeiten mit dem Computer bereitet mir Unbehagen“), (c) emotionale Belastung im schulischen Alltag („Ich fühle mich in der Schule oft erschöpft“); (2) Schulische Bedingungen: (a) Unterstützung durch die Schulleitung, das Kollegium und die Eltern (z. B. „Die Schulleitung ist zu Fragen des Unterrichts ansprechbar und kompetent“), (b) Anerkennung der ‚Online Diagnose‘ durch die Schulleitung, das Kollegium und die Eltern (z. B. „Die Online Diagnose ist im Kollegium ein akzeptiertes Hilfsmittel, um den Unterricht zu gestalten“); (3) Externen Bedingungen: (a) Anzahl der besuchten Fortbildungen zur ‚Online Diagnose‘, (b) externe Hilfen (z. B. Kurzanleitung: ‚Erste Schritte mit der Online Diagnose‘). Erste Analysen der bivariaten Zusammenhangsstruktur deuten darauf hin, dass eine positive computerbezogene Einstellung mit der Erfolgserwartung zusammenhängt ( $r = .21, p < .05$ ). Gleichzeitig korreliert die Erfolgserwartung ( $r = -.24, p < .01$ ) negativ mit der Ängstlichkeit von Lehrkräften im Umgang mit Computern. Ebenfalls negativ hängt die Computerängstlichkeit mit der Klarheit und Verständlichkeit der Darstellung ( $r = -.23, p < .01$ ) zusammen. Zwischen der emotionalen Belastung und der Erfolgserwartung zeigt sich eine moderate Korrelation ( $r = .20, p < .05$ ). Hinsichtlich der schulischen Bedingungen zeigen sich signifikante Zusammenhänge zwischen der Unterstützung durch die Schulleitung und der Erfolgserwartung ( $r = .27, p < .01$ ) sowie der Klarheit und Verständlichkeit der Darstellung ( $r = .22, p < .01$ ). Die Anerkennung der ‚Online Diagnose‘ im Kollegium ( $r = .32, p < .001$ ) und durch die Schulleitung ( $r = .21, p < .05$ ) korrelieren ebenfalls positiv mit dem Indikator Erfolgserwartung. Für die externen Bedingungen ergaben sich keine signifikanten Zusammenhänge. Mit multiplen Regressionsanalysen sowie Analysen des Zusammenhangs der Rezeption von Ergebnisrückmeldungen mit Aspekten der Motivation seitens der Lehrkräfte sollen diese Befunde weiter differenziert werden.

Hosenfeld, A. (2010). Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung?: Die Wirkung von videographischer und schriftlicher Rückmeldung bei Lehrkräften der vierten Jahrgangsstufe. Münster: Waxmann.

Hosenfeld, I., & Groß Ophoff, J. (2007). Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht. Empirische Pädagogik, 21 (4), 352-367.

Koch, U. (2011). Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten?: Datenkompetenz von Lehrkräften und die Nutzung von Ergebnisrückmeldungen aus Vergleichsarbeiten. Münster: Waxmann.

Müller, A. (2009). Rückmeldungen nach Vergleichsarbeiten im Kontext des Schulischen Qualitätsmanagements: Drei explorative Studien zu Gestaltung und Rezeption im Anschluss an KOALA-S. Berlin: Mensch und Buch Verlag.

Schneewind, J. (2006). Rückmeldungen als Motivator für die Teilnahme an Schulleistungsstudien?: Die Rezeptionsstudie von BeLesen. In H. Kuper, J. Schneewind, Rückmeldung und Rezeption von Forschungsergebnissen: Zur Verwendung wissenschaftlichen Wissens im Bildungssystem (S. 107-126). Münster: Waxmann.



Sebastian Vogel (Universität Kassel), Kay Achmetli (Universität Kassel), Janina Krawitz (Universität Kassel), Werner Blum (Universität Kassel)

## **Zentrale Lernstandserhebungen – zusätzliche Belastung oder Chance zur Förderung und Unterrichtsentwicklung?**

Seit 2009 finden alljährlich im Frühjahr deutschlandweit Lernstandserhebungen statt (kurz: LSE; auch „Vergleichsarbeiten“ genannt), u. a. für Mathematik in Klasse 8. Diese Tests sind im Bereich der Qualitätsentwicklung und Standardsicherung (Gesamtstrategie zum Bildungsmonitoring der KMK, 2009) anzusiedeln und stellen dort einen Teil der Evaluation der Kompetenzentwicklung der Schülerinnen und Schüler dar. Mit diesem Instrument sollen den Lehrkräften Orientierungen und Möglichkeiten zur Diagnose dieser Kompetenzentwicklung gegeben werden, um damit eine gezielte individuelle Förderung und eine Qualitätsentwicklung im Unterricht anzustoßen.

Die erfolgreiche Nutzung der LSE ist jedoch an bestimmte Gelingensbedingungen geknüpft. Nach Maier (2008) ist eine hohe Akzeptanz der Lehrkräfte notwendig, allerdings wird nur eine indifferente Haltung berichtet (vgl. auch Koch, 2011). Eine höhere Transparenz der Intentionen von LSE sowie zusätzliche Begleitinformationen können sich positiv auf die Akzeptanz der Lehrkräfte auswirken. Zentral für individuelle Förderung ist bei den LSE die diagnostische Kompetenz der Lehrkräfte (Kühle & Peek, 2007). Weitere Einflussfaktoren für die tatsächliche Nutzung der LSE zur individuellen Förderung und gezielten Unterrichtsentwicklung sind u.a. die wahrgenommene Lehrplanvalidität, Interpretationshilfen, die eine gute Balance zwischen Aufwand und Ertrag ermöglichen, sowie eine gute Zusammenarbeit und Unterstützung im Kollegium, um aus den Ergebnisse auch auf Jahrgangs- bzw. Fachkonferenzebene Schlüsse ziehen zu können.

Ziel des im Auftrag des Hessischen Kultusministeriums im Schuljahr 2012/13 in Hessen durchgeführten Projekts VELM-8 („Verbesserung der Effektivität der Lernstandserhebungen Mathematik Klasse 8“) ist es, exemplarisch zu zeigen, wie die LSE tatsächlich für die Unterrichtspraxis genutzt werden können. Im Laufe eines Schuljahres sollten die Einstellungen der Lehrkräfte zu den LSE und die konkrete Nutzung der LSE für die eigene Unterrichtsentwicklung sowie die individuelle Förderung durch gezielte Lehrerfortbildungen verbessert werden. In einem Prä-/Posttest-Design wurden die Einstellungen der Lehrkräfte bezüglich der LSE sowie deren Nutzung durch die Lehrkräfte erfasst. Genutzt wurden hierzu drei Skalen zu Einstellungen (Akzeptanz, Lehrplanvalidität und Belastung) und drei Skalen zur Nutzung (selektions- und förderdiagnostische Nutzung sowie Nutzung zur Unterrichtsgestaltung). Zusätzlich wurden die Mathematikleistungen der Schülerinnen und Schüler am Anfang und am Ende gemessen. Die Erhebungen fanden parallel in einer Projekt- und einer Wartekontrollgruppe statt. Zudem fanden bei einzelnen Lehrkräften Unterrichtsbeobachtungen nach der Lernstandserhebung statt. Im Laufe des Schuljahres fanden insgesamt vier Fortbildungen statt. Die Kontrollgruppe erhielt am Ende des Schuljahres eine Kompaktfortbildung.

Unsere Stichprobe umfasst 29 Lehrkräfte aus Nordhessen (11 aus der Wartekontroll- und 18 aus der Projektgruppe, knapp 60 % weiblich) mit einer durchschnittlichen Unterrichtserfahrung von etwa 14 Jahren. Gut die Hälfte der teilnehmenden Lehrkräfte hatte noch keine Vorerfahrungen mit den LSE. Etwa die Hälfte der Lehrkräfte kommt aus dem Gymnasium. Der Anteil der Haupt- und Realschullehrkräfte ist in etwa gleich stark vertreten.

Erste Auswertungen zeigen, dass die Fortbildungen (Projektgruppe) bzw. die Kompaktfortbildung (Wartekontrollgruppe) einen positiven Einfluss auf die Einstellungen der Lehrkräfte und die Nutzung der LSE haben. Ein ähnliches Bild zeichnen die individuellen Rückmeldungen durch die Lehrkräfte, die wir im Rahmen der Evaluation der einzelnen Fortbildungen schriftlich erhalten haben. Im Vortrag sollen diese ersten Ergebnisse ausgeschärft, der Zusammenhang mit anderen Faktoren (wie etwa Klassengröße, Unterrichtserfahrung etc.) ausführlicher analysiert und die Auswirkungen der Fortbildungen auf die Schülerleistungen untersucht werden. Ebenfalls gilt es die Unterrichtsbeobachtungen fallanalytisch zu betrachten. Ziel ist es, die Materialien, Ideen und Ergebnisse dieses Pilotprojekts im Schuljahr 2013/14 in Fortbildungen für Multiplikatoren umzusetzen

Koch, U. (2011). Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten?: Datenkompetenz von Lehrkräften und die Nutzung von Ergebnissrückmeldungen aus Vergleichsarbeiten. Münster [u.a.]: Waxmann.

Kühle, B. & Peek, R. (2007) Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnissrückmeldungen in Schulen. Empirische Pädagogik, 21(4), 428-447.

Kultusministerkonferenz (KMK) (2006). Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Verfügbar unter [http://www.kmk.org/fileadmin/veroeffentlichtungen\\_beschluesse/2006/2006\\_08\\_01-Gesamtstrategie-Bildungsmonitoring.pdf](http://www.kmk.org/fileadmin/veroeffentlichtungen_beschluesse/2006/2006_08_01-Gesamtstrategie-Bildungsmonitoring.pdf) [08.03.2013]

Maier, U. (2008). Rezeption und Nutzung von Vergleichsarbeiten aus der Perspektive von Lehrkräften. Zeitschrift für Pädagogik, 54(1), 95–117.

## **Sozialpsychologische Perspektiven auf den Unterstützungsbedarf von Schulleitungen zur selbstständigen Schulentwicklung**

In diesem Beitrag werden ausgewählte Ergebnisse einer qualitativen Querschnittuntersuchung mit N = 30 Schulleiter/-innen zu Themen der Steuerung und Entwicklung von Schulen dargestellt. Die Schulleitungstätigkeit stellt im 21. Jahrhundert eine Führungsaufgabe von wachsender Komplexität dar. Insbesondere Schulleitungen sind dazu angehalten, den Weg für eine immer eigenständiger agierende Schule zu ebnen und eine Neuausrichtung schulischen Denkens und Handelns auch im Lehrer/-innenkollegium zu verbreiten. Die neue Eigenständigkeit der Schulen wird dabei durch verschiedene Maßnahmen aus dem Bereich Educational Governance überprüft und unterstützt. Ein relativ neu eingeführtes Instrument der Analyse sowie der Rechenschaftslegung stellt dabei die externe Schulevaluation dar. Bezüglich externer Schulinspektionen existiert eine große Anzahl empirischer Studien in Ländern, die bereits seit längerer Zeit bemüht sind, ihr Schulsystem über evidenzbasierte Verfahren zu entwickeln (vgl. u.a. Plowright 2007, Janssens & Van Amelsvoort, 2008; Luginbuhl, Webbink, & De Wolf, 2009). In Deutschland bestehen derzeit noch Forschungsdesiderate im Hinblick auf Schulinspektionen. Diese sind mittlerweile zwar in allen Bundesländern flächendeckend eingeführt, jedoch ist dies erst seit wenigen Jahren der Fall (Müller, Dederig, & Bos, 2008).

Die Qualitätsanalyse kann dabei aus psychologischer Perspektive als Rückmeldung über die Leistung des Kollegiums einer Schule sowie deren Leitung verstanden werden. Insbesondere aus der psychologischen Forschung zum Selbstwert ist bekannt, dass Individuen bei Feedback dazu neigen, selbstwertdienliche Strategien anzuwenden (Semmer & Jacobshagen, 2010). Auch kann die Anordnung einer externen Evaluation sowie die aktuell starke Veränderung der Schullandschaft in NRW zu Unsicherheiten und zu Widerstandsprozessen führen (Steins, 2005, 2009). Dies gilt sowohl für die dadurch entstehende Erweiterung von Tätigkeiten der Schulleitungen als auch der Lehrkräfte. Auf Grund der starken Veränderung der Schullandschaft sowie der daran geknüpften vielfältigen Ansprüche an die neue Führungsverantwortung von Schulleitungen, ist es das Ziel der vorliegenden Untersuchung, herauszufinden, wie der Status Quo von Unterstützungsmöglichkeiten zur Schulentwicklung durch die Schulleitungen beurteilt wird und welcher zusätzliche Bedarf darüber hinaus besteht. Dies gilt einerseits hinsichtlich einer anstehenden Neuausrichtung der externen Schulevaluation in NRW, zeigt jedoch auch einen umfassenden Überblick über den aktuellen des modernen Schulleitungshandelns und -wissens sowie dessen (institutionell) gegebenen Grenzen auf. Des Weiteren wird ergründet, welche Ressourcen der Schulsteuerung innerhalb der Schulen bereits vorhanden sind und welche Faktoren Schulentwicklung begünstigen beziehungsweise verhindern. Die N = 30 Schulleitungen aller Schulformen aus ganz NRW wurden per Losverfahren ermittelt. Um die Einstellungen der Schulleiter/-innen in einer angenehmen Atmosphäre, ohne sozialen Druck und unter Gewährleistung von Anonymität, zu ermitteln, wurde das teilstandardisierte fokussierte Interview als Untersuchungsmethode gewählt (Bortz & Döring, 2006). Anschließend wurde das Interviewmaterial transkribiert und mit Hilfe der Software MAXQDA von zwei Expert/-innen kodiert. Die Kodierung erfolgte einerseits deduktiv auf der Grundlage des Interviewleitfadens und andererseits induktiv durch das Bilden von Kategorien bei der Durchsicht und Bearbeitung des Materials. Die Ergebnisse der Untersuchung deuten interessante Aspekte an, welche die Symbiose aus empirischer Bildungsforschung einerseits sowie einer sozialpsychologischen Perspektive andererseits anbieten. Unsere Ergebnisse zeigen, dass viele Schulen noch nicht ausreichende Kapazitäten entwickelt haben, um ihre Schulen selbstständig zu entwickeln. Dies bezieht sich beispielsweise auf Aspekte der Schulsteuerung durch Methoden der internen Evaluation. Die befragten Schulleiter/-innen der Stichprobe thematisieren außerdem die Anforderungen, welche an die moderne Leitungsfunktion einer Schule geknüpft sind. Des Weiteren zeigen sich außerdem explizit auf Schulleitungen bezogene Belastungsfelder und wichtige Ergebnisse hinsichtlich der Personalführung an Schulen. Die Chancen und Grenzen der Unterstützungsmöglichkeiten für Schulleitungen durch externe Schulevaluationen werden dargestellt und einer Diskussion unterzogen. Ein interessanter Befund ist darüber hinaus die von den Schulleitungen thematisierte hochkomplexe Personalführung an Schulen, welche genau beleuchtet wird.

Die Ergebnisse zeigen insgesamt, dass Schulen ein hochkomplexes Spannungsfeld darstellen, in welchem Unterstützungsbedarf besteht. Insbesondere auf den auf psychologischer Ebene angesiedelten Veränderungsbedarf wird in der Diskussion eingegangen.

Bortz, J. & Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Heidelberg: Springer.

Janssens, F. J.G./ van Amelsvoort, G. H.W.C.H. (2008). School self-evaluations and school inspections in Europe: An exploratory study. *Studies in Educational Evaluation*, 34, S. 15-23.

Luginbuhl, R./ Webbink, D./ de Wolf, I. F. (2009). Do Inspections Improve Primary School Performance? In: *Educational Evaluation and Policy Analysis*, Vol. 31, S. 221-237.

Müller, S., Dederig, K., & Bos, W. (2008). *Schulische Qualitätsanalyse in Nordrhein-Westfalen*. Köln: Kluwer.

Plowright, D. (2007). Self-evaluation and Ofsted Inspection. Developing an Integrative Model of School Improvement. *Educational Management Administration & Leadership*, Volume 35, No. 3, S. 373-393.

Semmer, N. K./ Jacobshagen, N. (2010). Feedback im Arbeitsleben – eine Selbstwert-Perspektive. In: *Gruppendynamik & Organisationsberatung*, Volume 41, S. 41.

Steins, G. (2005). *Sozialpsychologie des Schulalltags*. Stuttgart: Kohlhammer.

Steins, G. (2009). Widerstand von Lehrern gegen Evaluationen aus psychologischer Sicht. In: T. Bohl & H. Kiper (Hrsg.), *Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren*. (185-196). Bad Heilbrunn: Julius Klinkhardt.



## Wie lässt sich kollektive Innovationsbereitschaft von Lehrkollegien beeinflussen?

In ihrer Bedeutung als Reformträger erkannt, sind im Kontext der erweiterten Autonomie der Einzelschule in den letzten Jahren Voraussetzungen auf Seiten der Lehrkräfte ins Zentrum der Betrachtung gerückt (z.B. Seitz, 2008). Obgleich die Innovationsbereitschaft im Kollegium und damit Voraussetzungen auf kollektiver Ebene als ein wichtiger Erfolgsfaktor im Rahmen von Innovations- und Schulentwicklungsprozessen erkannt wurden (s. Holtappels, 1999), beschäftigt sich die Mehrheit einschlägiger Studien primär mit individuellen Voraussetzungen der Lehrenden für gelingende Reformumsetzung und Schulentwicklungsprozesse (z.B. Schumacher, 2008). Zugleich beschränken sich entsprechende Studien überwiegend auf eine Zustandsbeschreibung im Sinne einer Momentaufnahme, welcher eine längsschnittliche Betrachtung und damit auch eine Analyse einer möglichen Entwicklung dieser Voraussetzungen im Zeitverlauf (weitestgehend) verborgen bleibt. Um der Bedeutsamkeit der kollektiven Ebene des Lehrkollegiums einer Einzelschule Rechnung zu tragen, erscheint eine fundierte Auseinandersetzung mit kollektiver Innovationsbereitschaft von Lehrkollegien notwendig, in der insbesondere auch deren Entwicklung im Zeitablauf untersucht wird.

Kollektive Innovationsbereitschaft kann als ein kollektives, auf Innovationen ausgerichtetes, handlungslenkendes Bewusstsein verstanden werden. Sie lässt sich theoriegeleitet als ein mehrdimensionales Konstrukt modellieren (Buholzer, 2000; Klusemann, 2003; Gebert, 2004; Buske, 2013; Buske & Zlatkin-Troitschanskaia, 2009). In der im Vortrag präsentierten Studie wurden darüber hinaus Einflussfaktoren kollektiver Innovationsbereitschaft und deren Zusammenhänge zum Zielkonstrukt modelliert und querschnittlich sowie längsschnittlich untersucht. Die Modellierung der Beeinflussung kollektiver Innovationsbereitschaft basiert sowohl auf theoretischen Konzeptionen als auch empirischen Befunden verwandter einschlägiger Forschung. Sie verknüpft dabei u.a. schultheoretische Überlegungen, soziologische Grundpositionen, empirische Befunde der Schulentwicklungsforschung sowie der wirtschaftswissenschaftlichen Innovations- und Organisationsforschung in einem interdisziplinären Modellierungsansatz. Unter Rückgriff auf die Theorie des institutionellen Wandels (North, 1992) konnte zudem in die theoretische Modellierung der Beeinflussung kollektiver Innovationsbereitschaft ein weiterer Ansatz integriert werden, der eine Entwicklungsperspektive der Beeinflussung kollektiver Innovationsbereitschaft ermöglicht und damit erstmals eine Entwicklung kollektiver Innovationsbereitschaft im Zeitverlauf betrachten kann (s. Buske, 2013). Im Rahmen einer empirischen Überprüfung der Modellierung des Konstrukts kollektive Innovationsbereitschaft von Lehrkollegien kann seine mehrdimensionale Struktur auf Basis vorliegender Daten eines DFG-Projektes aus zwei Erhebungsrunden bestätigt werden. Beiden Erhebungen lagen schriftliche Befragungen des schulischen Personals an insgesamt 15 Berliner beruflichen Schulen knapp ein (t1) sowie zweieinhalb Jahre (t2) nach Inkrafttreten eines neuen Schulgesetzes mittels standardisierten Fragebogens zugrunde (N=896). Kollektive Innovationsbereitschaft lässt sich als ein mehrdimensionales formatives Konstrukt abbilden. In Bezug auf die theoretische Modellierung der Beeinflussung kollektiver Innovationsbereitschaft zeigen multiple lineare Regressionsanalysen modellkonform signifikante Einflüsse mehrerer theoriegeleiteter Prädiktoren auf der Ebene der Schule und der Schulleitung auf kollektive Innovationsbereitschaft (s. Buske, 2013). Zur Ausweitung der methodischen Analysen zur empirischen Überprüfung wurden Berechnungen mittels Strukturgleichungsmodellierung angeschlossen, die insbesondere der Abbildung des komplexen Zusammenhangsgefüges zwischen kollektiver Innovationsbereitschaft sowie deren potentiellen Einflussfaktoren Rechnung tragen. Im Rahmen des Vortrages wird sowohl auf Basis der bisherigen Befunde als auch der weiterführenden Analysen latenter Zusammenhänge mittels Strukturgleichungsmodellierung die übergeordnete Frage beantwortet werden, wodurch sich kollektive Innovationsbereitschaft von Lehrkollegien beeinflussen lässt. Hierzu werden im Anschluss an die vorzustellenden theoretischen Modellierungen erstens die ermittelten Befunde hinsichtlich der Erfassung und Existenzprüfung der theoretischen Modellierung kollektiver Innovationsbereitschaft präsentiert. In einem zweiten Schritt werden die Ergebnisse der Zusammenhänge zwischen kollektiver Innovationsbereitschaft und schulischen Faktoren diskutiert. Hinsichtlich der Veränderung kollektiver Innovationsbereitschaft werden drittens die längsschnittlich ermittelten Zusammenhänge kollektiver Innovationsbereitschaft und ihrer Einflussfaktoren vorgestellt und kritisch beleuchtet. Damit greift der Vortrag sowohl theoretisch als auch methodisch verschiedene Perspektiven auf die Thematik auf und führt sie zusammen, um schließlich Ansatzpunkte zur Förderung kollektiver Innovationsbereitschaft ableiten zu können.

Buholzer, A. (2000). Das Innovationsklima in Schulen. Aarau: Bildung Sauerländer.

Buske, R. (2013). Die Entwicklung kollektiver Innovationsbereitschaft von Lehrkollegien - Eine theoretische Modellierung und empirische Untersuchung im Längsschnitt. Landau: Verlag Empirische Pädagogik. (im Druck)

Buske, R. & Zlatkin-Troitschanskaia, O. (2009). Kollektive Innovationsbereitschaft von Lehrkollegien als Moderator für die Wirksamkeit bildungspolitischer Steuerung. In Böttcher, W., Dicke, J. N. & Ziegler, H. (Hrsg.), Evidenzbasierte Bildung. Wirkungsevaluation in Bildungspolitik und pädagogischer Praxis. Münster: Waxmann, 239-250.

Gebert, D. (2004). Innovation durch Teamarbeit: Eine kritische Bestandsaufnahme. Stuttgart: Kohlhammer.

Holtappels, H. G. (1999). Zeitorganisation und Schulkonzept als Schulentwicklungsfaktoren -Entwicklungsprozesse in Halbtagschulen. In Hofmann, J., Weishaupt, H., & Zedler, P. (Hrsg.), Organisationsentwicklung in Schulen, in Unternehmen und im sozialen Bereich. Erfurt: Pädagogische Hochschule, 83-110.

Klusemann, J. (2003). Typologie der Innovationsbereitschaft: Messung und Erklärung der Innovationsbereitschaft in Gruppen und Organisationseinheiten. Bern: Huber.

North, D. C. (1992). Institutionen, institutioneller Wandel und Wirtschaftsleistung. Tübingen: Mohr.

Schumacher, L. (2008). Wodurch wird die Bereitschaft von Lehrkräften zur Mitarbeit an Schulentwicklungsprojekten beeinflusst? In Lankes, E.-M. (Hrsg.), Pädagogische Professionalität als Gegenstand empirischer Forschung. Münster: Waxmann, 279-290.

Seitz, S. (2008). Der Lehrer als Innovator von Schule: Ein neues Professionsverständnis? Historische und empirische Analysen zum beruflichen Selbstverständnis von Lehrkräften unter dem Aspekt der Schulentwicklung. Hamburg: Kovac.





Sebastian Kempert (Humboldt-Universität zu Berlin), Regina Götz (Universität Würzburg), Catharina Tibken (Universität Würzburg), Andrea Beinicke (Universität Würzburg), Kristine Blatter (Otto-Friedrich-Universität Bamberg), Petra Stanat (Humboldt-Universität zu Berlin), Wolfgang Schneider (Universität Würzburg), Cordula Artelt (Otto-Friedrich-Universität Bamberg)

## **Wirksamkeit von Fördermaßnahmen der phonologischen Bewusstheit im Vergleich: Erste Ergebnisse eines musikalischen Trainings im Vorschulalter**

Sprache, insbesondere der Schriftsprache, wird eine besondere Rolle für die akademische Entwicklung beigemessen (Baumert & Schümer, 2001). Ihre Bedeutung für nahezu alle Schulfächer wurde nicht zuletzt in großen Schulleistungsstudien wiederholt nachgewiesen (z.B. Müller & Stanat, 2006). Es besteht weitgehende Einigkeit darüber, dass die Weichen für einen erfolgreichen Schriftspracherwerb bereits im Vorschulalter gestellt werden und eine Förderung in dieser Altersstufe daher besonders wirkungsvoll erscheint. Bemühungen, Kinder mit schwachen sprachlichen Voraussetzungen mit Förderprogrammen entsprechend zu unterstützen, zeigten jedoch bisher nur bedingt Erfolg (vgl. Roos, Polotzek & Schöler, 2010).

Für den Erwerb der Schriftsprache hat sich die phonologische Bewusstheit (pB) als eine zentrale Komponente erwiesen (Schneider & Näslund, 1993; Wagner et al., 1997). Diese beschreibt die Fähigkeit, seine Aufmerksamkeit auf formale Eigenschaften von Sprache auszurichten und sie in ihren lautlichen Eigenschaften zu analysieren. Für die pB zeigt sich das konkret im Erkennen von Reimen sowie in der Identifizierung von Wörtern als Teile von Sätzen, Silben als Teile von Wörtern und letztlich auch einzelnen Lauten als Teile von Wörtern. Typischerweise wird in der Literatur zwischen pB im weiteren Sinne (z.B. Reime und Silben erkennen) und pB im engeren Sinne (z.B. Anlauterkennung und Phonemsynthese) unterschieden (Skowronek & Marx, 1989).

In einigen Studien zeigte sich, dass die pB durch ein strukturiertes Training (Hören, lauschen lernen 1 und 2 [HLL 1 und 2]; Küspert & Schneider, 2008; Plume & Schneider, 2004) im Vorschulalter gefördert werden kann (Schneider, 2008; Schneider & Marx, 2008; Souvignier, 2003). Neuere Studien weisen zudem darauf hin, dass auch ein musikalisches Training geeignet sein könnte, entsprechende Trainingseffekte zu erzielen (Degé & Schwarzer, 2011). Dies wird darauf zurückgeführt, dass sprachlichen und musikalischen Verarbeitungsprozessen ähnliche Mechanismen der Lautkategorisierung zugrunde liegen (Lamb & Gregory, 1993). Ziel der aktuellen Studie ist es, zum einen die Wirksamkeit eines musikalischen Trainings zu bestätigen und zum anderen zu prüfen, ob sich eine Kombination aus musikalischer Förderung und HLL als besonders effektiv erweist. In einer längsschnittlich angelegten Trainingsstudie werden daher die Effekte einer alleinigen Förderung mit HLL sowie einer sukzessiven Kombination aus musikalischer Förderung und HLL überprüft und miteinander verglichen. Dabei werden wichtige Hintergrundvariablen und das sprachliche Ausgangsniveau der Kinder kontrolliert. Bei der Stichprobe der Studie handelt es sich um ca. 400 Kindergartenkinder in Würzburg, Bamberg und Berlin, die zwei unterschiedlichen Interventionsbedingungen und einer Kontrollgruppe zugewiesen wurden: Trainingsgruppe 1 nimmt im vorletzten Kindergartenjahr an einem musikalischen Training in Anlehnung an das Unterrichtswerk Musik und Tanz für Kinder (Nykrin et al., 2007) über 16 Wochen teil, welches dreimal die Woche à 20 Minuten durchgeführt wird. Im letzten Kindergartenjahr schließt sich zusätzlich das etablierte Verfahren HLL 1 und 2 zur Förderung der pB an, welches über 20 Wochen jeweils 10 bis 15 Minuten täglich durchgeführt wird. Trainingsgruppe 2 wird ausschließlich im letzten Kindergartenjahr mit dem Verfahren HLL 1 und 2 gefördert. Die Kontrollgruppe nimmt an den regulären Aktivitäten der Einrichtungen teil und wird nicht spezifisch gefördert. Insgesamt werden die Kinder somit über zwei Jahre im Rahmen von vier Messzeitpunkten untersucht. Zum jetzigen Zeitpunkt wurden zwei Messzeitpunkte realisiert. Neben der Erfassung der pB werden mit einer Testbatterie u.a. die kognitiven Grundfähigkeiten, sprachliche Voraussetzungen, der sozio-ökonomische Hintergrund und die musikalischen Fähigkeiten der Kinder mit erhoben, um für Vortestunterschiede kontrollieren zu können und mögliche Interaktionen zu identifizieren. Der vorliegende Beitrag fokussiert auf die Wirksamkeit des musikalischen Trainings unter Berücksichtigung der sprachlichen Ausgangslagen und des Zuwanderungshintergrunds der teilnehmenden Kinder. Vorläufige Analysen weisen auf die Wirksamkeit des musikalischen Trainings in Teilbereichen der pB hin. Ob sich die Ergebnisse ergänzend zu bisherigen Förderansätzen verhalten und somit ein Inkrement zu etablierten Verfahren bieten, wird im Beitrag diskutiert.

Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich (S. 323–407). Opladen: Leske + Budrich.

Küspert, P. & Schneider, W. (2008). Hören, lauschen, lernen: Sprachspiele für Kinder im Vorschulalter: Würzburger Trainingsprogramm zur Vorbereitung auf den Erwerb der Schriftsprache. Göttingen: Vandenhoeck & Ruprecht.

Müller, A. G. & Stanat, P. (2006). Schulischer Erfolg von Schülerinnen und Schülern mit Migrationshintergrund: Analysen zur Situation von Jugendlichen aus der ehemaligen Sowjetunion und der Türkei. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), Herkunftsbedingte Disparitäten im Bildungswesen (S. 221–256). Wiesbaden: VS Verlag.

Nykrin, R., Grüner, M. & Widmer, M. (2007). Musik und Tanz für Kinder. Mainz: Schott.

Plume, E. & Schneider, W. (2004). Hören, lauschen, lernen 2: Spiele mit Buchstaben und Lauten für Kinder im Vorschulalter: Würzburger Buchstaben-Laut-Training. Göttingen: Vandenhoeck & Ruprecht.

Roos, J., Polotzek, S. & Schöler, H. (2010). EVAS – Evaluationsstudie zur Sprachförderung von Vorschulkindern: Abschlussbericht. Heidelberg. Pädagogische Hochschule

Schneider, W. (2008). Prävention von Lese-Rechtschreibschwierigkeiten. In W. Schneider & M. Hasselhorn (Hrsg.), Handbuch der Pädagogischen Psychologie (S. 348-359). Göttingen: Hogrefe.

Schneider, W. & Marx, P. (2008). Früherkennung und Prävention von Lese-Rechtschreibschwierigkeiten. In F. Petermann & W. Schneider (Hrsg.), Enzyklopädie der Psychologie, Serie V (Entwicklungspsychologie), Band 7: Angewandte Entwicklungspsychologie (S. 237-273). Göttingen: Hogrefe.

Schneider, W. & Näslund, J. C. (1993). The impact of early metalinguistic competencies and memory capacity on reading and spelling in elementary school: Results of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC). European Journal of Psychology of Education, 8(3), 273-287.

Skowronek, H. & Marx, H. (1989). Die Bielefelder Längsschnittstudie zur Früherkennung von Risiken der Lese-Rechtschreibschwäche: Theoretischer Hintergrund und erste Befunde. Heilpädagogische Forschung, 15(1), 38-49.





## ***Divergent denken und kreativ schreiben in der Grundschule?***

Divergentes Denken beschreibt die Fähigkeit, möglichst viele und unterschiedliche, sowie ungewöhnliche Einfälle zu produzieren (Cropley, 2006; Preiser, 2006). Im Gegensatz zu konvergentem Denken, womit häufig eine schnelle und zielgerichtete Problemlösung angestrebt wird, sollen beim divergenten Denken ausdrücklich entfernte Assoziationen und ungewöhnliche Problemlösungsansätze ausprobiert werden, die letztlich - verglichen mit konvergenten Problemlösungen - zu einem elaborierteren Ergebnis führen können.

In Modellen zur Erklärung kreativen Verhaltens gilt die Fähigkeit, divergent zu denken, neben anderen individuellen und externalen Merkmalen als eine Determinante von Kreativität (z. B. Urban, 2004), wobei divergentes Denken nicht mit Kreativität gleichzusetzen ist (ebd.; Runco & Acar, 2012). Divergentes Denken kann durch Techniken wie beispielsweise Brainstorming geschult werden (Osborn, 1957) und wird in der betrieblichen Weiterbildung eingesetzt, um Output zu optimieren (z. B. Preiser & Buchholz, 2008). Innerhalb der sprachlichen Domäne existieren empirische Hinweise darauf, dass sich divergentes Denken positiv auf Outcome-Variablen auswirkt: Wang (2012) fand an einer studentischen Stichprobe positive Zusammenhänge zwischen Aspekten divergenten Denkens und Einstellungen zum Lesen bzw. Schreiben. Baer (1996) untersuchte mit einem experimentellen Design die Wirksamkeit eines domänenspezifischen Trainings zur Förderung divergenter Denkstrategien auf spätere Schreibprodukte von Siebtklässlern und berichtet von positiven Effekten des Trainings. Studien dieser Art sind bislang allerdings rar und es stellt sich die Frage, ob sich derartige Zusammenhänge bereits in früheren Schulstufen zeigen. Mit der präsentierten Studie soll an einer Stichprobe von Grundschulern den folgenden Fragestellungen nachgegangen werden:

Lassen sich Facetten divergenten Denkens (Flüssigkeit und Flexibilität) im Grundschulalter erfassen? Zeigen sich (unter Kontrolle von kognitiven Grundfähigkeiten) Zusammenhänge zwischen divergentem Denken und der Wortschatzleistung von Grundschulkindern?

Zur Überprüfung der Fragestellung dienen Daten aus der Grundschulstudie PERLE (Lipowsky, Faust & Kastens, 2013). Verbale Flüssigkeit und verbale Flexibilität als Facetten divergenten Denkens wurden mit einem eigenentwickelten Instrument erfasst (Berner, Theurer & Lipowsky, 2012), das sich an vorhandenen Testverfahren zur Erfassung divergenter Denkfähigkeiten anlehnt. Dieses kam am Ende des dritten und am Ende des vierten Schuljahrs zum Einsatz. Die Wortschatzleistung wurde mit dem entsprechenden Subtest des KFT (Heller & Perleth, 2000) und die kognitive Grundfähigkeit mit der Skala 1 des CFT (Weiß & Osterland, 1997) gemessen. Erste Analysen zeigen, dass eine Messung divergenter Denkfähigkeiten gelungen ist. Sowohl die Flüssigkeit als auch die Flexibilität der Schülerantworten konnte von mehreren Ratern übereinstimmend beurteilt und für beide Klassenstufen faktorenanalytisch mit akzeptablem Fit separiert werden (3. Schuljahr:  $\chi^2/df=4.72$ ;  $p<.01$ ;  $RMSEA=.08$ ;  $SRMR=.06$ ;  $CFI=.93$ ;  $TLI=.90$ ; 4. Schuljahr:  $\chi^2/df=2.23$ ;  $p<.01$ ;  $RMSEA=.05$ ;  $SRMR=.05$ ;  $CFI=.96$ ;  $TLI=.95$ ). Korrelationsanalysen zeigen im dritten und vierten Schuljahr niedrige bis moderate Zusammenhänge zwischen der Wortschatzleistung und beiden Facetten divergenter Denkfähigkeit ( $r$  von  $.11^*$  bis  $.24^{**}$ ). Regressionsanalysen deuten an, dass auch unter Kontrolle vorheriger Wortschatzleistung und des IQ die Facette Flexibilität des divergenten Denkens zwar einen geringen, aber dennoch unique Anteil an Varianz der Wortschatzleistung am Ende der Grundschulzeit erklären kann ( $\beta=.10^*$ ). Flüssigkeit besitzt hingegen keine Vorhersagekraft für die Wortschatzleistung am Ende der Grundschulzeit.

Die Ergebnisse der Studie deuten darauf hin, dass Aspekte divergenten Denkens sich bereits im Grundschulalter positiv auf Schreibprodukte auswirken können. In weiterführenden Analysen soll deshalb die Annahme geprüft werden, dass divergentes Denken einen Effekt auf die Qualität von Schreibprodukten, welche von den Kindern im zweiten, dritten und vierten Schuljahr angefertigt worden sind, hat.

- Baer, J. (1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior*, 30(3), 183-187.
- Berner, N., Theurer, C. & Lipowsky, F. (2012). Ist Kreativität messbar? Zur Erfassung kreativer Fähigkeiten im Forschungsprojekt PERLE. *Erziehung & Unterricht*, 162(5-6), 442-453.
- Cropley, A. (2006). Kreativität und Kreativitätsförderung. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (pp. 362-370). Weinheim: Beltz.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeits-Test (Rev.) für 4. Klassen (KFT 4 R)*. Göttingen: Beltz-Testgesellschaft.
- Lipowsky, F., Faust, G. & Kastens, C. (2013). *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren*. Münster: Waxmann.
- Osborn, A. F. (1957). *Applied imagination*. New York: Charles Scribner's Sons.
- Preiser, S. (2006). *Kreativität*. In K. Schweizer (Ed.), *Leistung und Leistungsdiagnostik* (pp. 51-66). Heidelberg: Springer.
- Preiser, S. & Buchholz, N. (2004). *Kreativität. Ein Trainingsprogramm für Alltag und Beruf*. Heidelberg: Asanger.
- Runco, M. A. & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24(1), 66-75.
- Urban, K. K. (Ed.) (2004). *Kreativität: Herausforderung für Schule, Wissenschaft und Gesellschaft*. Münster: LIT.
- Wang, A. (2012). Exploring the relationship of creative thinking to reading and writing. *Thinking Skills and Creativity*, 7, 38-47.
- Weiß, R. H. & Osterland, J. (1997). *CFT1. Grundintelligenztest Skala 1*. Göttingen: Hogrefe.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## **Mode-Effekte beim Rechtschreiben? Ein Vergleich eines computer- und papierbasierten Rechtschreibtests in Klasse Fünf**

Rechtschreibtests werden sowohl in Large-Scale-Assessments als auch im Rahmen der Einzeldiagnostik vielfach eingesetzt. Der Großteil an Test- und Diagnoseinstrumenten, die auf einem empirisch validierten Kompetenzmodell basieren, erfolgt papierbasiert (Frahm & Blatt, 2011). Dies hat den Nachteil, dass die Transkription und Auswertung der Tests zeit- und kostenintensiv sind und linguistische Grundkenntnisse voraussetzen. Computerisierte Tests vereinfachen dagegen nicht nur die Auswertung, sondern auch die Entwicklung adaptiver Tests sowie die Generierung von Feedback (z. B. Tierney & Charland, 2007). Hinzu kommt, dass Computer mittlerweile einen integralen Teil der Lebenswelt der Schülerschaft darstellen (z. B. Senkbeil & Wittwer, 2007). Bevor jedoch papierbasierte Tests zu Computertests umgewandelt werden, ist zu klären, inwiefern der Testmodus Effekte, sog. Mode-Effekte, hervorruft. Für den deutschsprachigen Raum liegen hierzu keine Ergebnisse vor (Frahm, 2012). Im Rahmen des Nationalen Bildungspanels (NEPS) wird daher erforscht, inwieweit computerisierte Testverfahren genutzt werden können, um die Rechtschreibkompetenz genauso zu messen wie papierbasierte Tests. Hierzu wird der Sprachsystematische Rechtschreibtest (SRT) eingesetzt. Dieser Test basiert auf einem fünfdimensionalen Kompetenzmodell (vgl. Blatt et al., 2011) und wird im Rahmen der NEPS-Studie in den Klassenstufen fünf bis neun eingesetzt. Dieser Untersuchung, die Kern des Posterbeitrages ist, liegen folgende Forschungsfragen zugrunde:

- Bewirkt die Veränderung des Schreibmodus von handschriftlichen zu computerisierten SRTs eine Veränderung der Rechtschreibleistung insgesamt bzw. in den Dimensionen des Kompetenzmodells und wenn ja, in welcher Weise?
- Verändert sich das zugrunde liegende Kompetenzmodell in Abhängigkeit vom Schreibmodus und wenn ja, wie?
- Entstehen durch die Veränderung des Schreibmodus vom handschriftlichen zum computerisierten Test Nachteile für bestimmte Schülergruppen?

Um die Forschungsfragen zu beantworten, wurde eine Mode-Effekt-Studie in Klassenstufe fünf durchgeführt. Die Stichprobe (N=531) wurde randomisiert in zwei Gruppen unterteilt, die den identischen Test mit unterschiedlichen Modus bearbeiteten (papierbasiert vs. computerisiert). Ergänzend wurde ein Fragebogen zur Erhebung der Konstrukte Demographie, Motivation, Computererfahrung und Rechtschreibmotivation und –Selbstkonzept sowie ein Test zur Feststellung der Tippfertigkeit eingesetzt.

Die Daten des papierbasierten Tests wurden zunächst transkribiert. Alle Testdaten wurden daraufhin mit einem eigens für den SRT entwickelten Kodierprogramm kodiert (Frahm, 2012). Neben deskriptiven Analysen werden auf dem Poster auch Skalierungsergebnisse auf IRT-Basis sowie qualitative Einzelfallanalysen präsentiert. Die Skalierungsergebnisse geben zum einen Aufschluss über den Itemfit und die Reliabilität des computerisierten Tests. Zum anderen liefern sie vergleichend Informationen zur Itemschwierigkeit sowie zur Personenfähigkeit beider Testmodi.

Die Ergebnisse der Studie zeigen, dass die Leistung beider Gruppen signifikant unterschiedlich ist. Die Kompetenzstruktur dagegen ist in beiden Gruppen vergleichbar. Hinsichtlich der geprüften Hintergrundvariablen lassen sich keine systematischen Benachteiligungen zeigen. Ausschlaggebend für die schwächere Leistung ist vor allem eine mangelnde Tippfertigkeit und damit einhergehend ein sehr langsames Tippen.

Die Ergebnisse liefern eine Grundlage für eine Nutzung computerisierter Tests, insbesondere im unterrichtlichen Kontext, z. B. im Rahmen von formativen Assessment (Black & William, 1998). Darüber hinaus bestätigen sie unter Berücksichtigung des internationalen Forschungsstandes, dass Mode-Effekte eindeutig stichproben- und testabhängig sind, die Ergebnisse also nicht verallgemeinerbar sind (Kröhne & Martens, 2011).

Auf dem Poster sollen die Forschungsfragen, das Studiendesign sowie die Ergebnisse veranschaulicht werden.

Black, P. & William, D (1998). Assessment and Classroom Learning. Assessment in Education 5(1) pp. 7-71.

Blatt, I., Voss, A., Kowalski, K. & Jarsinski, S. (2011). Empirische Modellierung und Messung von Rechtschreibkompetenz. U. Bredel (edt) Weiterführender Orthographieunterricht (Deutschunterricht in Theorie und Praxis). Baltmannweiler: Schneider Verlag Hohengehren.

Frahm, Sarah (2012). Computerbasierte Testung der Rechtschreibleistung in Klasse fünf - eine empirische Studie zu Mode-Effekten im Kontext des Nationalen Bildungspanels. Berlin: Logos-Verlag.

Frahm, Sarah & Blatt, Inge (2011): "Rechtschreibtests" In: Bredel, U. (Hrsg.): Weiterführender Orthographieunterricht. Baltmannweiler: Schneider Verlag Hohengehren. (DTP (Deutschunterricht in Theorie und Praxis)).

Kröhne, U. & Martens, T (2011). Computer-based competence tests in the national educational panel study. Zeitschrift für Erziehungswissenschaft. Special issue, 14 (2) pp. 169-186.

Senkbeil, M. & Wittwer, J. (2007). Die Computervertrautheit von Jugendlichen und Wirkungen der Computernutzung auf den fachlichen Kompetenzerwerb. PISA-Konsortium Deutschland (eds.), PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie. Waxmann. pp. 277-307.

Tierney, R. D., & Charland, J. (2007). Stocks and prospects: Research on formative assessment in secondary classrooms. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.



## **Untersuchung der Eignung von Fibeln für einen systematischen Schriftspracherwerb -- Analyse von Fibeltexten durch automatische Sprachverarbeitungsmethoden**

Fibeln haben die Funktion, Kindern systematisch schriftsprachliche Kompetenzen zu vermitteln. Sie präsentieren Inhalte in Form von Schrift als Orthographie mit dem Ziel der Aneignung. Dabei stellt die memorierende Aneignung der graphischen Formen der Buchstaben eine geringere Aufgabe für das Lesen und Schreiben lernen dar als das Erlernen ihrer unterschiedlichen Funktionen in verschiedenen strukturierten Wörtern (Vgl. <r> in <Rot>, <Tor>, <Roter>). Da die orthographischen Zeichensystem Standardsprache symbolisieren, die im Gesprochenen in dieser expliziten Form nicht vorkommen, haben die Kinder mit dem Schrifterwerb zugleich die Chance, ihre alltäglichen Spracherfahrungen im Sinne eines umfassenderen Sprachausbaus zu erweitern und dabei die Systematik der Standardsprache zu entdecken und, bei entsprechender Hinführung zur Schrift, Ihre Kompetenzen im Umgang mit Symbolen zu festigen (literacy).

Lernen erfolgt generell in einem Kontinuum der Aneignung, in dem der Grad der Komplexität allmählich zunimmt und dabei die jeweils nächste Form an das bereits erworbene Wissen anknüpft. An die vorschulisch erworbenen Fähigkeiten der Kinder anzuschließen, erfordert eine didaktische Reduktion am Schrifanfang (also für die Gestaltung der Fibeln), die den Kindern Gelegenheit gibt, systematisch die Schrift mit Bezug auf ihr bereits erworbenes phonologisches/silbisches Wissen als Einstieg in die orthographische Systematik kennenzulernen (die Differenzierung nach betonten und unbekanntem Silben sowie die Wahrnehmung der vier phonologischen Gestalten deutscher Wörter, vgl. Maas, 2003; Eisenberg, 1995; Röber, 2009). Da die morphologischen Markierungen weitgehend die phonologischen Markierungen (Schärfungs- und Dehnungsmarkierung) aufnehmen, setzt die Thematisierung der Morphologie mit ihren unterschiedlichen Facetten das didaktische Kontinuum fort. Daran schließen sich die syntaktischen und textuellen orthographischen Markierungen (Verbflexionen, Markierungen nominaler Gruppen mit der Großschreibung, Textgestaltung durch Zeichensetzung) an und weiten damit zugleich das Sprachwissen der Lerner systematisch aus. Betreffen diese Kriterien die Angebote für die kognitive Arbeit der Kinder, ist Lernen mit dem Ziel der Automatisierung als kognitive Entlastung zugleich an Üben gebunden, so dass die Frage nach der orthographischen Systematik in den Fibeln mit der nach ihrer methodischen Gestaltung zusammenhängt.

Anhand von Instrumenten aus der automatischen Sprachverarbeitung werden Texte aus Fibeln entsprechend der genannten Kriterien auf die Progression ihrer sprachlichen Angebote für die Kinder (Systematik, Übungsmöglichkeiten) hin untersucht und verglichen. Für die analytische Arbeit wird Balloon, ein Textverarbeitungsinstrument für die deutsche Sprache, eingesetzt. Das methodische Vorgehen sowie erste Ergebnisse der Analysen werden im Vortrag vorgestellt.

Auszüge aus Analysen auf Silben- und Satzebene sind hier kurz erwähnt.

- Die ausgewählten Fibeln lassen sich in der Aufnahme von Wiederholungen stark unterscheiden: Eine der untersuchten Fibeln baut schnell einen größeren Wortschatz auf, der erfordert, dass schon früh eine große sprachliche Komplexität verarbeitet wird, und reduziert damit Üben durch Wiederholungen. Ein Vergleich der Fibeln lässt einen quantitativen Unterschied von etwa 300 Wörtern auf den ersten 54 Seiten erkennen - ein Resultat, das Aussagen über die didaktische Konzeption der einzelnen Lehrbücher zulässt.
- Einer besonderen Analyse bedürfen die neuen Silbenfibeln, die im Kontrast zu den herkömmlichen Fibeln mit einem dezidierten sprachanalytischen Anspruch auftreten. Detailliertere Statistiken resp. der vier phonologischen Gestalten ermöglichen eine Kontrolle, inwieweit dieser Anspruch eingehalten wurde. Ein Vergleich zeigt, dass einige Fibeln stärker von Silbenwiederholung profitieren können als andere, detaillierte Analysen werden folgen.
- Komplexität auf Satzebene wurde informationstheoretisch erfasst. Anhand der Kreuzentropie eines Satzes mit einem Wortbigramm-Wahrscheinlichkeitsmodell kann auf die Vorhersagbarkeit eines Worts gegeben sein linker Textkontext rückgeschlossen werden. Weiterhin können Satzlänge, Anzahl Sätze pro Seite, und andere Werte automatisch ermittelt werden. Die untersuchten Fibeln unterscheiden sich hinsichtlich des Einstiegsniveaus sowie des Anstiegs der Komplexität.

Berkling, K., "A Case Study Using Data Exploration of Spelling Errors Towards Designing Automated Interactive Diagnostics," in Proc. 3rd Workshop on Child, Computer Interaction (Interspeech 2012 Post-Conf. Workshop), Portland, USA, September 2012.

Berkling, K.; Fay, J.; Stüker, S. Speech Technology-based Framework for Quantitative Analysis of German Spelling Errors in Freely Composed Children's Texts. Interspeech, Florenz, 2011.

Eisenberg, P.: Das Wort. Der Buchstabe und die Schriftstruktur des Wortes. In: Drosdowski, G. (Hrsg.): Duden. Bd.4: Grammatik der deutschen Gegenwartssprache. Mannheim: Dudenverlag 1995 S.56-84

Lorz, J.. Die Präsenz der silbenanalytischen Methode in den Silbenfibeln "Piri" und "ABC der Tiere". GRIN Verlag, 2012.

Lynne GD, D., Seymour, Ph., and Hill, Sh. "A small-to-large unit progression in metaphonological awareness and reading?." The Quarterly Journal of Experimental Psychology: Section A 53.4 (2000): 1081-1104.

Maas, U.: Orthographie und Schriftkultur. (Nicht nur im Deutschen). Universität Osnabrück: Arbeitsskript zur Vorlesung im SS 2003.

Ravid, D. "Morphological Scaffolding in Learning to Spell: A Cross-Linguistic Review." Spelling Morphology. Springer US, 2012. 41-55.

Reichel, U.D. PerMA and Balloon: Tools for string alignment and text processing, Proc. Interspeech 2012, Portland, Oregon. Paper 346.

Röber, Ch., Die Leistungen der Kinder beim Lesen- und Schreibenlernen. Grundlagen der Silbenanalytischen Methode, Schneider Verlag Hohengehren 2009.



## „Schreibunterricht, so wie ich ihn sehe.“- Ergebnisse einer Lehrer/-innen- und Schüler/-innenbefragung zum Schreibunterricht an den österreichischen AHS und BHS im Jahre 2011

Die neue zentralisierte schriftliche Reife- und Diplomprüfung Deutsch in Österreich (SRDP) (Start 2014/15) ist das Resultat eines internationalen bildungspolitischen Trends, der auf Standardisierung sowie landesweite und internationale Vergleichbarkeit der Abschlüsse zielt.

Die Fragebogenerhebung für Lehrer/-innen und Schüler/-innen 2011 zum Schreibunterricht in der Sekundarstufe II der AHS (Allgemeinbildende Höhere Schule) und der BHS (Berufsbildende Höhere Schule) ist Teil der begleitenden Forschung bzw. eines umfassenden Maßnahmenpaketes zur Einführung der zentralisierten schriftlichen Reife- und Diplomprüfung Deutsch in Österreich (SRDP). Die Fragebogenerhebung soll dazu beitragen, für die Implementation eines veränderten Schreibunterrichts und die damit einhergehende Lehrer/-innen- und -fortbildung relevante Auskünfte zu erhalten. Die didaktischen Überlegungen, die dem Fragebogen zu Grunde liegen, sind mit dem theoretischen Konzept der an Kompetenzen orientierten Reifeprüfung akkordiert (vgl. [www.bifie.at](http://www.bifie.at)).

Das Konzept der Erhebung ist Annahmen geleitet, beispielsweise was die Verbreitung der Textsorten und die Verbindung von Lesen/Schreiben anlangt. Es wird Fragen nachgegangen, wieweit sich die Schüler/-innen in integrierenden Lese-Schreibaktivitäten üben, mit „Hilfs- und Transfertexten“ (Bräuer & Schindler, 2011) und „Lernformen in engerem Sinn“ (Pohl & Steinhoff, 2010) arbeiten und wie die Aufgabenstellungen wahrgenommen werden. Die reflexive Praxis des Schreibunterrichts, also Peer-Feedback und kooperative Schreibformen (Becker-Mrotzek & Böttcher, 2006), interessiert ebenso wie der Erwerb von digitaler Literalität (Alvermann, 2009) und Fördermaßnahmen.

An der bundesweit durchgeführten Erhebung haben sich 1880 SchülerInnen und 94 Lehrpersonen beteiligt.

Die in beiden Fragebögen inhaltlich identischen Fragen erlauben es, die Aussagen der LehrerInnen in Beziehung zu den Aussagen der SchülerInnen zu setzen und die Ergebnisse deskriptiv darzustellen und zu vergleichen.

In der Auswertung der SchülerInnenendaten wurden Merkmale herausgearbeitet zu: Unterricht von SchülerInnen mit/ohne Migration, m/w, in den Schultypen AHS/BHS und – in Verbindung mit den im Fragebogen erhobenen Leistungsnoten im Fach Deutsch – schwacher und sehr guter SchülerInnen (immer laut Selbstausskunft der SchülerInnen). Die Daten erlauben auch einen Vergleich von Aussagen von SchülerInnen mit unterschiedlicher Einstellung/Assoziation zum Schreiben (positive/negative und neutrale bzw. nicht zuzuordnenden Assoziationen).

Beispiele für Ergebnisse aus den SchülerInnenendaten:

Erörterungen und Interpretationen sind die Textsorten, die in der Wahrnehmung der Schüler(-innen) am häufigsten verfasst wurden; Aufgabenstellungen werden an den BHS expliziter formuliert als an den AHS, Sachtexte werden an den BHS häufiger bearbeitet als an den AHS. Buben geben signifikant häufiger als Mädchen an, medientypische Aspekte beim Schreiben anzuwenden.

DAZ-Schülerinnen hingegen geben signifikant häufiger an, weniger Feedback seitens der Mitschüler/innen oder Lehrer/innen zu erhalten, weniger zu recherchieren, aber häufiger Schreibblockaden zu haben. DAZ- Schüler/innen haben insgesamt eine geringe Selbsteinschätzung.

SchülerInnen mit positiven Assoziationen scheinen sich grundsätzlich besser daran zu erinnern, unterschiedliche Textsorten erklärt bekommen zu haben, als SchülerInnen mit negativen Assoziationen. Die Ergebnisse der SchülerInnen spiegeln in Beziehung zu den LehrerInnenendaten die sehr unterschiedlichen Sichtweisen der beiden Befragtengruppen wider. Im Vortrag sollen nun ausgewählte Aspekte im Vergleich vorgestellt werden.

Besonders in den Blick genommen werden dabei das Zusammenspiel von Teilaufgaben in der Textproduktion, situierte Aufgabenstellungen, die Arbeit mit unterschiedlichen Textsorten, die Aneignung von literalen Kompetenzen in Aufgabensettings im Umgang mit Texten sowie Textkompetenz in der Verschränkung von Lesen und Schreiben.

Das sich abzeichnende Bild der großen Unterschiede zwischen LehrerInnen und SchülerInnen wirft die grundsätzliche Frage auf, in welchen Szenarien die SchülerInnen nicht nur nachhaltig lernen, sondern auch, wie diese Lernumwelten gestaltet werden müssen, damit sie einer bewussten Wahrnehmung und Reflexion zugänglich sind.

Fußnote: Zur begleitenden Forschung der zentralisierten schriftlichen Reife- und Diplomprüfung Deutsch in Österreich (SRDP) gehört zudem die Sammlung und Analyse von SchülerInnenendaten, Aufgabenstellungen (und Leistungsbeurteilungen) aus der alten Matura 2009, des Feldtests der neuen Matura 2012 und des ersten Schulversuchs 2013 (Gesamtleitung seit 2012 Margit Böck, AAU Klagenfurt).

Literatur: Abraham, Ulf/ Saxalber, Annemarie: Typen sprachlichen Handelns („Operatoren“) in der neuen Reifeprüfung Deutsch. In: *ide. Informationen zur Deutschdidaktik. Zeitschrift für den Deutschunterricht in Wissenschaft und Schule*, 36. Jg., Heft 1 2012: Reifeprüfung Deutsch. Innsbruck – Wien – Bozen: Studienverlag. S. 36 – 40.

Alvermann, Donna E. (2009): *New literacies. Schnittmengen der Interessen von Heranwachsenden und der Wahrnehmung von Lehrerinnen und Lehrern*. In: Bertschi-Kaufmann/Rosebrock (Hrsg.), S. 91-103.

Becker-Mrotzek, Michael; Böttcher, Ingrid (2006): *Schreibkompetenz entwickeln und beurteilen*. Berlin: Cornelsen.

Bertschi-Kaufmann, Andrea; Rosebrock, Cornelia (Hrsg.) (2009): *Literalität. Bildungsaufgabe und Forschungsfeld*. Weinheim; München: Juventa.

Bräuer, Gerd; Schindler, Kirsten (2011): *Authentische Schreibaufgaben – ein Konzept*. In: Bräuer, Gerd; Schindler, Kirsten (Hg.): *Schreibarrangements für Schule, Hochschule, Beruf*. Freiburg im Breisgau: Fillibach, S. 12-63.

Pohl, Thorsten; Steinhoff, Torsten (2010): *Textformen als Lernformen*. In: Pohl; Steinhoff (Hrsg.): *Textformen als Lernformen*, S. 5-26

Pohl, Thorsten; Steinhoff, Torsten (Hrsg.) (2010): *Textformen als Lernformen*. Duisburg: Gilles & Francke.

Saxalber, Annemarie/Estler, Ursula (Hg. 2010): *Schreibprozesse begleiten. Vom schulischen zum universitären Schreiben*. Publikation in der Reihe *ide-extra: Band 17 (= Reihe des AECC Deutsch Klagenfurt)*

Saxalber, Annemarie /Werner Wintersteiner (Hg.): *ide. Informationen zur Deutschdidaktik. Zeitschrift für den Deutschunterricht in Wissenschaft und Schule*, 36. Jg., Heft 1/2012: Reifeprüfung Deutsch. Innsbruck – Wien – Bozen: Studienverlag

Witschel, Elfriede/ Saxalber, Annemarie: *Von fremden Texten zu eigenen Texten. Folgerungen aus einer LehrerInnenbefragung für den Schreibunterricht in der Sekundarstufe II*. In: *ide. Informationen zur Deutschdidaktik. Zeitschrift für den Deutschunterricht in Wissenschaft und Schule*, 36. Jg., Heft 1 2012: Reifeprüfung Deutsch. Innsbruck – Wien – Bozen: Studienverlag. S. 74 – 84.



Ina Plöger (WWU Münster), Miriam Leuchter (WWU Münster)

## ***Diagnose von Inhalts- und Prozesswissen bei Kindern und Jugendlichen in Mathematik und Naturwissenschaften***

Chair(s): Miriam Leuchter (WWU Münster, Didaktik des Sachunterrichts)

Diskutant/in: Henrik Saalbach (Universität des Saarlandes, Bildungswissenschaften)

Ziel pädagogisch-psychologischer Diagnostik ist eine möglichst akkurate Beurteilung der aktuellen Ausprägung von Personenmerkmalen, beispielsweise von Fähigkeiten der Lernenden, sowie von Aufgabenmerkmalen (Wild & Krapp, 2001). Diagnostische Urteile beziehen sich auf bestimmte Domänen wie z.B. Mathematik oder Naturwissenschaften, und betreffen Inhaltswissen, etwa Algebra oder Stabilität, oder Prozesswissen, zum Beispiel Hypothesentestung oder Erkenntnisgewinnung.

Da Wissensdiagnostik Entscheidungen auf der Mikroebene (z.B. Lernprozess-Steuerung) und der Makroebene (z.B. längerfristige Bildungsentscheidungen) beeinflusst, ist eine zutreffende Einschätzung der Fähigkeiten von Lernenden und der Aufgabenschwierigkeit von hoher Qualität und Güte wichtig.

Formelle Diagnosen stellen eine Möglichkeit dar, Fähigkeiten von Lernenden sowie Aufgabenschwierigkeiten mithilfe wissenschaftlich erprobter Methoden gezielt und systematisch zu beurteilen. Beispielsweise liegen zur Erfassung der Intelligenz bereits Tests vor, die empirisch erprobt sind. In vielen anderen Bereichen ist hingegen offen, welches Wissen Kinder und Jugendliche bezogen auf bestimmte Aufgaben mitbringen und wie es erfasst werden kann. Im Symposium werden Möglichkeiten vorgestellt, Fähigkeiten von Kindern und Jugendlichen in Mathematik und Naturwissenschaften hinsichtlich des Inhalts- und Prozesswissens zu diagnostizieren. Damit zusammenhängende Analysen der Aufgabenschwierigkeiten werden präsentiert.

In Studie 1 (Ina Plöger) werden Einschätzungen von Kindern zwischen 4 und 8 Jahren zur Stabilität von Bauklotz-Anordnungen analysiert. Es wird untersucht, ob sich Merkmale von Bauklotz-Anordnungen identifizieren lassen, die mit den Leistungen der Kinder zusammenhängen.

Studie 2 (Peter Edelsbrunner) analysiert das Verständnis von Dritt- und Viertklässlern hinsichtlich experimenteller Designs. Ziel der Studie ist es, ein Multiple-Choice-Instrument zur Messung der Aspekte Variablenkontrolle und Hypothesentestung zu entwickeln.

Studie 3 (Susanne Mannel) untersucht und erprobt ein Testinstrument zur Erfassung des Prozesswissens im Bereich Erkenntnisgewinnung von Lernenden zu Beginn der Jahrgangsstufe 5.

Studie 4 (Esther Ziegler) erforscht das Vorwissen von Kindern in der sechsten Klasse zur Algebra und dessen Auswirkungen auf spätere Lernergebnisse.



## ***Einschätzungen zur Stabilität von Bauklötz-Anordnungen: Unterschiedliche Merkmale und deren Zusammenhang mit den Leistungen von Kindern zwischen 4 und 8 Jahren***

### Theoretischer Hintergrund

Beim Bauen mit Bauklötzen sammeln Kinder bereits früh Erfahrungen zu Aspekten der Stabilität von Bauwerken, wie z.B. Standfestigkeit und Gleichgewicht. Bisherige Studien erfassen die intuitiven Vorstellungen zur Stabilität anhand von unterschiedlich geformten Einzelobjekten mit zwei Methoden: a) Kinder balancieren Einzelobjekte aus oder b) beurteilen abgebildete Einzelobjekte hinsichtlich ihrer Stabilität auf einer Plattform. Kategorisierungsaufgaben zeigen, dass Kinder dazu tendieren, nach oberflächlichen Prinzipien zu klassifizieren (vgl. Gentner & Namy, 1999). Dieser Befund lässt sich auf die Studien zu intuitiven Vorstellungen der Stabilität von Objekten übertragen:

a) Das Ausbalancieren von unterschiedlich geformten Einzelobjekten gelingt Kindern zwischen 3 und 8 Jahren mit kontinuierlichem, altersabhängigem Leistungszuwachs. Dabei fällt es ihnen leichter, symmetrische als asymmetrische Objekte auszubalancieren (Krist et al., 2005), da bei Letzteren der Massenmittelpunkt als tieferliegendes Prinzip berücksichtigt werden muss. Mit 7.4 Jahren beachten Kinder den Massenmittelpunkt; mit 6.8 Jahren richten sie asymmetrische Objekte an dem oberflächlichen Merkmal „geometrische Mitte“ aus (Bonawitz et al., 2012), welches bei symmetrischen Objekten mit dem Massenmittelpunkt übereinstimmt.

b) Vergleichbare Befunde ergeben sich bei der beobachtungs-basierten Beurteilung, ob Einzelobjekte auf einer Plattform liegen bleiben, über die sie seitlich hinausragen: Eine Studie von Krist zeigt, dass 5- und 6-Jährige eine Sensibilität für die geometrische Mitte eines Objekts fast perfekt entwickelt haben, aber die Sensibilität für den Massenmittelpunkt bei 6-jährigen Kindern kaum vorhanden ist (Krist, 2010). Dies überrascht, da Studien auf Basis von Blickzeitenmessungen bereits Säuglingen ein intuitives Verständnis für den Massenmittelpunkt zusprechen (Baillargeon & Hanco-Summers, 1990). Ergebnisse von Studien beider Methoden zeigen, dass es Kindern schwer fällt, tieferliegende, nicht offensichtliche Eigenschaften zu berücksichtigen. Bisherige Studien haben dies mit Einzelobjekten analysiert; kaum untersucht wurden Anordnungen aus mehreren Objekten, wie z.B. unterschiedlich angeordnete Bauklötze. Offen bleibt, ob Kinder bei der Einschätzung der Stabilität von unterschiedlich angeordneten Bauklötzen ebenfalls auf äußere Merkmale fokussieren (z.B. geometrische Mitte).

### Fragestellung

Lassen sich Merkmale von Bauklötz-Anordnungen identifizieren, die mit den Leistungen von Kindern zwischen 4 und 8 Jahren zusammenhängen?

### Methode

Zur Diagnose des Wissens zur Stabilität von Bauklötz-Anordnungen sehen die Kinder Fotos, bei denen sie einschätzen, ob die Anordnung stabil bleibt, wenn ein bestimmter Bauklötz entfernt wird. 20 verschiedene Anordnungen variieren in ihrer Komplexität (z.B. Anzahl der Bauklötze, Lage der geometrischen Mitte/des Massenmittelpunkts). Die Stichprobe besteht aus N=205 Kindern (Alter in Jahren M=6.7; SD=.38).

### Ergebnisse

Erste Ergebnisse deuten darauf hin, dass die Kinder durchschnittlich 63.42% aller Bauklötz-Anordnungen bezüglich ihrer Stabilität richtig einschätzen und somit signifikant über der Ratewahrscheinlichkeit liegen ( $t(204)=16.176$ ,  $p\leq.001$ ,  $d=0.75$ ). Die Itemschwierigkeit variiert zwischen  $p=.11$  und  $p=.94$ .

Differenzierte Ergebnisse zu einzelnen Merkmalen der Bauklötz-Anordnungen und ihr deren Zusammenhang mit der Leistung der Kinder werden präsentiert und zur Diskussion gestellt.

Peter A. Edelsbrunner (ETH Zürich, Lehr- und Lernforschung), Anne Deiglmayr (ETH Zürich, Lehr- und Lernforschung), Lennart Schalk (ETH Zürich, Lehr- und Lernforschung)

## ***Das Verständnis experimenteller Designs im Kindesalter: Analyse und Überarbeitung eines Erhebungsinstruments***

Lernen durch Experimentieren wurde in den letzten Jahrzehnten als effektives Instrument akzeptiert, um SchülerInnen reichere und stärker wissenschaftlich basierte Lernerfahrung zu ermöglichen (Kuhn et al., 2000). Als wichtige Grundlage hierfür gilt das sogenannte wissenschaftliche Denken, zu dem auch die Kompetenz gezählt wird, Experimente valide zu designen (Zimmerman, 2007). Im Mittelpunkt von Untersuchungen standen bisher vorrangig Interviewmethoden und Beobachtungen von Kindern und Jugendlichen während unterschiedlicher Phasen des Experimentierens (z.B. Chen & Klahr, 1999; Schauble, Klopfer, & Raghavan, 1991). Die vorliegende Studie hatte das Ziel, ein Multiple-Choice-Instrument zur Messung des Verständnisses experimenteller Designs im Kindesalter zu entwickeln, welches auch in großen Stichproben ökonomisch einsetzbar ist.

Das Instrument soll zwei zentrale Aspekte des Verständnisses experimenteller Designs messen: Variablenkontrolle und konklusive Hypothesentestung. Etwa ab dem 9. Lebensjahr beginnen Kinder ein Verständnis dieser Aspekte zu entwickeln (Zimmerman, 2007; Kuhn, 2010), wobei in Bezug auf Variablenkontrolle vermutet wird, dass die selbstständige Planung von kontrollierten Experimenten schwieriger ist als die Bewertung dieser Eigenschaft an vorgegebenen Versuchsplänen (Bullock & Ziegler, 1999).

Für den Fragebogen zum Verständnis experimenteller Designs (FreD) wurden 14 Multiple-Choice-Aufgaben mit jeweils zwei bis vier Antwortmöglichkeiten entwickelt, von denen zehn zur Erfassung des Aspekts der Variablenkontrolle und vier zur Erfassung der konklusiven Hypothesentestung dienen. Sprachlich und inhaltlich wurde der Fragebogen auf die Messung des Verständnisses experimenteller Designs von Zehn- bis Zwölfjährigen ausgelegt.

Der FreD wurde im Rahmen der Schweizer MINT Studie (siehe [www.educ.ethz.ch/mint/natwiss](http://www.educ.ethz.ch/mint/natwiss)) bisher 1787 SchülerInnen aus Primarschulklassen (1. bis 6. Klassenstufe) der deutschsprachigen Schweiz vorgegeben (50.4% Mädchen, Alter=8.68, SD=1.33, 50% 3.-4. Schulstufe). Zur psychometrischen Überprüfung der Fragen wurde an einer zufälligen Stichprobenhälfte eine exploratorische maximum likelihood-Faktorenanalyse unter Verwendung von numerischer Integration für kategoriale Daten durchgeführt. Eine Analyse des Screeplots zeigte einen starken Erstfaktor (Eigenwert: 4.21), einen Knick nach dem ersten Faktor sowie insgesamt 3 Faktoren mit Eigenwerten  $> 1$ . Zur statistischen Extrahierung der Faktorenanzahl wurde eine minimum average partial-Analyse durchgeführt, die für eine Ein-Faktoren-Lösung sprach. Inhaltlich schien jedoch eine Drei-Faktoren-Lösung mit den Faktoren Planung kontrollierter Experimente, Bewertung von Variablenkontrolle in vorhandenen Versuchsplänen und Überprüfung konklusiven Testens plausibel (unter Ausschluss von vier nicht zuordenbaren Items). Diese zwei statistisch und theoretisch möglichen Faktoren-Lösungen wurden an der zweiten Stichprobenhälfte konfirmatorisch verglichen. Es zeigte sich eine adäquate absolute Passung sowie eine bessere relative Passung der Drei-Faktoren-Lösung gegenüber der Ein-Faktoren-Lösung. Die erhaltenen drei Faktoren beschreiben die Planung kontrollierter Experimente (5 Items; interne Konsistenz:  $\lambda_4=.61$ ), die Bewertung von Variablenkontrolle in vorhandenen Versuchsplänen (3 Items; interne Konsistenz:  $\lambda_4=.49$ ) und die Überprüfung konklusiven Testens (2 Items; interne Konsistenz:  $\lambda_4=.44$ ). In Anbetracht der geringen Anzahl an Items sind die internen Konsistenzen vertretbar.

Eine genauere Inspektionen der Faktorwertverteilungen sowie Berechnungen von Mischverteilungsmodellen zeigten eine Klasse von etwa einem Drittel der Kinder, die volles Verständnis aller drei Teilfähigkeiten etwa ab dem 9. Lebensjahr entwickeln. Um Deckeneffekte durch diese individuelle Entwicklung zu verhindern und um die Faktoren 2 und 3 breiter abzudecken, werden derzeit in einem Revisionsprozess neue, schwierigere Aufgaben entwickelt, die die vier unzureichend validen Aufgaben der Erstversion ersetzen sollen.

Theoretisch deuten die berichteten Ergebnisse darauf hin, dass die Planung kontrollierter Experimente und die Bewertung von Variablenkontrolle in vorhandenen Versuchsplänen nicht nur wie vermutet unterschiedlich schwierige Facetten der Variablenkontrolle darstellen (Bullock & Ziegler, 1999), sondern dass es sich hierbei sogar um distinkte Fähigkeiten im Kindesalter handelt. Zudem bestätigen die Ergebnisse, dass bei Kindern ab etwa dem 9. Lebensjahr eine starke Selbstentwicklung in ihrem Verständnis experimenteller Designs stattfindet. Der FreD zeigt derzeit noch eine unzureichende interne Konsistenz, dem durch die Ersetzung der vier unpassenden Aufgaben entgegengewirkt werden soll. Bereits jetzt aber weist der Test eine gute Konstruktvalidität auf und lässt sich durch das Multiple-Choice-Format ökonomisch einsetzen, um das Verständnis experimenteller Designs in Kindern zu untersuchen.

## **Erfassung von Kompetenz im Bereich Erkenntnisgewinnung zu Beginn der Jahrgangsstufe 5**

### Theoretischer Hintergrund und Ziel der Studie

Zu einer naturwissenschaftlichen Grundbildung gehören nach Bybee (1997) Fachwissen und Prozesswissen. Beide Komponenten werden z. B. in IGLU-E (2003) erfasst, wobei für das Prozesswissen (Bereich Erkenntnisgewinnung) ausschließlich Aufgaben vorliegen, die sich post hoc dem höchsten Kompetenzniveau zuordnen lassen (Bos et al., 2003). Aus den Daten geht daher nicht hervor, über welche basalen Kenntnisse Schülerinnen und Schüler in diesem Bereich verfügen. Untersuchungen, in denen speziell das Prozesswissen von Kindern im Grundschulalter bzw. zu Beginn der Sekundarstufe I durch Tests oder Interviews erfasst wird, legen meist nur einen sehr kleinen Aufgabenpool zugrunde (z. B. Bullock & Ziegler, 1999; Schrempf & Sodian, 1999). Eine weitere Differenzierung des Konstrukts „Erkenntnisgewinnung“ ist auf Basis eines solch reduzierten Aufgabenpools schwierig. Deshalb wurde ein schriftlicher Test (NAW-II Test) zur Erfassung des Prozesswissens bei jüngeren Kindern (Jahrgangsstufe 5) auf Basis eines Kompetenzmodells entwickelt und evaluiert, mit dem die Schwierigkeit von Aufgaben im Bereich Prozesswissen systematisch variiert werden kann und zwar in Bezug auf die Dimensionen Aufgabenkomplexität und Grad der Hilfestellung bei der Aufgabenbearbeitung. Das Modell ist eine Adaption des ESNaS-Modells zur Evaluation der Bildungsstandards in den Naturwissenschaften (Walpuski et al., 2010).

### Forschungsfragen

1. Erfasst das Testinstrument das Prozesswissen der Probandinnen und Probanden?
2. Kann der hierarchische Aufbau der Modelldimensionen Komplexität und Grad der Hilfestellung empirisch bestätigt werden?

### Methode

Für jede Zelle des Kompetenzmodells ist ein Aufgabenpool entwickelt worden; 140 Aufgaben insgesamt. Alle Aufgaben liegen im Multiple-Choice Format (1 Attraktor, 3 Distraktoren) vor. Die Testaufgaben sind per Mult-Matrix-Design in 18 verschiedene Testhefte mit 18 Aufgaben pro Heft verteilt worden. Gemäß der Empfehlung von Rost (2004) ist jede Testaufgabe von mindestens 100 Schülerinnen und Schülern (Hauptschüler: N=613, Gymnasiasten: N=521) bearbeitet worden. Zu Validierungszwecken sind zusätzlich Lesegeschwindigkeit/ Leseverständis (ELFE; Lenhard & Schneider, 2006) und kognitive Fähigkeiten (KFT, Heller & Perleth, 2000) erfasst worden.

### Ergebnisse

Die Daten aus dem Test sind zunächst einer DIF Analyse unterzogen worden. Die Analyse wurde in ConQuest durchgeführt (Wu, Adams, & Wilson, 1998) mit dem Ziel, alle Testaufgaben von weiteren Analysen auszuschließen, die schulformspezifischen DIF erzeugen. Im nächsten Schritt sind die verbleibenden Aufgaben hinsichtlich ihrer Passung zum eindimensionalen Raschmodell hin überprüft worden. Insgesamt 123 Testaufgaben (von 140) entsprechen den Fit-Kriterien ( $0,75 > MNSQ > 1,2$ ;  $-2,0 > T < 2,0$ ). Die Reliabilität des Testinstruments (EAP/PV reliability: .76) sowie die Varianz (1,26) können als zufriedenstellend angesehen werden. Mit Blick auf die Validität des Testinstruments ist zunächst das Antwortverhalten der Schülerinnen und Schüler untersucht worden. Hierbei konnte gezeigt werden, dass die Probandinnen und Probanden, die die Aufgaben nicht korrekt lösen, insbesondere die Distraktoren wählen, die Misskonzepte im Bereich der Erkenntnisgewinnung widerspiegeln (z. B. fehlende Variablenkontrolle). Dies spricht dafür, dass das Konstrukt Prozesswissen mit den Testaufgaben erfasst wird. Im zweiten Schritt wurden bivariate Korrelationen zwischen dem NAW-II Test und den anderen Testinstrumenten (ELFE, KFT) berechnet. Die Ergebnisse zeigen, dass mit den Instrumenten verschiedene Konstrukte gemessen werden (Mannel, 2011). Um die hierarchische Struktur des Kompetenzmodells zu überprüfen, ist der Einfluss der Modellfaktoren Komplexität und Grad der Hilfestellung auf die Aufgabenschwierigkeit untersucht worden. Die hierzu durchgeführte ANOVA konnte statistisch signifikante Haupteffekte für beide Faktoren aufzeigen (Komplexität :  $F(2, 115) = 31,304$ ,  $p = .003$ ;  $\eta^2 = .353$ ; Grad der Hilfestellung :  $F(2, 115) = 6,224$ ,  $p \leq .001$ ;  $\eta^2 = .098$ ). Die Berechnung von Kontrasten hat die Abfolge der Stufen für den Faktor Komplexität bestätigt. In Bezug auf den Faktor Grad der Hilfestellung zeigen sich signifikante Unterschiede in der Aufgabenschwierigkeit zwischen den Extrembereichen (Mannel, 2011). Die Befunde deuten darauf hin, dass das Konstrukt Erkenntnisgewinnung mit den Aufgaben erfasst und die hierarchische Struktur des Kompetenzmodells bestätigt werden.

## **Bereits wenig Vorwissen bewirkt ein Unterschied, wie sehr von einer Algebraeinführung profitiert wird**

### Theoretischer Hintergrund

Vorwissen in einem Gebiet spielt eine wichtige Rolle, wenn neue Konzepte gelernt werden müssen (Schneider, Grabner, & Paetsch, 2009). Es konnte oft gezeigt werden, dass domänenspezifisches Wissen Lerngewinne besser vorhersagen kann als die Intelligenz (Grabner, Neubauer, & Stern, 2006; Stern, 2001). Nach dem Schweizer Mathematiklehrplan wird Algebra erst in der Oberstufe eingeführt, d.h. Schüler in der Primarschule haben noch keinen formalen Algebraunterricht bekommen. Für die Testung eines Algebralernprogramms wurden daher Sechstklässler ausgewählt, die sich im 2. Halbjahr und somit kurz vor ihrem Wechsel in die Oberstufe befanden. Schüler in diesem Alter besitzen das notwendige Arithmetikwissen welches als Voraussetzung gilt, um Algebra zu lernen. Wir untersuchten daher in dieser Studie, ob die Schüler schon vor dem Beginn des Algebra-Unterrichts intuitives Verständnis von Algebra zeigten. Des Weiteren war von Interesse, ob sich solch ein mögliches Vorwissen auf die Lernergebnisse auswirkten würde.

### Methode

253 Sechstklässler (Alter = 12.3 Jahre, SD = 0.5) bekamen eine Einführung in die Algebra mit einfachen Termumformungen. Jeder Schüler nahm an 2-stündigen Trainingssessionen an vier aufeinanderfolgenden Tagen teil. Dabei bearbeiteten sie ein Selbstlernprogramm mit insgesamt 9 Lerneinheiten.

Das Algebra Vorwissen wurde mit 8 einfachen Additions- und Multiplikationsaufgaben erhoben (" $a+a+a+a=$ ", " $5+a+a+5+a=$ ", " $c \cdot c \cdot c=$ ", " $2 \cdot 2 \cdot z \cdot 2 \cdot z=$ ", " $7b+7b=$ ", " $7b \cdot 7b=$ ", " $ab \cdot 4ab=$ ", " $xy+xy+xy+xy=$ "). Nach dem Training nahmen alle Teilnehmer an drei Posttests teil (1 Tag, 1 Woche und 10 Wochen später), bei denen wiederum ihr Algebrawissen erhoben wurde (58 Aufgaben, z.B. " $a2 \cdot a \cdot ay \cdot 4a=$ ", " $5ab+b+3b+2ab+2b=$ "). An den Daten des Pretests sollte festgestellt werden, ob die Leistung in diesem mit unterschiedlichen Lernmustern einhergeht.

### Resultate

Wir führten eine 3 x 3 ANOVA mit Messwiederholung durch mit den Faktoren Algebra vorwissen (keines: 0 Punkte; minimales: 1-2 Punkte; moderates: 3-4 Punkte) und Zeitpunkt des Posttests (T1: 1 Tag, T2: 1 Woche, T3: 10 Wochen).

Das Algebra Vorwissen war im Durchschnitt sehr tief ( $M = 0.98$ ,  $SD = 1.12$ ); 12 % der Schüler lösten 3 Aufgaben, 2 % 4 Aufgaben (Tabelle 1) und sogar nur 7 Schüler lösten eine der Multiplikation ( $c \cdot c \cdot c=$ ) richtig. Wie erwartet zeigte sich ein Effekt des Vorwissens auf das Algebrawissen ( $p = .027$ ,  $\eta^2 = .024$ ; siehe Abbildung 1). Posthoc-Tests zeigten einen signifikanten Unterschied zwischen wenig Vorwissen und keinem Vorwissen,  $p = .017$ , sowie minimalem Vorwissen,  $p = .041$ .

Genauere Betrachtung der falschen Lösungen zeigte 2 Muster: Im ersten Muster wird die Variable in irgendeiner Form mit in die Lösung einbezogen (für „ $5+a+a+5+a=$ “ z.B. 10a, 13a, 10 3a), im zweiten Muster werden die Variablen zu einer Zahl umgerechnet („ $5+a+a+5+a=$ “ z.B. 13, 25) oder zu einem neuen Buchstaben im Alphabet („ $5+a+a+5+a=$ “ z.B. m, 10c).

Schüler, welche im Pretest mehr Aufgaben lösen konnten, produzierten signifikant mehr falsche Lösungen nach dem ersten Muster. Hingegen gab es im Produzieren falscher Muster keinen Hinweis auf Intelligenzunterschiede (erhoben per 2 Skalen Logisches Denken aus dem LPS (Horn, 1983)) oder Unterschiede zwischen denjenigen, die besser oder schlechter im Algebrawissen abschnitten.

### Diskussion

Die Vorgabe eines Algebra tests am Ende der sechsten Schulstufe zeigte, dass kaum intuitives Vorwissen vorhanden ist, v.a. nicht zu Multiplikationsaufgaben, welche nicht von Masseinheiten ableitgeleitet werden können. Trotzdem zeigte sich ein Einfluss dieses minimalen Vorwissens auf die Vorhersage der weiteren Lerngewinne.

Die Subanalyse zeigt, dass zwar das Algebra vorwissen einen Einfluss auf das spätere Algebra wissen hat, dass sich jedoch aus der Art der Fehler, welche die Schüler im Pretest produzieren, keine Tendenz für den Lernerfolg zeigt. Schüler, welche die einfachen Additionsaufgaben im Algebra pretest lösen konnten, verwendeten die Variablen jedoch eher im Sinne von Masseinheiten, was zeigt, dass sie sich auf ihr Vorwissen bezüglich der Einheiten stützten. Andere Schüler hingegen zeigten die Tendenz, für die Variablen Zahlen einzusetzen, was ein Hinweis sein kann, dass sie sich eher auf ihr Arithmetikwissen stützten.

## Lernverläufe mathematischer Kompetenzen im 1. Schuljahr

Theoretischer Hintergrund: Mathematische Kompetenzen entwickeln sich bereits ab einem Alter von zwei Jahren und werden in den ersten Jahren formaler Erziehung stetig erweitert. Basale (Unterscheidung zwischen Mengen, Zahlen als Mengeneinheiten begreifen) sowie fortgeschrittene Vorläuferfertigkeiten (Zahlenmuster erkennen, Zahlen auf dem Zahlenstrahl verorten) sollten bereits vor Schuleintritt weitgehend ausgebildet sein, da sie die Grundlage für die weitere Entwicklung mathematischer Kompetenzen bilden. Rechenfertigkeiten wie Addition und Subtraktion schließlich bilden ein curriculares Ziel für die ersten Schuljahre. Wissen über Lernverläufe mathematischer Kompetenzen kann dabei helfen, individuelle Förderung besser zu planen und anzuwenden sowie Risikogruppen frühzeitig zu identifizieren und entsprechende Fördermaßnahmen zeitnah zu initiieren. Während im Bereich des Lesens kumulative Verläufe mit zunehmendem Auseinanderdriften der Leistungen starker und schwacher SchülerInnen und auch kompensatorische Mechanismen durch verschiedene Forschungsarbeiten gut beschrieben werden, sind entsprechende Untersuchungen bezüglich mathematischer Kompetenzen selten (z.B. Aunola, Leskinen, Lerkkanen & Nurmi, 2004; Geary, Hoard, Nugent & Bailey, 2012; Jordan, Kaplan, Locuniak & Ramineni, 2007; Jordan, Kaplan, Oláh & Locuniak, 2006). Einige dieser Studien finden hauptsächlich kumulative Trends. Es gibt aber auch Hinweise darauf, dass einige SchülerInnen eher mittelmäßig oder schwach beginnen und dann zunehmend aufholen. Wenig Klarheit herrscht dabei über die Anzahl und die tatsächlichen Verlaufsmuster unterschiedlicher latenter Gruppen hinsichtlich des Lernverlaufs mathematischer Kompetenzen, wobei oft a priori Zuordnungen von SchülerInnen zu unterschiedlich leistungsstarken Gruppen vorgenommen werden, anstatt die Anzahl und Struktur der möglichen Gruppen direkt aus den Daten abzuleiten (z.B. Morgan, Farkas & Wu, 2009). Unklarheiten bestehen auch bezüglich verschiedener Subkompetenzen, da die meisten Studien sich auf Gesamtscores konzentrieren. Zudem wurde zwar anhand längerer Messintervalle die Stabilität der Entwicklung über insgesamt mehrere Jahre hinweg dokumentiert, nicht aber die (teils temporeiche) Entwicklung innerhalb eines Schuljahres über kleinschrittige Messintervalle (z.B. monatlich) verfolgt.

Fragestellung: Welche Lernverläufe lassen sich für mathematische Subkompetenzen im ersten Schuljahr über kurze Messintervalle hinweg identifizieren? Gibt es Gruppen von SchülerInnen, die sich unterschiedlich entwickeln? Gibt es Hinweise auf die Stabilität der Zuordnung von SchülerInnen zu starken oder schwachen (sowohl bezüglich Ausgangswert als auch Wachstum) Klassen in den verschiedenen Subkompetenzen?

Methode: Zur Beantwortung dieser Fragen wurden die mathematischen Kompetenzen von 153 ErstklässlerInnen (46% weiblich, im Mittel 6,4 Jahre alt beim ersten Messzeitpunkt) mit einem onlinegestützten Messinstrument über sieben Messzeitpunkte untersucht. Erfasst wurden basale (Zahlendiskrimination, Symbolmengendiskrimination, Zahlenidentifizierung) und fortgeschrittene Vorläuferfertigkeiten (Zahlenreihen, Zahlenstrahl) sowie Rechenkompetenzen (Addition, Subtraktion, Gleichungen). Diese Subkompetenzen wurden außerdem zu einem Gesamtscore zusammengefasst. Vier parallele Testversionen A, B, C und D wurden konstruiert und in der Reihenfolge A-D, A-C mit einem mittleren Abstand von 3 Wochen innerhalb des ersten Schuljahres vorgegeben. Jeder einzelne Test benötigte eine Bearbeitungszeit von 10 (siebter Testzeitpunkt) bis 15 (erster Testzeitpunkt) Minuten. Parallelität und Reliabilität erwiesen sich in einer vorangegangenen Studie als gut. Die Auswertung erfolgte mittels Wachstumskurvenmodellen innerhalb einer Multilevel-Strukturgleichungsumgebung. Durch die Anwendung von Full Information Maximum Likelihood-Schätzungen wurden alle vorhandenen Datenpunkte genutzt.

Ergebnisse: Latent Growth Curve-Analysen (LGC; z.B. Bollen & Curran, 2006) zeigten für den Gesamtscore und auch für die drei einzelnen Subkompetenzen signifikant von Null verschiedene Startwerte und Wachstumsparameter. Latent Class Growth-Analysen (LCGA; z.B. Jung & Wickrama, 2008) ergaben für die basalen Vorläuferkompetenzen drei und für die fortgeschrittenen Vorläuferkompetenzen, Rechenkompetenzen und Gesamtscores fünf sinnvoll interpretierbare und bezüglich Fit-Indizes vertretbare latente Klassen mit jeweils sehr unterschiedlichen Verläufen. Dabei fanden sich neben starken und schwachen Gruppen auch stets abgrenzbare Klassen von SchülerInnen, die nach einem schwachen Start zunehmend zu den besseren Gruppen aufholen konnten. Die Zuordnung zu den latenten Klassen erwies sich über die Subkompetenzen hinweg als stabil. Die Ergebnisse zeigen, dass eine Statusmessung zu Beginn eines Schuljahres den möglichen unterschiedlichen Entwicklungspotenzialen der SchülerInnen auch bezüglich der genannten Subkompetenzen nicht gerecht wird. Deshalb sollten wiederholte Kompetenzmessungen mit relativ kurzen Messintervallen der reinen Statusdiagnostik möglichst vorgezogen werden.

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, 96(4), 699-713. doi:10.1037/0022-0663.96.4.699

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: a structural equation perspective*. New York, NY: Wiley.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*, 104(1), 206-223. doi:10.1037/a0025398

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research & Practice*, 22(1), 36-46. doi:10.1111/j.1540-5826.2007.00229.x

Jordan, N. C., Kaplan, D., Oláh, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77(1), 153-175. doi:10.2307/3696696

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302-317. doi:10.1111/j.1751-9004.2007.00054.x

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-Year Growth Trajectories of Kindergarten Children With Learning Difficulties in Mathematics. *Journal of Learning Disabilities*, 42(4), 306-321. doi:10.1177/0022219408331037





Susanne Koerber (Pädagogische Hochschule Freiburg), Christopher Osterhaus (Pädagogische Hochschule Freiburg), Knut Schwippert (Universität Hamburg), Beate Sodian (LMU München)

## **Kompetenzmodelle und intraindividuelle Entwicklungsverläufe des wissenschaftlichen Denkens in der Grundschule**

Im Rahmen des Kooperationsprojektes Science-P (z.B. Hardy et al., 2010; Koerber, Kropf, Mayer, Sodian, & Schwippert, 2011), in dem die Entwicklung der naturwissenschaftlichen Kompetenz im Grundschulalter untersucht wird, werden in diesem Vortrag Befunde einer längsschnittlichen Untersuchung zur Struktur und der Entwicklung des wissenschaftlichen Denkens von der dritten bis zur vierten Klassenstufe vorgestellt.

Wissenschaftliches Denken ist eine zentrale Komponente der naturwissenschaftlichen Kompetenz und umfasst das Verständnis und die Anwendung des wissenschaftlichen Erkenntnisprozesses, mit einem Fokus auf dem Hypothese-Evidenz Verständnis (z.B. Kuhn, 2010). Somit beinhaltet es grundlegende, zugleich jedoch sehr breite Fähigkeiten im Bereich des Verständnisses und der Anwendung der experimentellen Methode, der Interpretation von Daten unter Rückbezug auf eine Hypothese sowie ein metakognitives Verständnis von Wissenschaft und ihren Zielen (Nature of Science) und der Bedeutung von Rahmentheorien. Während frühere Studien (Koerber, Mayer, Osterhaus, Schwippert, & Sodian, submitted) bereits Kompetenzzuwächse im Grundschulalter zeigen konnten, ist bisher wenig über intraindividuelle Entwicklungsverläufe in dieser jungen Altersstufebekannt. Darüber hinaus gibt es bis dato für diese Altersstufe kein theoretisch fundiertes und zugleich empirisch überprüftes Kompetenzmodell, das die Entwicklung des wissenschaftlichen Denkens quantitativ und qualitativ beschreibt.

Basierend auf Befunden von Mayer, Sodian, Koerber und Schwippert (2014), die eine eindimensionale Struktur des wissenschaftlichen Denkens bei Viertklässlern gefunden haben, geht unser Modell von einer konzeptuellen Kohärenz des Konstrukts und somit einer eindimensionalen Struktur auch für jüngere Kinder ab der dritten Klasse aus. Wenig ist jedoch darüber bekannt, inwieweit sich qualitativ verschiedene Kompetenzlevels des wissenschaftlichen Denkens innerhalb einzelner Aufgaben unterscheiden und abbilden lassen. Bezugnehmend auf Befunde aus der Forschung zur Entwicklung des wissenschaftlichen Denkens und zum Konzeptwechsel in naturwissenschaftlichen Inhaltsbereichen (z.B. Carey, Evans, Honda, Jay & Unger, 1989; Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001) postulieren wir, dass sich das wissenschaftliche Denken ausgehend von naiven Vorstellungen über Zwischenvorstellungen hin zu wissenschaftlich adäquaten Vorstellungen entwickelt.

Zur Untersuchung dieser Fragestellungen wurde N=1559 Drittklässlern jeweils 21 paper-und-pencil Aufgaben zum wissenschaftlichen Denken präsentiert, sowie Tests zu Intelligenz, zum Leseverständnis und zum SES Status der Eltern. Erste Analysen zeigten für den ersten Messzeitpunkt eine gute Modellpassung. Es konnten die oben postulierten qualitativ unterschiedlichen Vorstellungen modelliert werden. Gemäß der Erwartung ließen sich interindividuelle Unterschiede zwischen den Kindern auf ihre Intelligenz, ihre sprachlichen Fähigkeiten sowie den Bildungshintergrund der Eltern zurückführen, jedoch nicht auf diese reduzieren.

Erste Ergebnisse der längsschnittlichen Betrachtung der Kompetenzentwicklung deuten an, dass Kompetenzzuwächse in allen Komponenten des wissenschaftlichen Denkens aufgedeckt werden konnten. Zudem werden in dem Vortrag Ergebnisse vorgestellt, die die intraindividuellen Entwicklungsverläufe näher beschreiben. Die Ergebnisse der Studie werden im Zusammenhang mit Kompetenzstrukturmodellen zum wissenschaftlichen Denken in höheren Altersbereichen und mit der Kompetenzmessung in der Grundschule diskutiert.

Bullock, M., Sodian, B., & Koerber, S. (2009). Doing experiments and understanding science: Development of scientific reasoning from childhood to adulthood. In W. Schneider, & M. Bullock (Eds.), Human development from early childhood to early adulthood. Findings from the Munich Longitudinal Study (pp. 173–197). Mahwah, NJ: Lawrence Erlbaum.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works. A study of junior high school students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11, 514–529.

Hardy, I., Kleickmann, T., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., Sodian, B., & Schwippert, K. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. In E. Klieme, D. Leutner, M. Kenk. Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. 56. Beiheft der Zeitschrift für Pädagogik (S. 115-125), Weinheim u.a.: Beltz.

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., & Sodian, B. (submitted). The development of early scientific reasoning. *Child Development*.

Koerber, S., Sodian, B., Kropf, N., Mayer, D., & Schwippert, K. (2011). Die Entwicklung des wissenschaftlichen Denkens im Grundschulalter: Theorieverständnis, Experimentierstrategien, Dateninterpretation. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*. 43(1), 16-21.

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of childhood cognitive development* (2nd ed., pp. 472–523). Oxford, UK: Blackwell.

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific Reasoning in Elementary School Children: Assessment and Relations with Cognitive Abilities. *Learning & Instruction*, 29, 43-55. doi:10.1016/j.learninstruc.2013.07.005

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, 15, 317-419.



2007 wurde am Institut für Qualitätsentwicklung im Bildungswesen (IQB) das fächerübergreifende Projekt „Evaluation der Standards in den Naturwissenschaften der Sekundarstufe 1“ (ESNaS) mit dem Ziel eingerichtet, eine gemeinsame Grundlage für die Aufgabenentwicklung in den naturwissenschaftlichen Fächern zu erarbeiten, auf deren Grundlage Testaufgaben zur Überprüfung des Erreichens der Bildungsstandards entwickelt werden können. Kompetenzen, die für den Erwerb einer naturwissenschaftlichen Grundbildung im Fach Chemie bedeutsam sind, werden in den nationalen Bildungsstandards (NBS) beschrieben. Es wird zwischen den vier Kompetenzbereichen Fachwissen, Erkenntnisgewinnung, Kommunikation & Bewertung unterschieden (KMK, 2005 a, b, c). Mit der Ausweisung des Kompetenzbereichs Erkenntnisgewinnung haben die deutschen Bildungsstandards Anschluss an internationale Standards gefunden. Insbesondere in angelsächsischen Curricula (England, Kanada, USA) ist naturwissenschaftliche Erkenntnisgewinnung traditionell stärker verankert als in deutschen Lehrplänen der Vergangenheit. Sie ist ein wesentlicher Teil des Konstrukts „Science“ (vgl. NAGB, 2004).

In den auf der Weinert'schen Kompetenzdefinition (Weinert, 2001) basierenden deutschen Bildungsstandards liegt der Schwerpunkt im Bereich Erkenntnisgewinnung zum einen auf der Förderung von Kompetenzen, die das Denken in und mit Modellen verlangen. Ein weiterer Schwerpunkt liegt auf der problemlösenden Prüfung naturwissenschaftlicher Hypothesen und dem Aufzeigen möglicher Auswirkungen neuer Erkenntnisse (KMK, 2005 a,b,c). Während für den Kompetenzbereich Fachwissen bereits durch die nationalen Bildungsstandards eine Unterteilung durch die Beschreibung von vier Basiskonzepten erfolgte, wurde für den Kompetenzbereich Erkenntnisgewinnung im Projekt ESNaS in Anlehnung an Mayer (2007) eine solche Unterteilung vorgenommen. Die drei, auf Basis theoretischer und empirischer Vorarbeiten definierten, Teilbereiche lauten daher Naturwissenschaftliche Untersuchungen, Naturwissenschaftliche Modellbildung und Wissenschaftstheoretische Reflexion. Die einzelnen Teilbereiche werden wiederum in einzelne Aspekte untergliedert. Für den ersten Teilbereich Naturwissenschaftliche Untersuchungen handelt es sich dabei um die Aspekte Fragestellung, Hypothese, Untersuchungsdesign und Datenauswertung. Der Teilbereich Naturwissenschaftliche Modellbildung wird in die Aspekte Funktionalität, Anwendung und Grenzen unterteilt, während der dritte Teilbereich Wissenschaftstheoretische Reflexion in die zwei Aspekte Eigenschaften und Entwicklung untergliedert wird (Walpuski et al., 2010; Wellnitz et al., 2012). Die in den NBS formulierten acht Standards zum Kompetenzbereich Erkenntnisgewinnung zeigen eine Schwerpunktsetzung in den Teilbereichen Naturwissenschaftliche Untersuchungen, und Naturwissenschaftliche Modellbildung. International (z. B. NRC, 1996) werden in der Regel alle Bereiche gleichwertig beschrieben. Ein internationaler Vergleich von Schülerkompetenzen im dritten Teilbereich, Wissenschaftstheoretische Reflexion, wurde von Neumann (2011) im Fach Physik durchgeführt und zeigte, trotz der Schwerpunktsetzung auf die anderen Teilbereiche, eine höhere Kompetenzausprägung auf Seiten der deutschen SuS. Untersuchungen in den Fächern Biologie und Chemie zur Dimensionalität zwischen einzelnen Aspekten im Teilbereich Naturwissenschaftliche Untersuchungen liefern bisher keine einheitlichen Ergebnisse (Hammann, Phan & Bayrhuber, 2007; Henke, 2006; Klos et al., 2008; Mayer, Grube & Möller, 2008; Wellnitz, 2012). Ziel dieser Studie ist, eine detaillierte Analyse der Schülerkompetenzen für das Fach Chemie im Hinblick auf die deutsche Schwerpunktsetzung im Bereich Erkenntnisgewinnung durchzuführen. Dazu sollen – in Ergänzung zu ESNaS – die Schülerkompetenzen in allen Kompetenzteilbereichen und Aspekten erhoben werden und die Tests auf Dimensionalität (Teilbereiche und Aspekte) geprüft werden. Die Forschungsfragen dazu lauten: FF1: Lassen sich die theoretisch hergeleiteten Kompetenzteilbereiche als einzelne Skalen empirisch trennen?; FF2: Lassen sich die theoretisch hergeleiteten Aspekte der Kompetenzteilbereiche als einzelne Skalen empirisch trennen?; FF3: Weisen Schülerinnen und Schüler in den einzelnen Kompetenzteilbereichen und Aspekten unterschiedliche Fähigkeiten auf? Damit jeder Aspekt von einer ausreichend großen Anzahl an Aufgaben abgebildet wird, müssen neue Aufgaben entwickelt werden. Diese Aufgaben wurden im Sommer 2012 in der Pilotstudie eingesetzt. Die Daten werden derzeit eingegeben und werden bis zur Tagung ausgewertet sein (Itemfit, Schwierigkeit, Reliabilität etc). Im Frühjahr 2013 findet die Hauptdatenerhebung mit ca. 800 Schülerinnen und Schülern statt. In der Hauptstudie werden die besten Aufgaben aus der Pilotierung durch den Einsatz von Ankeritems mit den ESNaS Daten (Normierungsstudie) in Verbindung gebracht. Hierbei werden außerdem die Kontrollvariablen Geschlecht, Schulnote, Interesse, sozialer Hintergrund, Lesefähigkeit und kognitive Fähigkeit erhoben.

Hammann, M., Phan, T. H. & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS- Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? In M. Prenzel, I. Gogolin & H.-H.

Henke, C. (2006). Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe – Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven. Berlin: Logos.

Klos, S., Henke, C., Kieren, C., Walpuski, M. & Sumfleth, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen – zwei verschiedene Kompetenzen. Zeitschrift für Pädagogik, 54(3), 304-321.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005a). Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Jahrgangsstufe 10). München: Luchterhand.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005b). Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss (Jahrgangsstufe 10). München: Luchterhand.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005c). Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Physik für den Mittleren Schulabschluss (Jahrgangsstufe 10). München: Luchterhand.

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger, & H. Vogt, Theorien in der biologiedidaktischen Forschung (S. 177-186). Berlin/Heidelberg: Springer.

Mayer, J., Grube, C. & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In U. Harms & A. Sandmann (Hrsg.), Lehr- und Lernforschung in der Biologiedidaktik. (Band 3). Ausbildung und Professionalisierung von Lehrkräften (S. 63-79). Innsbruck: StudienVerlag.

NAGB / National Assessment Governing Board (2004). Science Framework for the 2005 National Assessment of Educational Progress. U.S. Department of Education.

Neumann, I. (2011). Beyond Physics Content Knowledge. Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge. Berlin: Logos.

NRC / National Research Council (1996). National science education standards. Washington, DC: National Academy Press.

Walpuski, M., Kauertz, A., Kampa, N., Fischer, H. E., Mayer, J., Sumfleth, E. et al. (2010). ESNaS – Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I. In A. Gehrman, U. Hericks, & M. Lüders (Hrsg.), Bildungsstandards und Kompetenzmodelle – Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht (S. 171–184). Bad Heilbrunn: Julius Klinkhardt.

Weinert, F.E. (2001). Leistungsmessungen in Schulen. Weinheim: Beltz Verlag.

Wellnitz, N. (2012). Kompetenzstruktur und -niveaus von Methoden naturwissenschaftlicher Erkenntnisgewinnung. Berlin: Logos.

Wellnitz, N., Fischer, H. E., Kauertz, A., Mayer, J., Neumann, I., Pant, H. A., Sumfleth, E. & Walpuski, M. (2012). Evaluation der Bildungsstandards – eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung. Zeitschrift für Didaktik der Naturwissenschaften, 8,261-291.

Montag,  
03.03.2014,  
17:30 - 18:30,

Casino-Gebäude  
Raum 823



## **Analyse des Zusammenhangs zwischen chemischem Fachwissen und der Fähigkeit, Phänomene auf submikroskopischer Ebene zu erklären.**

Theoretischer Hintergrund: Studien konnten zeigen, dass Lernende verschiedenen Alters Schwierigkeiten bei der Verknüpfung der Repräsentationsebenen von chemischen Inhalten haben (submikroskopisch, makroskopisch, symbolisch; Johnstone, 1991) (z.B. Dori & Kaberman, 2012). Ein Beispiel hierfür stellt die Interpretation von Phänomenen auf Grundlage der Teilchenebene dar (z.B. Meijer, Bulte & Pilot, 2009). Hieraus resultieren häufig Fehlvorstellungen, die auf der Übertragung makroskopischer Eigenschaften auf submikroskopische Konzepte basieren (z.B. Barke, 2006; Busker, 2010). Die Verknüpfung der verschiedenen Repräsentationsebenen ist jedoch eine zentrale Grundlage für ein tieferes Verständnis chemischer Konzepte (z.B. Jaber & BouJaoude, 2012). Um diese Verknüpfungen herstellen zu können, muss ebenenbezogenes Wissen zunächst reproduziert werden. Die submikroskopische und die symbolische Ebene stehen dabei eng miteinander in Beziehung, da die nicht wahrnehmbaren submikroskopischen Konzepte oft durch symbolische Visualisierungen veranschaulicht und auf diese Weise reproduziert werden (z.B. Treagust, Chittleborough & Mamiala, 2003). Danach muss das Wissen auf den drei Repräsentationsebenen integriert und zu einem sinnhaften mentalen Modell verknüpft werden. Ziel dieses Projekts ist die Entwicklung eines Testinstruments, das die Fähigkeit zur Verknüpfung von Phänomenen und submikroskopischen Zusammenhängen (VPs) misst. Dieses Testinstrument wird dazu genutzt, mögliche Zusammenhänge zwischen dem Wissenszuwachs im Fach Chemie und der Fähigkeit, Phänomene und submikroskopische Zusammenhänge zu verknüpfen, aufzuklären.

Fragestellungen: In wie weit sind Erstsemester-Studierende im Fach Chemie im Modul „Allgemeine Chemie“ in der Lage, ihr Wissen über Phänomene und submikroskopische Zusammenhänge zu verknüpfen?

Welchen Einfluss hat die Fähigkeit, chemische Phänomene und submikroskopische Zusammenhänge zu verknüpfen, auf den Fachwissenszuwachs von Erstsemester-Studierenden des Fachs Chemie im Modul „Allgemeine Chemie“?

Welchen Einfluss hat das Fachwissen von Erstsemester-Studierenden des Fachs Chemie im Modul „Allgemeine Chemie“ auf den Zuwachs der Fähigkeit, chemische Phänomene und submikroskopische Zusammenhänge zu verknüpfen?

Methode: Im Rahmen dieses Projekts wurde ein Testinstrument entwickelt, das die Fähigkeit zur Verknüpfung von Phänomenen und submikroskopischen Zusammenhängen messen soll (VPs). Der Test besteht dazu aus drei Itemtypen: Einerseits Items, bei denen das Wissen auf makroskopischer (erster Itemtyp) oder auf submikroskopischer bzw. symbolischer Ebene (zweiter Itemtyp) reproduziert werden muss, und andererseits Items, bei denen das Wissen auf den verschiedenen Ebenen miteinander verknüpft werden muss (dritter Itemtyp). Zu einem chemischen Konzept wurde jeweils ein Item der drei Itemtypen entwickelt, sodass der Test insgesamt aus 30 Items zu zehn chemischen Konzepten besteht. Bei der Pilotstudie im Wintersemester 2012/13 wurden der neu entwickelte Test, sowie ein bereits vorhandener Fachwissenstest (Freyer, in Vorb.) jeweils zu Beginn und zum Ende des Semesters eingesetzt. Hierbei konnten 61 Erstsemesterstudierende des Lehramts Chemie in einer Übung zum Modul „Allgemeine Chemie“ an der Universität Duisburg-Essen getestet werden. Ergebnisse: Die Kennwerte für den neu entwickelten VPs-Test lagen bei der Pilotstudie im Wintersemester 2012/13 im zufriedenstellenden Bereich ( $N=61$ ). Für die weitere Auswertung wurden, mit Hilfe des Cross-Lagged-Panel-Designs, die partiellen Rangkorrelationen der einzelnen Variablen berechnet. Es zeigte sich, dass es einen signifikanten Zusammenhang zwischen VPs-Fähigkeit zu Beginn des Semesters und Fachwissen am Ende des Semesters gibt ( $r_s = .49$ ;  $p < .001$ ), aber keinen signifikanten Zusammenhang zwischen dem Fachwissen zu Beginn des Semesters und der VPs-Fähigkeit am Ende des Semesters gibt ( $r_s = .21$ ;  $p = .103$ ), jedoch war der Unterschied dieser Korrelationen nur knapp nicht signifikant ( $p = .086$ ).

Barke, H.-D. (2006). Chemedidaktik: Diagnose und Korrektur von Schülervorstellungen. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Busker, M. (2010). Entwicklung einer adressatenorientierten Übungskonzeption im Übergang Schule-Universität auf Basis empirischer Analysen von Studieneingangsvoraussetzungen. Tönning, Lübeck, Marburg: Der Andere Verlag.

Dori, Y. J. & Kaberman, Z. (2012). Assessing high school chemistry students' modeling sub-skills in a computerized molecular modeling learning environment. *Instructional Science*, 40 (1), 69-91.

Freyer, K. (in Vorb.). Schwierigkeiten Studierender des 1. Semesters im Fach Chemie, Universität Duisburg-Essen. Essen

Jaber, L. Z. & BouJaoude, S. (2012). A Macro-Micro-Symbolic Teaching to Promote Relational Understanding of Chemical Reactions. *International Journal of Science Education*, 34 (7), 973-998.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning* (7), 75-83.

Meijer, M. R., Bulte, A. M. & Pilot, A. (2009). Structure-Property Relations Between Macro and Micro Representations: Relevant Meso-levels in Authentic Tasks. In J. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education. Models and Modeling in Science Education* (S. 195-213). Dordrecht: Springer-Verlag.

Treagust, D., Chittleborough, G. & Mamiala, T. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, 25 (11), 1353-1368.



## Ein Basistraining für erfolgreiche Nachhilfe

Theoretischer Hintergrund: Der Nachhilfemarkt boomt - eine unstrittige Tatsache. Dafür verantwortlich ist u.a. eine Unzufriedenheit mit der derzeitigen schulischen Situation (Rudolph, 2002). Auf der anderen Seite wird immer wieder Kritik an den fachlichen und pädagogischen Fähigkeiten der Nachhilfelehrer geäußert: „So ist zum Eröffnen eines Nachhilfeeinstituts lediglich ein Gewerbeschein notwendig, es ist jedoch keinerlei Überprüfung der pädagogischen Inhalte und Qualifikationen erforderlich“ (Dohmen, Erbes, Fuchs und Günzel, 2008, S. 70).

Dabei zeigt ein Blick in der Literatur zur individuellen Förderung, dass über strukturierte Programme „Novizen“ durchaus erfolgreich sein können: Elbaum, Vaughn, Hughes und Moody (2000) legen eine Metaanalyse zur Effektivität von Einzelförderung für Grundschul Kinder mit Lernschwierigkeiten im Lesen durch erwachsene Förderlehrkräfte vor. Dabei steigt die Effektivität der Förderung durch ehrenamtliche Helfer deutlich an, wenn diese im Vorfeld der Förderung ein Training erhalten. In die gleiche Richtung weisen die Forschungsbefunde der Autorengruppen Ritter, Barnett, Denny und Albin (2009) sowie Elliott, Arthurs und Williams (2000), dass nichtprofessionelle Förderlehrer unter der Voraussetzung erfolgreich sind, dass sie Hilfen der Strukturierung erhalten.

Empirisch konnte belegt werden, dass es bei Nachhilfe vor allem auf den Einsatz von stützenden Maßnahmen ankommt, dass sich der Nachhilfelehrer individuell auf seinen Nachhilfeschüler einlassen können muss (Streber, Haag & Götz, 2011). Dabei konnten fünf Wirkfaktoren identifiziert werden: Vorwissen sichern, Selbstkontrolliertes Lernen, Time on task, Lernstrategien, Individuelle Bezugsnorm. Gemeinsam in einem Team mit einer Lehrkraft, einem Psychologen und den beiden Autoren („Die Perspektiven verbinden“) wurde ein Praxisleitfaden für Nachhilfelehrer erarbeitet, in den diese Wirkfaktoren eingearbeitet wurden (Kessel & Haag, 2011).

Fragestellung: Hier geht es um die Forschungsfrage, inwieweit sich die Bearbeitung des Praxisleitfadens positiv auf Schüler- und Lehrerseite auswirkt, konkret:

1. Schätzen die Nachhilfelehrer ihren Unterricht nach erfolgreich durchgearbeiteten Leitfaden positiver ein, was die für erfolgreiche Nachhilfe zentralen Wirkfaktoren betrifft?
2. Schätzen die Nachhilfeschüler den Unterricht ihrer Nachhilfelehrer nach erfolgreich absolviertem Basistraining positiver ein, was die für erfolgreiche Nachhilfe zentralen Wirkfaktoren betrifft?

Methode: Stichprob: Von einer Nachhilfeeinrichtung konnten an fünf Standorten insgesamt 33 Nachhilfelehrkräfte gewonnen werden, die den Leitfaden (Basistraining) bearbeiteten. Außerdem nahmen an der Studie 135 NachhilfeschülerInnen (Sekundarstufe) teil, die Nachhilfeunterricht in Kleingruppen zwischen drei und fünf Teilnehmern erhielten. Bei der Stichprobenauswahl dieser Studie handelt es sich, wie bei vielen Projekten im Bereich der Schul- und Unterrichtsforschung, somit nicht um eine Zufalls-, sondern um eine Gelegenheitsstichprobe (Rost, 2007, S. 91), ein sog. convenience sample.

Erhebungsinstrumente: Es wurden zwei Fragebogen eingesetzt, ein Lehrer- und ein Schülerfragebogen, in denen die fünf Wirkfaktoren abgefragt wurden. Zu jedem der fünf Wirkfaktoren wurden Items zu Skalen zusammengefasst.

Datenerhebung: Beide Gruppen erhielten einen ersten Fragebogen, bevor die Lehrkräfte mit ihrem Training begonnen hatten. Als Startpunkt wurde Mitte Oktober gewählt, ein Monat, in dem nach Schuljahresbeginn nach Angaben der Nachhilfeeinrichtungen mit vielen Neuanmeldungen zu rechnen ist. So wurden auch nur Neuanmeldungen in die Studie aufgenommen. Damit diese neuangemeldeten Nachhilfeschüler auch sinnvoll ihre Nachhilfelehrer einschätzen konnten, wurde der Fragebogen erst nach zwei Wochen Verweildauer in der Nachhilfeeinrichtung ausgegeben. Zur Bearbeitung des Praxisleitfadens hatten die Nachhilfelehrer gut zwei Wochen Zeit. Um auch sicherzustellen, dass alle Nachhilfelehrer den Praxisleitfaden durchgearbeitet haben, wurden sie alle an einen Nachhilfestandort eingeladen, an dem anhand eines Paper-Pencil Tests das Wissen abgeprüft wurde. Ca. drei Monate später im Januar erhielten beide Gruppen einen mit dem ersten identischen zweiten Fragebogen.

Ergebnisse: Insgesamt zeigen die Skalen ausreichende Reliabilitäten (Cronbachs alpha). Bei den Lehrerskalen ergibt eine multivariate Varianzanalyse mit Messwiederholung keinen signifikanten Effekt, wohl aber bei den Schülerskalen. Nachgeordnete t-Tests für abhängige Stichproben zeigen bei einzelnen Skalen durch das Training Verbesserungen von mittleren bzw. kleinen Effekten. Mit dem konzipierten Praxisleitfaden für Nachhilfelehrer liegen hier erste Evaluierungsergebnisse vor, die auf die Wirksamkeit des Trainings hindeuten.

Dohmen, D., Erbes, A., Fuchs, K. & Günzel, J. (2008). Was wissen wir über Nachhilfe? – Sachstand und Auswertung der Forschungsliteratur zu Angebot, Nachfrage und Wirkungen. Berlin: Forschungsinstitut für Bildungs- und Sozialökonomie.

Elbaum, B., Vaughn, S., Hughes, M. T. & Moody, S. W. (2000). How Effective Are One-to-One Tutoring Programs in Reading for Elementary Students at Risk for Reading Failure? A Meta-Analysis of the Intervention Research. *Journal of Educational Psychology*, 92 (4), 605-619.

Elliott, J., Arthurs, J. & Williams, R. (2000). Volunteer Support in the Primary Classroom: the long-term impact of one initiative upon children's reading performance. *British Educational Research Journal*, 26 (2), 227-244.

Ritter, G. W., Barnett, J. H., Denny, G. S. & Albin, G. R. (2009). The Effectiveness of Volunteer Tutoring Programs for Elementary and Middle School Students: A Meta-Analysis. *Review of Educational Research*, 79 (1), 3-38.

Rost, D. H. (2007). Interpretation und Bewertung pädagogisch-psychologischer Studien. 2., überarbeitete und erweiterte Auflage. Weinheim: Beltz.

Rudolph, M. (2002). Nachhilfe – gekaufte Bildung? Empirische Untersuchung zur Kritik der außerschulischen Lernbegleitung. Bad Heilbrunn: Klinkhardt.

Streber, D., Haag, L. & Götz, T. (2011). Erfolgreiche Nachhilfe – Kann das jeder oder bedarf es besonderer Qualifikationen? *Empirische Pädagogik*, 25, 342-357.

Streber, D., van Kessel, M. & Haag, L. (2014). Ein Basistraining für erfolgreiche Nachhilfelehrer. *Empirische Pädagogik*





Balasundaram Krisanthan (Goethe-Universität Frankfurt), Sabine Fabriz (Goethe-Universität Frankfurt), Christian Müller (Goethe-Universität Frankfurt), Tim Engartner (Goethe-Universität Frankfurt), Hanna Mach (Goethe-Universität Frankfurt), Gerhard Minnameier (Goethe-Universität Frankfurt), Holger Horz (Goethe-Universität Frankfurt), Rico Hermkes (Goethe-Universität Frankfurt)

## **Der Wandel von Einstellungen im Kontext sozialwissenschaftlicher Bildung am Beispiel der ökonomischen Kategorie „Gewinn“**

Bislang ist die Erfassung von Lernprozessen – insbesondere die des Wissensaufbaus und der Einstellungsentwicklung – in den sozialwissenschaftlichen Didaktiken weitgehend unberücksichtigt geblieben (vgl. weiterführend Kunter & Voss, 2011). Insofern soll das Forschungsprojekt einen Beitrag zu der Frage leisten, ob – und wenn ja, inwieweit – ökonomisches Wissen einerseits und Einstellungen bezüglich relevanter Kategorien der sozialwissenschaftlichen Bildung andererseits korrelieren. Dabei fokussiert das Projekt im Westlichen zwei übergeordnete Fragestellungen:

- (1.) Welchen Einfluss hat die Vermittlung unterschiedlich konnotierter Gewinnbegriffe (positiv vs. negativ) auf die Einstellungsentwicklung von Schülerinnen und Schülern?
- (2.) Haben beobachtbare Basisvariablen der Unterrichtsqualität sowie beobachtbare Aktivitäten der Schülerinnen und Schüler einen (signifikanten) Einfluss auf die Einstellungsentwicklung?

Da sich Kompetenzen nicht allein in kognitiven Aspekten der Wissensanwendung erschöpfen, sondern affektive Faktoren einschließen (Engartner, 2012; Seeber et al., 2012; Weinert, 2001; Beck, 1989), sollen diese im Rahmen des interdisziplinär angelegten Teilforschungsprojektes aus der Perspektive der Einstellungsforschung beleuchtet werden, um diese bedeutende Kompetenzfacette in den Fokus der sozialwissenschaftlichen Bildung zu rücken. Einstellungen werden hierbei als „gegenstandsbezogene, erfahrungsbedingte und systemabhängige Verhaltensdeterminanten“ (Roth, 1967, S. 43) verstanden, die durch das Individuum erlernt bzw. erworben werden (müssen). Insbesondere werden Aspekte der Lernerfahrung expliziert und akzentuiert. Die Betrachtung von Einstellungen, die sowohl als Lernvoraussetzungen wie auch als Bildungs- und Erziehungsziele fungieren, ist umso bedeutsamer, als nahezu alle sozialwissenschaftlichen (Rahmen-)Lehrpläne über die Fachkompetenz hinaus Entscheidungs-, Urteils- und Handlungskompetenzen fokussieren. So zielt etwa der Erwerb ökonomischer Urteilskompetenz auf die Fähigkeit der Lernenden, „sich zu wirtschaftlichen Problemstellungen im eigenen Lebensbereich (Konsum, Vorsorge, Arbeit, Kreditaufnahme) eine begründete Meinung bilden zu können“ und „hinter ökonomischen Situationen und Argumentationen stehende Interessen zu erkennen“ (Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen, 2004, S. 16). Zudem ist der unterrichtliche Rahmen auch deshalb von Bedeutung, weil neben den Einstellungsänderungen bei den Lernenden prozessbezogene Basisvariablen der Unterrichtsqualität (kognitive Aktivierung und konstruktive Unterstützung) videographisch erfasst werden. Hierbei ist von Interesse, ob – und wenn ja, inwiefern – ein Zusammenhang zwischen den beiden Untersuchungsaspekten besteht. Im Rahmen des Forschungsvorhabens werden Unterrichtseinheiten zum Themenschwerpunkt „Gewinn“ bzw. „Unternehmensziele“ aufgegriffen und an hessischen Gymnasien mit Lernenden der 8. und 9. Jahrgangsstufe durchgeführt. Bei den Unterrichtseinheiten handelt es sich um im Internet frei zugängliche Angebote der Initiative Neue Soziale Marktwirtschaft (positiv konnotierter Gewinnbegriff) auf der einen und des globalisierungskritischen Netzwerks attac (negative Konnotation) auf der anderen Seite. Beide Unterrichtsmaterialsets sind für insgesamt 180 Minuten konzipiert und gescrriptet, um das Lehrerverhalten weitestgehend zu normieren und diesbezügliche Abweichungen zwischen den einzelnen Lerngruppen zu vermeiden. Hinsichtlich der Sozialform werden die Unterrichtsmaterialien dergestalt adaptiert, dass sie kooperativ und diskursiv angelegt sind, um u. a. die von den Schülerinnen und Schülern vorgetragenen Deutungs- und Argumentationsmuster videografisch erfassen zu können. Zur Erfassung der Einstellungen der Schülerinnen und Schüler wird im Vorfeld der Hauptuntersuchung jeweils vor und nach den Unterrichtseinheiten ein Fragebogen mit vierstufiger Likert-Skalierung entwickelt und erprobt. Neben den neu konzipierten und standardisierten Items werden etablierte Vignetten aus der englischsprachigen wirtschaftspsychologischen Forschung adaptiert (vgl. dazu u. a. Kahneman et al., 1986). Ausgehend von der ersten Erhebungsphase der Hauptstudie sollen im Rahmen der GEBF-Tagung erste Ergebnisse der Fragebogenerhebungen sowie der Ratings des Videomaterials von zwei bis drei Klassen präsentiert werden (N ≥ 45). Zugleich greift die Studie die aktuelle Kontroverse um die Verwendung von Unterrichtsmaterialien privater Content Anbieter wie „Wirtschaft und Schule“, „Handelsblatt macht Schule“, „My Finance Coach“, „hoch-im-kurs“ etc. in den sozialwissenschaftlichen Unterrichtsfächern auf (vgl. hierzu u.a. Hedtke, 2012; Lehmann, 2012; Schmidt, 2012; Schönwitz, 2012), so dass Rückschlüsse auf die curriculare (Weiter-)Entwicklung der ökonomischen Bildung ermöglicht werden (sollen).

Beck, K. (1989). „Ökonomische Bildung“ - Zur Anatomie eines wirtschaftspädagogischen Begriffs. Zeitschrift für Berufs- und Wirtschaftspädagogik, 85(7), S. 570-596.

Engartner, T. (2012). Die Bedeutung der Einstellungsforschung am Beispiel einer explorativen Studie zur Wettbewerbs- und Fairnessaffinität von Schüler(inne)n. In T. Retzmann (Hrsg.), Ökonomische Allgemeinbildung in der Sekundarstufe II (S. 113-127). Schwalbach/Ts.: Wochenschau Verlag.

Hedtke, R. (2012). Die Wirtschaft in der Schule. Agendasetting, Akteure, Aktivitäten. Working Paper Nr. 3. Bielefeld: Fakultät für Soziologie der Universität Bielefeld. Zugriff am 28. Oktober 2013. [http://www.uni-bielefeld.de/soz/ag/hedtke/pdf/Hedtke\\_Working-Paper\\_3.pdf](http://www.uni-bielefeld.de/soz/ag/hedtke/pdf/Hedtke_Working-Paper_3.pdf).

Kahneman, D., Knetsch, J. L., Thaler, R. H. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. The American Economic Review, 76(4), S. 728-741

Kunter, M. & Voss Thamar (2011). Das Modell der Unterrichtsqualität in COACTIV. Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV (S. 85-113). Münster: Waxmann.

Lehmann, A. (2012). Experten aus der Arbeitswelt im Unterricht. die tageszeitung vom 30.03.2012, S. 3.

Ministerium für Schule Jugend und Kinder Landes Nordrhein-Westfalen (2004). Rahmenvorgabe für die ökonomische Bildung in der Sekundarstufe I. Düsseldorf.

Roth, E. (1967). Einstellung als Determination individuellen Verhaltens. Göttingen: Hogrefe.

Schmidt, M. (2012). Tafeldienst der Spitzenberater. ftd.de vom 15.04.2012 Zugriff am 27. Dezember 2012. [http://www.ftd.de/wissen/bildung-tafeldienst-der-spitzenberater/7\\_0022288.html](http://www.ftd.de/wissen/bildung-tafeldienst-der-spitzenberater/7_0022288.html).

Schönwitz, D. (2012). Werbung im Klassenzimmer. Die Zeit, Nr. 17 vom 19.4.2012, S. 35.

Seeber, G., Retzmann T., Remmele, B. & Jongbloed, H.-C. (2012). Bildungsstandards der ökonomischen Allgemeinbildung. Kompetenzmodell, Aufgaben, Handlungsempfehlungen. Schwalbach/Ts.: Wochenschau-Verlag.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In ders. (Hrsg.), Leistungsmessungen in Schulen (S. 17-31). Weinheim/Basel: Beltz.





## **Aufklärung zum Thema Organspende – Kann man Schüler überzeugen, ihren Willen schriftlich zu dokumentieren?**

Der Organmangel (DSO, 2013) ist ein gesellschaftliches Problem, dem durch Aufklärungsarbeit und die Einführung der neuen Organspenderegelung (Entscheidungslösung; BMG, 2012) begegnet wird. Seit November 2012 erhält jeder deutsche Bürger ab 16 Jahren regelmäßig Informationsmaterial von seiner Krankenkasse und wird aufgefordert, eine persönliche Entscheidung auf dem enthaltenen Organspendeausweis zu dokumentieren, um Angehörige und medizinisches Personal zu entlasten (ibd.). Dadurch wird das Thema zu einem alltagsrelevanten Thema für Schüler der Sekundarstufe 2, das Bewertungskompetenz (KMK, 2004) erfordert. Die Entscheidungslösung gibt potenziellen Spendern die Möglichkeit, selbstbestimmt zu entscheiden und so aus reiner Solidarität zu spenden (BÄK, 2011; Vollmann, 2012). Der Organspendeausweis bietet Optionen, einer Organentnahme nach dem Tod zuzustimmen oder sie abzulehnen. Es war trotzdem ebenfalls erklärtes Ziel der Einführung der Entscheidungsregelung, mehr potenzielle Organspender zu gewinnen (BMG, 2012), denn repräsentative Befragungen hatten mehrfach eine starke Diskrepanz zwischen Einstellung und Verhalten ergeben (bspw. Watzke & Stander, 2010). Zudem zeigte sich, dass involvierte und gut informierte Personen häufiger einen Organspendeausweis ausgefüllt hatten (ibd.). Die BZgA publiziert Aufklärungsmaterial, das Organspende gemäß einer normativen Setzung (Vollmann, 2012) nicht ethisch neutral propagiert. Neben Broschüren, Webseiten und FAQs stellt die BZgA auch einen Unterrichtsfilm zur Verfügung, der besonders Jugendliche ansprechen und in das Thema involvieren soll.

Das Ziel der vorliegenden Studien war es, die Wirkung von Aufklärungsinterventionen mit Organspendeausweisen zur Mitnahme und von der BZgA bereitgestellten Informationen zu evaluieren. Es wurde hierbei die Wirkung einer Unterrichtsstunde mit dem Unterrichtsfilm im Vergleich mit einer neutral gestalteten PowerPoint-Präsentation mit demselben Inhalt verglichen. Es wurde untersucht, ob die Interventionen die Diskrepanz zwischen Einstellung und Verhalten minimieren oder (zusätzlich) eine Einstellungsänderung hervorrufen. In zwei Studien, Studie 1 vor und Studie 2 nach Einführung der Entscheidungslösung, wurden 188 (Durchschnittsalter 16,41 Jahre, 64,9% Mädchen) bzw. 103 (Durchschnittsalter 16,40 Jahre, 62,1% Mädchen) Schüler mit dem Film oder der PowerPoint-Präsentation unterrichtet. Wissen, Involviertheit, Einstellung, Verhaltensintention und Verhalten der Schüler wurden in Vor- und Nachtests erhoben. Im Nachtest wurden sie zusätzlich gebeten, eine Begründung für ihr Verhalten in offener Form zu geben (vgl. Bewertungskompetenz). In beiden Studien lernten die Schüler unabhängig vom Treatment signifikant. Außerdem verfügte ein signifikant höherer Prozentsatz Schüler nach der Intervention über einen Organspendeausweis. Die Schüler waren außerdem involvierter, änderten jedoch nicht ihre Einstellung zum Thema. Die Schüler, die nach der Intervention über einen Organspendeausweis verfügten, waren nicht involvierter oder lernten mehr über das Thema Organspende als Schüler ohne Ausweis. Zudem waren ihre Begründungen für ihr Verhalten nicht von höherer argumentativer Qualität. Die Schüler mit Organspendeausweis hatten jedoch bereits vor der Intervention eine positivere Einstellung und eine höhere Verhaltensintention. Der Unterrichtsfilm involvierte die Schüler zudem nicht mehr als die neutral gestaltete PowerPoint-Präsentation. Betrachtet man die Entscheidung der Schüler mit Organspendeausweis nach der Intervention genauer, so zeigt sich, dass in beiden Studien ein Teil der Schüler einer Organentnahme nach dem Tod widerspricht. Diese Schüler hatten bereits vor der Intervention eine negative (Studie 1) oder neutrale (Studie 2) Einstellung. Aufklärung zum Thema Organspende überzeugt Schüler also nicht davon, ihren Willen schriftlich zu dokumentieren oder einer Organentnahme zuzustimmen. Hingegen führt sie zur Verringerung einer vorhandenen Diskrepanz zwischen Einstellung und Verhalten. Das trifft auch auf den Unterrichtsfilm zu, der besonders involvierend gestaltet und auf Jugendliche zugeschnitten ist. Es lässt sich annehmen, dass die unterrichteten Schüler selbstbestimmt in ihrer Entscheidung für oder gegen den Organspendeausweis bzw. die Organentnahme blieben. Die Anteile Schüler, die einen Organspendeausweis ausfüllten bzw. einer Organentnahme zustimmten, waren in Studie 2, die nach Einführung der Entscheidungslösung durchgeführt wurde, jeweils kleiner. Im selben Jahr waren mehrere Organspendeskandale publik geworden, die zu einem Rückgang an Organspenden geführt hatten (DSO, 2012). Es gab jedoch auch in Studie 2 kein Schüler den potenziellen Organmissbrauch als Grund für seine Entscheidung an.

BÄK (Bundesärztekammer) (2011). Modell einer Selbstbestimmungslösung zur Einwilligung in die Organ- oder Gewebespende.

BMG (Bundesministerium für Gesundheit). (2012). Entscheidungslösung bei der Organspende tritt in Kraft. Pressemitteilung Nr. 74. Berlin.

DSO (Deutsche Stiftung Organtransplantation) (2012). Statement zur Entwicklung der Organspende im dritten Quartal 2012. Presseinformation vom 11.10.2012. Frankfurt a. M.

DSO (Deutsche Stiftung Organtransplantation) (2013a). Organspende und Transplantation in Deutschland: Jahresbericht 2012. Frankfurt a. M.

KMK (2004). Bildungsstandards im Fach Biologie für den Mittleren Bildungsabschluss. München, Neuwied: Luchterhand.

Vollmann, J. (2012). Ethische Aspekte der Gesundheitskommunikation über die postmortale Organspende. In Bundeszentrale für gesundheitliche Aufklärung (BZgA) (Hrsg.), Aufklärung zur Organ- und Gewebespende in Deutschland: Neue Wege in der Gesundheitskommunikation (Forschung und Praxis der Gesundheitsförderung Nr. 40, S. 108–113). Köln.

Watzke, D. & Stander, V. (2010). Einstellung, Wissen und Verhalten der Allgemeinbevölkerung zur Organ- und Gewebespende. Köln.



## **Die Benotung des Arbeits- und Sozialverhaltens von Schülerinnen und Schülern in Nordrhein-Westfalen. Befunde aus der Längsschnittstudie PARS**

Nachdem die Bewertung des Arbeits- und Sozialverhaltens von Schülerinnen und Schülern (umgangssprachlich: ‚Kopfnoten‘) in den 1960er und 1970er in den meisten Ländern der Bundesrepublik Deutschland abgeschafft wurde (nach 1989 auch in den neuen Ländern), kam es in den vergangenen Jahren in zahlreichen Ländern zu deren Wiedereinführung (Brahm, 2006). Diese bildungspolitische Maßnahme löste heftige Diskussionen seitens der Medien, der Eltern- und Lehrerverbände, der Direktorenvereinigungen sowie der Schülervertretungen aus. Während Befürworter dieser Bewertungsform neben der Sozialisationsfunktion der Schulen auch auf mögliche Chancen für Schülerinnen und Schüler mit schlechten Fachnoten verweisen, sehen Kritiker neben dem erheblichen zeitlichen Mehraufwand für die Lehrerkollegien vor allem Probleme hinsichtlich der Objektivität und Validität der Benotung (Brahm, 2006). Trotz der Aktualität des Themas gibt es von wissenschaftlicher Seite bislang kaum empirische Studien zu diesem Thema (vgl. zum Überblick Frey & Bonsen, 2013). Für Nordrhein-Westfalen – wo im Jahr 2007 Kopfnoten zunächst eingeführt und bereits 2010 wieder abgeschafft wurden – liegen nun Daten vor, die für diesen Beitrag ausgewertet werden.

Zwei zentrale Forschungsfragen werden bearbeitet: zum einen hinsichtlich der Einstellungen der beteiligten Personen zur Bewertung des Arbeits- und Sozialverhaltens und zum anderen hinsichtlich des Zusammenhanges von Kopfnoten mit anderen Variablen. Konkret wird gefragt, inwieweit Lehrkräfte sowie Schülerinnen und Schüler eher positive, negative oder neutrale Einstellungen zu diesem Thema haben, beziehungsweise welche Chancen und Risiken sie sehen. Vertiefend werden Schülermerkmale (z. B.

Fachleistungen und Selbsteinschätzungen) und Lehrermerkmale (z. B. Einstellungen zu Kopfnoten) auf ihren Zusammenhang mit den vergebenen Kopfnoten untersucht.

Die Daten für die Analysen stammen aus der nordrhein-westfälischen Längsschnittstudie PARS (Panel Study at the Research School ‚Education and Capabilities‘ in North Rhine-Westphalia), die von 2009 bis 2012 vom Institut für Schulentwicklungsforschung (IFS) der TU Dortmund durchgeführt wurde. An dieser Studie haben sich 50 Sekundarschulen aus Nordrhein-Westfalen beteiligt. Befragt wurden zu jeweils drei Messzeitpunkten Schülerinnen und Schüler in zwei Kohorten (beginnend mit den Jahrgangsstufen 5 und 9) sowie die Lehrkräfte der Schulen, die Schulleitungen und das weitere pädagogisch tätige Personal. Die Eltern der teilnehmenden Schülerinnen und Schüler wurden einmalig befragt. Ergänzend zu der Befragung haben die Schülerinnen und Schüler Kompetenztests in unterschiedlichen Domänen bearbeitet (vgl. Bos et al., 2013).

Die Einstellungen der beteiligten Personen (Schulleitungen, Lehrkräfte, Schülerinnen und Schüler) zum Thema ‚Bewertung des Arbeits- und Sozialverhaltens‘ wurden mithilfe einer im Rahmen einer Vorstudie selbstentwickelten Skala erhoben (Gebauer & Stubbe, 2009, September). Es liegen Daten aus dem Schuljahr 2009/10 (vor Abschaffung der Kopfnoten) und aus dem Schuljahr 2010/11 (nach Abschaffung der Kopfnoten) vor. Zusätzlich wurden die Kopfnoten der teilnehmenden Schülerinnen und Schüler in den Teilbereichen ‚Leistungsbereitschaft‘, ‚Zuverlässigkeit / Sorgfalt‘ und ‚Sozialverhalten‘ (sowie deren Noten in den Fächern Deutsch, Mathematik und Englisch) von den Lehrkräften erfragt. Schließlich liegen noch Selbsteinschätzungen der Schülerinnen und Schüler zu ihrem Sozialverhalten vor (Frey, 2013).

Mit Hilfe von Latent Class-Analysen werden in diesem Beitrag zunächst Einstellungstypen hinsichtlich der Bewertung des Arbeits- und Sozialverhaltens unterschieden (getrennt für Lehrkräfte beziehungsweise Schülerinnen und Schüler). Es wird erwartet, dass sich (wie in der bereits erwähnten Vorstudie) jeweils drei Gruppen unterscheiden lassen: bei den Lehrkräften ‚Ablehnung von Kopfnoten‘, ‚Kritischer Umgang mit Kopfnoten‘ und ‚Befürworten von Kopfnoten‘; bei den Schülerinnen und Schülern ‚Chancen durch Kopfnoten‘, ‚Gleichgültigkeit‘ und ‚Angst vor Kopfnoten‘. In weiteren multivariaten Modellen (z. B. Mehrebenenregressionen) wird der Zusammenhang der vergebenen Kopfnoten mit (1) den Fachnoten, (2) den Kompetenzen in Lesen und Mathematik, (3) den Selbsteinschätzungen der Schülerinnen und Schüler zu ihrem Sozialverhalten und (4) den Lehrertypen hinsichtlich der Einstellung zu Kopfnoten analysiert werden (ergänzt durch Kontrollvariablen wie Geschlecht, soziale Herkunft etc.). Erwartet wird ein relativ enger Zusammenhang mit Fachleistungen (Noten und Kompetenzen) aber auch mit den Selbsteinschätzungen und den Kontrollvariablen.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



Chair(s): Barbara Flunger (Universität Tübingen)

Diskutant/in: Birgit Spinath (Universität Heidelberg)

Die Untersuchung von Möglichkeiten, wie Interesse und Motivation gesteigert werden können, gehört zu den zentralen Anliegen in der empirischen Bildungsforschung, den Fachdidaktiken, der Erziehungswissenschaft und der pädagogischen Psychologie. Insbesondere im naturwissenschaftlichen Unterricht und im Fach Mathematik hat hierbei die Motivationsentwicklung in jüngerer Zeit erneut viel Aufmerksamkeit gefunden, da auch und gerade in diesen Fächern über die Zeit ein Motivationsabfall zu verzeichnen ist (vgl. z.B. Daniels, 2008), der die Schülerleistung beeinflussen kann, und ein hoher Leistungsstand in den Naturwissenschaften und Mathematik oft als zentraler Grundpfeiler prosperierender Gesellschaften bezeichnet wird.

Interventionen zur Steigerung von Motivation und/oder Leistung in Mathematik und den Naturwissenschaften werden in unterschiedlichen Forschungskontexten mit divergierenden theoretischen Bezügen entwickelt. Noch immer zu selten werden die Ansätze aus unterschiedlichen Fachdisziplinen verglichen bzw. interdisziplinäre Ansätze verfolgt. Das Symposium trägt deshalb Erkenntnisse aus vier Interventionsstudien aus unterschiedlichen Perspektiven zusammen. Die Studien wurden in Kooperation der Fachdidaktiken (Englischdidaktik und Mathematikdidaktik) sowie der empirischen Bildungsforschung und pädagogischen Psychologie durchgeführt. Verbindende Elemente sind die Interdisziplinarität und die anspruchsvollen experimentellen Designs, die eine randomisierte Zuweisung zu Kontroll- und Treatmentgruppen aufweisen und relativ große Stichprobengrößen umfassen. Zwei Beiträge untersuchen Interventionseffekte auf Schülerleistungen, zwei weitere Beiträge nehmen die Schülermotivation stärker in den Blick. Der erste Beitrag erforscht, wie bilingualer Naturwissenschafts-Unterricht auf die Entwicklung von Wissen und Motivation im naturwissenschaftlichen Bereich wirkt. Dabei werden Unterschiede im Lernzuwachs von Schüler/innen aus jeweils 15 bilingual und monolingual unterrichteten Klassen verglichen. Der zweite Beitrag untersucht Gestaltungsmerkmale von Lernumgebungen im Mathematikunterricht, genauer im Statistikerunterricht, wobei verschiedene Förderansätze und ihr Einfluss auf das Statistikwissen verglichen werden. Die Stichprobe umfasst 25 Treatment-Klassen und 4 Baseline-Kontrollklassen. Der dritte Beitrag basiert auf einem Ansatz zur Förderung elementarer mathematischer Konzepte, wobei ein Training evaluiert wird, das die stochastische Kompetenz, Entscheidungen unter Unsicherheit zu treffen, steigern soll. Verglichen werden dabei die Lernzuwächse von acht Klassen, die ein Treatment erhielten, mit denen von vier Kontrollklassen. Der vierte Beitrag untersucht basierend auf der Erwartungs-Wert-Theorie von Eccles (1983), inwieweit durch eine Intervention mit Fokus auf der Nützlichkeit der Mathematik die Motivation von Schüler/innen aus 82 Klassen von 25 Gymnasien für dieses Fach gesteigert werden kann.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823

Daniels, Z. (2008). Entwicklung schulischer Interessen im Jugendalter (Vol. 69). Waxmann Verlag.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives* (S. 74-146). San Francisco, CA: W. H. Freeman.

Nicole Piesche (PH Ludwigsburg), Kathrin Jonkmann (Eberhard Karls Universität Tübingen), Jörg-U. Keßler (PH Ludwigsburg), Jan Hollm (PH Ludwigsburg)

## **COMBIS (= Competences in the Bilingual Science Classroom): Effekte des bilingualen Unterrichts auf das sachfachliche Lernen im naturwissenschaftlichen Unterricht an Realschulen in Baden-Württemberg**

### Theoretischer Hintergrund

Bilinguales Lernen und Lehren etabliert sich immer stärker als alternative Form des fremdsprachlichen, aber auch inhaltlichen Lernens und ist ein seit Jahren beforschtes Gebiet. So liegen Studien über die fremdsprachendidaktischen Grundlagen (z. B. Bach, 2008), Möglichkeiten und Grenzen des Einsatzes von Englisch als Arbeitssprache (z. B. Biederstädt, 2008; Butzkamm, 2008) und Diskussionen einer bilingualen Didaktik (z. B. Breidbach, 2008) vor. Inzwischen kann ebenfalls auf vielfältige Arbeiten zur Kompetenzentwicklung in der Fremdsprache zurückgegriffen werden, die zumeist positive Effekte auf den Fremdspracherwerb zeigen (vgl. z.B. Nold et al., 2008; Zydariß, 2007). Lange Zeit fanden der bilinguale Unterricht und die damit verbundene Forschung fast ausschließlich an Gymnasien in hinsichtlich der individuellen Lernvoraussetzungen (z.B. kognitive Grundfähigkeiten, Vorleistungen, familiärer Hintergrund) positiv selektierten Klassen statt. Heute haben jedoch in Baden-Württemberg auch 53 Realschulen einen bilingualen Zug (vgl. <http://www.kultusportal-bw.de/Lde/Startseite/schulebw/Standorte+mit+bilin-gualen+Zuegen>, 05.10. 2013). Wie sich der bilinguale Unterricht an dieser Schulform und in nicht positiv selektierten Klassen auswirkt, ist bislang jedoch weitgehend unerforscht.

Ein noch größerer Forschungsbedarf zeigt sich hinsichtlich der Effekte des bilingualen Unterrichts auf die Lernzuwächse im Sachfach (vgl. Hollm et al., 2013), insbesondere im naturwissenschaftlichen Unterricht (vgl. Bohn & Doff, 2010), der nur selten bilingual angeboten wird.

Bisherige Studien, kommen zu uneindeutigen Ergebnissen. Marsh et al. (2000) fanden in Hong Kong stark negative Effekte der mittleren bis späten Immersion auf die Leistungen in den Naturwissenschaften.

Kondring und Ewig (2005) zeigten keine signifikanten Unterschiede im Lernzuwachs zwischen bilingual und monolingual unterrichteten Schülern einer gymnasialen Jahrgangsklasse 9 auf (siehe auch Haagen-Schützenhöfer, Mathelitsch, & Hopf, 2011).

Koch und Bündler (2006) fanden in naturwissenschaftlichen Arbeitsgemeinschaften am Gymnasium in der 5. und 6. Klasse heraus, dass die englischsprachig unterrichteten Schüler höhere fachliche Lernzuwächse verzeichneten. Selektionseffekte wurden kontrolliert.

Osterhage (2007) kommt mittels der PISA 2003 - Daten zu ähnlichen Befunden, jedoch ließen sich hier Selektionseffekte nicht ausschließen.

Auch Genderfragen wurde im bilingualen Unterricht bislang kaum Beachtung geschenkt. Auf Basis des Modells des dynamischen Selbst (vgl. Hannover, 1997) ist es denkbar, dass das Fach Physik weiblicher konnotiert wird, wenn es in einer Fremdsprache abgehalten wird, und so für Mädchen, die zumeist niedrigere Leistungen und Motivation v.a. in Physik aufweisen (z.B. PISA-Konsortium, 2007), zugänglicher erscheint.

### Fragestellungen

Vor diesen Hintergründen setzt sich die Studie COMBIS v.a. mit folgenden Fragestellungen auseinander:

1. Wie wirkt sich bilingualer NWA-Unterricht in einer unselektierten Stichprobe auf die Entwicklung von Kompetenzen im naturwissenschaftlichen Sachfach aus?
2. Kann durch die Verwendung der Fremdsprache die Motivation der Mädchen für den naturwissenschaftlichen Unterricht gesteigert werden?

### Methode

Die Stichprobengröße belief sich auf 30 sechste Klassen mit insgesamt 722 Schülerinnen und Schüler (48,7 % Mädchen). Die Klassen stammten aus zehn Realschulen in Baden-Württemberg, die kein bilinguales Angebot hatten. Die Klassen wurden randomisiert den beiden Treatmentbedingungen zugeteilt. Diese bestanden aus einer 10-stündigen Unterrichtseinheit zum Thema „Sinken und Schwimmen“ (vgl. Möller, 2005), die entweder auf Deutsch oder bilingual deutsch-englisch unterrichtet wurde. Alle Klassen wurden von derselben Lehrkraft unterrichtet. Zu Beginn der Studie fand ein 90-minütiger Prä-Test statt, dem sich die 3-wöchige Unterrichtsintervention anschloss. In der Folgestunde bearbeiteten die Schülerinnen und Schüler einen Post-Test von 60-minütiger Dauer. 6 Wochen darauf wurde die Studie durch einen 45-minütigen Follow-up-Test abgeschlossen.

### Ergebnisse

Erste Analysen zeigen, dass die Randomisierung erfolgreich war. Die Treatmentgruppen unterscheiden sich in ihren Ausgangsbedingungen (u.a. Englischkompetenzen, physikalisches Vorwissen, thematisches Vorwissen, Motivation und Selbstkonzept) nicht signifikant voneinander.

Erste Analysen zur Leistungsentwicklung im Sachfachwissen, deuten an, dass die monolingual deutsch unterrichteten Schüler im Posttest signifikant höhere Werte ( $d = 0.25$ ) erzielen, sich die Treatmentgruppen im Follow-up-Test jedoch nicht mehr signifikant voneinander unterscheiden.

Die Analysen zur Genderaspekt sind zum momentanen Zeitpunkt noch nicht abgeschlossen.

## ***Einflussfaktoren auf Statistical Literacy***

### Theoretischer Hintergrund

Der Begriff Statistical Literacy bezieht sich im Sinne des Kompetenzbegriffs von Weinert (2001) auf die Partizipation im Bereich der Statistik –, hier sind unter anderem das Treffen von Entscheidungen unter Unsicherheit sowie das Verstehen und Interpretieren statistischer Daten von Bedeutung (Wallman, 1993). Wissen um mögliche Einflussfaktoren auf Statistical Literacy könnte dazu beitragen, Voraussetzungen von Lernenden differenzierter zu beschreiben und dementsprechend wirksame Lernmaterialien zu erstellen.

Innerhalb von Statistical Literacy gibt es Teilaspekte, die entscheidend für den Aufbau dieser Kompetenz sind. Dazu zählt als ein zentrales Element der adäquate Umgang mit statistischer Variabilität, der eine Modellierung der Daten unter Einbezug des Kontextes zum Zweck der Erklärung, Steuerung oder Vorhersage verlangt (Wild & Pfannkuch, 1999). Entsprechend enthält das Kompetenzmodell von Watson und Callingham (2003) Variabilität als Kernelement von Statistical literacy.

Um angesichts statistischer Variabilität adäquat modellieren zu können, müssen Daten zugunsten eines besseren Überblicks häufig verdichtet werden (Kröpfl, Peschek, & Schneider, 2000). Dass dabei in der Regel Information verloren geht, sollte hierbei nicht außer Acht gelassen werden. Ergebnisse von Reading (2002) zeigen, dass Fähigkeiten im Bereich Darstellung und Reduktion von Daten als Indikator für die Ausprägung von Statistical Literacy betrachtet werden können.

Ein Kompetenzmodell im Bereich von Statistical Literacy (Kuntze, Lindmeier, & Reiss, 2008; Kuntze, 2013) verbindet diese beiden Ideen und verortet Schülerfähigkeiten auf fünf hierarchischen Niveaus. Die Rasch-Skalierbarkeit dieses Modells konnte bisher in mehreren Studien empirisch bestätigt werden (vgl. z.B. Kuntze et al., 2008; Fröhlich, Kuntze, & Lindmeier, 2007).

Empirische Evidenz zur spezifischen Rolle von Wissen zu Reduktion und Variabilität innerhalb von Statistical Literacy fehlt bislang. Ebenso gibt es kaum Erkenntnisse zum Einfluss von Bedingungsvariablen wie kognitiven Grundfähigkeiten oder Leseverständnis. Die hier vorgestellte Studie fokussiert daher auf Einflussfaktoren auf Statistical Literacy sowohl im Bereich allgemeiner Lernerdispositionen, als auch auf spezifische Faktoren innerhalb dieser Kompetenz.

### Fragestellungen

Das Projekt ReVa-Stat zielt in zwei Teilstudien auf die folgenden Forschungsfragen ab:

1. Welchen Einfluss haben allgemeine Lernerdispositionen wie Leseverständnis oder kognitive Grundfähigkeiten auf die Ausprägung von Statistical Literacy?
2. Wie hängt Statistical Literacy mit weiteren Kovariaten zusammen?
3. Inwieweit verbessert die Förderung durch Prompting im Bereich von Reduktion bzw. Variabilität die Kompetenz von Lernenden im Bereich von Statistical Literacy?

### Methode

Diesem Projekt liegen Daten von über 600 Achtklässlern (12 – 16 Jahre) aus 29 Klassen an 8 baden-württembergischen Realschulen zu Grunde. Während zweier aufeinanderfolgender Schulstunden wurden sie zum „Nutzen von Darstellungen und Modellen in statistischen Kontexten“ (vgl. Kuntze, 2013) und zu weiteren individuellen Dispositionen getestet bzw. befragt: Das Arbeitsgedächtnis wurde in Anlehnung an den „Verbalen Lern- und Merkfähigkeitstest“ (Helmstädter, Lendt & Lux, 2001), verbale und nonverbale kognitive Fähigkeiten über Subskalen des „Kognitiven Fähigkeitstests 4-12+R“ (Heller & Perleth, 2000) sowie das Leseverständnis über den „Leseengeschwindigkeits- und -verständnistest“ (Schneider, Schlagmüller, & Ennemoser, 2007) gemessen. Während einer vierstündigen Intervention wurden die Lernenden entsprechend eines 2x2-Designs im Bereich von Reduktion und/oder Variabilität gefördert. Der einstündige Nachtest folgte nach nur wenigen Tagen, ein Follow-Up-Test fand nach acht bis neun Wochen statt.

Die Daten wurden mit mehrebenenanalytischen Methoden ausgewertet.

### Ergebnisse

Die Ergebnisse zeigen, dass verbale und nonverbale kognitive Fähigkeiten über andere Lernerdispositionen wie Arbeitsgedächtnis und Leseverständnis dominieren. Auf individueller Ebene konnte die Mathematiknote, auf Klassenebene der durchschnittliche sozioökonomische Status (Paulus, 2009) zusätzlich Varianz aufklären.

Auswertungen zur dritten Forschungsfrage bezüglich spezifischer Einflussfaktoren innerhalb von Statistical Literacy ermöglichen Einblicke in die Bedeutung domänenspezifischer Begriffswissensaspekte.

### Diskussion

Die moderaten Zusammenhänge mit allgemeinen kognitiven Fähigkeiten sprechen dafür, dass es sich bei der untersuchten Kompetenz „Nutzen von Darstellungen und Modellen in statistischen Kontexten“ um ein eigenständiges, bereichsspezifisches Konstrukt handelt. Mit Hilfe der Befunde zu den begriffswissensbezogenen Einflussfaktoren Variabilität und Reduktion soll Statistical Literacy als Kompetenzkonstrukt konkreter modelliert werden.



## „Risk Literacy“ – Risiko und Entscheidungen unter Unsicherheit im Mathematikunterricht der Grundschule

### Theoretischer Hintergrund

Kompetenzen im Umgang mit statistischer Information spielen in der heutigen datenüberfluteten Welt eine übergeordnete Rolle. Dazu gehören vor allem auch Fähigkeiten, Risiken einzuschätzen, zu beschreiben sowie argumentativ mit Risiken umgehen zu können (Schiller & Kuntze, 2012). Befunde aus der kognitionspsychologischen Forschung deuten darauf hin, dass Erwachsene Defizite in diesen Bereichen vorweisen und statistische Information über Risiken in der Medizin und der Finanzwelt oft nicht verstehen (Gaissmaier et al., 2011; Gigerenzer, 2013; Spiegelhalter, 2011).

Die mathematische Modellierung von Risiko gründet sich auf einem Bündel an elementaren mathematischen Konzepten wie dem des Proportionsbegriffs, des Erwartungswerts und der bedingten Wahrscheinlichkeit (Martignon & Krauss, 2009). Durch eine didaktische Reduktion ist es möglich, Vorstufen dieser Konzepte anhand geeigneter Lernmaterialien und Darstellungsformate mit Grundschulkindern zu behandeln. Durch den Einsatz enaktiver Materialien in Form von Steckwürfeln können und sollten Kinder lernen, wie man elementare Stochastik nutzen kann, um Risiko zu modellieren und zu quantifizieren (Kurz-Milcke et al., 2011; Latten et al., 2011). Diese Quantifizierung beruht auch auf der ikonischen Darstellung von Häufigkeiten in Form von Piktogrammen (Brase, 2008). Dies ermöglicht Kindern - ohne den Gebrauch von Bruchzahlen – Wahrscheinlichkeiten und Risiken abzuwägen und einzuschätzen. Im Vortrag werden die Ergebnisse einer Interventionsstudie vorgestellt, die zum Ziel hat, die Effektivität einer Unterrichtseinheit zum Thema „Risiko und Entscheidungen unter Unsicherheit“ zu evaluieren. Die Studie soll Antworten auf die Frage liefern, ob Kinder durch eine kurze Lerneinheit zum Thema Risiko elementare mathematische Konzepte erlernen können. Die Ideen für die Inhalte der Unterrichtseinheit entstanden an der Pädagogischen Hochschule Ludwigsburg in Zusammenarbeit mit dem Harding Center for Risk Literacy am Max-Planck Institut in Berlin.

### Fragestellungen

1. Sind bei Schülerinnen und Schülern in Klasse 4 Intuitionen zu verschiedenen stochastischen Konzepten vorhanden, mit denen Risiko und Entscheidungen unter Unsicherheit modelliert werden können?
2. Lassen sich elementare stochastische Konzepte zu Risiko bereits durch eine kurze Lerneinheit anhand enaktiver Materialien fördern?
3. Ist bei bestimmten stochastischen Konzepten ein erhöhter Lernzuwachs erkennbar?

### Methode

An der Studie, die zwischen Dezember 2012 und Juli 2013 stattfand, nahmen insgesamt 244 Schülerinnen und Schüler im Alter zwischen 8 und 12 Jahren aus sechs Grundschulen im Umkreis von Ludwigsburg teil. Die 12 Klassen wurden in 8 Versuchs- und 4 Kontrollklassen aufgeteilt. Die Versuchsklassen erhielten ein vierstündiges Training, während die Kontrollklassen nicht unterrichtet wurden. Vor der Intervention wurde das Vorwissen der Schülerinnen und Schüler durch einen 30-minütigen Test abgefragt. Nach der Unterrichtseinheit schloss ein Nachtest an, in welchem Aufgaben zu den gelernten Konzepten gelöst werden sollten. Der Vergleich von Vor- und Nachtestergebnissen sollte letztendlich Aussagen über den Wissenszuwachs ermöglichen. Mögliche Langzeiteffekte des Trainings wurden anhand eines Nachhaltigkeitstest nach drei Monaten festgestellt.

### Ergebnisse

Die Analyse der Daten deutet auf ein außerschulisches Vorwissen der Kinder bezüglich der trainierten Konzepte in Form von „mathematischen Intuitionen“ hin. Die Schülerinnen und Schüler konnten in allen drei mathematischen Bereichen „Proportionen“, „Erwartungswert“ und „bedingte Wahrscheinlichkeit“ einfache Aufgaben lösen. Schwierigkeiten ergaben sich bei Aufgaben, bei welchen es um das mathematisch korrekte Erweitern von Proportionen ging (Bsp.: „2 von 3“ entspricht „4 von 6“). Das Vorwissen zu den untersuchten Konzepten wurde durch die Intervention erfolgreich gefördert und es ergab sich somit ein signifikant höherer Lernzuwachs der Kinder aus den Versuchsklassen gegenüber dem der Kontrollklassen. Gute Prädiktoren für die Voraussage für den Lernerfolg waren das jeweilige Vortestergebnis und allgemeine mathematische Fertigkeiten. Auch drei Monate nach der Intervention unterschieden sich die Testleistungen von Kontroll- und Versuchsgruppe signifikant voneinander, weswegen von einem nachhaltigen Lernzuwachs gesprochen werden kann. Weitere Analysen sollen Aufschluss darüber geben, auf welche mathematischen Konzepte die Intervention den größten/geringsten Lernzuwachs bewirkt hat.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



Hanna Gaspard (Eberhard Karls Universität Tübingen), Anna-Lena Dicke (Eberhard Karls Universität Tübingen), Barbara Flunger (Eberhard Karls Universität Tübingen), Brigitte Schreier (Eberhard Karls Universität Tübingen), Isabelle Häfner (Eberhard Karls Universität Tübingen), Ulrich Trautwein (Eberhard Karls Universität Tübingen), Benjamin Nagengast (Eberhard Karls Universität Tübingen)

## ***Effekte einer Nützlichkeitsintervention im Fach Mathematik auf Wertüberzeugungen von Schülerinnen und Schülern***

Wertüberzeugungen im Sinne der Erwartungs-Wert-Theorie von Eccles und Kollegen (1983) sind zentrale Prädiktoren für lernbezogenes Verhalten und akademische Entscheidungen (z.B. Meece, Wigfield, & Eccles, 1990; Simpkins, Davis-Kean, & Eccles, 2006). Die Erwartungs-Wert-Theorie unterscheidet dabei vier Komponenten der Wertüberzeugungen: intrinsischen Wert, Wichtigkeit, Nützlichkeit und Kosten. Während der Sekundarstufe nehmen Wertüberzeugungen insbesondere für das Fach Mathematik ab (z.B. Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Watt, 2004). Weiterhin zeigen sich in der Sekundarstufe stereotypkonforme Geschlechtsunterschiede in Wertüberzeugungen im Fach Mathematik (Frenzel, Pekrun & Goetz, 2007; Steinmayr, & Spinath, 2008; Watt, 2004).

Im Hinblick auf diese Befunde stellt sich die Frage, wie einem solchen Motivationsabfall entgegengewirkt werden kann und wie dabei insbesondere auch Mädchen angesprochen werden können. Im Vergleich zu den anderen Wertkomponenten scheint die Nützlichkeitsüberzeugung, die mit extrinsischer Motivation in Verbindung gebracht werden kann (Wigfield, Tonks & Klaua, 2009), leichter von außen beeinflussbar zu sein. In verschiedenen Studien konnte durch Nützlichkeitsinterventionen das Interesse und die Leistung bei Studierenden in den Bereichen Mathematik und Psychologie und bei Schülerinnen und Schülern im Bereich Naturwissenschaften gesteigert werden (Hulleman, Godes, Hendricks & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009).

Es bleiben verschiedene offene Fragen, die in dieser Studie beantwortet werden sollen. Erstens ist unklar, inwiefern die Ergebnisse früherer Interventionsstudien auf den deutschen Schulkontext und das Fach Mathematik übertragen werden können. Die Studie untersucht daher die zentrale Frage, ob sich Wertüberzeugungen von Schülerinnen und Schülern der neunten Klasse in diesem Fach durch Interventionen im Schulkontext fördern lassen. Zweitens stellt sich die Frage, welche Interventionsmaterialien eine optimale Förderung der Motivation ermöglichen. In der vorliegenden Studie werden zwei verschiedene Strategien zur Förderung der Nützlichkeitsüberzeugung miteinander verglichen. Drittens sollen aufgrund von den in der bisherigen Forschung gezeigten Geschlechtsunterschieden in Wertüberzeugungen im Fach Mathematik differentielle Effekte der Intervention für Mädchen und Jungen untersucht werden.

An der Studie nahmen insgesamt 82 neunte Klassen aus 25 Gymnasien in Baden-Württemberg teil. Die teilnehmenden Klassen wurden randomisiert einer von zwei Interventionsbedingungen oder einer Wartekontrollgruppe zugewiesen. Die Intervention bestand aus einer zweistündigen Unterrichtseinheit zur Nützlichkeit der Mathematik, die aus einem interaktiven Vortrag und einem Arbeitsauftrag bestand. In der ersten Bedingung verfassten die Schülerinnen und Schüler am Ende der Einheit einen Text zur persönlichen Relevanz der Mathematik. In der zweiten Bedingung wurden Interviewzitate zur Nützlichkeit von Mathematik bewertet. Vor der Intervention (Prätest) sowie sechs Wochen (Posttest) und vier Monate nach der Intervention (Follow-Up) füllten insgesamt 1978 Schülerinnen und Schüler (53.6% weiblich) Fragebögen aus. Darin wurden die Wertüberzeugungen im Fach Mathematik in Bezug auf die vier Komponenten des Erwartungs-Wert-Modells erfasst. Die Interventionseffekte wurden in Mehrebenen-Regressionsanalysen unter Kontrolle der Ausgangswerte geprüft.

Im Vergleich zur Wartekontrollgruppe fanden sich zum Posttest positive Effekte der Zitate-Bedingung auf die Nützlichkeit und die Wichtigkeit. Zum Follow-Up nahmen Klassen der Zitate-Bedingung im Vergleich zur Wartekontrollgruppe Mathematik immer noch als nützlicher und wichtiger wahr, außerdem berichteten sie nun auch einen höheren intrinsischen Wert der Mathematik. Die Text-Bedingung zeigte sowohl zum Posttest als auch zum Follow-Up ebenfalls positive, wenn auch geringere Effekte auf die wahrgenommene Nützlichkeit, nicht jedoch auf die anderen Wertkomponenten. Weiterhin fanden sich für die Effekte auf Nützlichkeit und intrinsischen Wert Interaktionen mit dem Geschlecht, die darauf hindeuten, dass beide Interventionsbedingungen für Mädchen stärkere positive Effekte hatten als für Jungen.

Durch eine zweistündige Intervention im Klassenkontext konnte in dieser Studie die Motivation von Schülerinnen und Schülern der neunten Klasse nachhaltig gefördert werden, auch wenn die Interventionseffekte von ihrer Effektstärke her eher gering waren. Dabei erwies sich die Beurteilung von Zitaten als für die Motivation förderlicher als das Schreiben von Texten. Schließlich zeigte sich die Intervention als besonders geeignet zur Förderung der Motivation von Mädchen.

**GEBF** TAGUNG  
2014

E P5 d

Poster-  
Symposiums-  
beitrag

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823

2

Unterricht und schulbezogene Interventionen

## **Konstruktivistisches Unterrichten ist Einstellungssache: Eine Mediationsanalyse des Zusammenhangs von Wissen und Unterrichtspraxis.**

Theorie: Die Frage nach dem Einfluss von Pedagogical Content Knowledge (PCK), aber auch Pedagogical Knowledge (PK) auf die Unterrichtspraxis ist ein zentraler Angelpunkt in der Bildungsforschung. Ungeklärt scheint jedoch der Zusammenhang zwischen Wissen und konstruktivistisch orientierter Unterrichtspraxis zu sein. Einerseits kann man davon ausgehen, dass PCK prädiktiv für kognitiv aktivierende Instruktionen und Lernunterstützung ist (Kunter et al., 2013). Andererseits können beim forschend-entdeckenden Lehren in den Naturwissenschaften Einstellungsfaktoren (bspw. Selbstwirksamkeit oder die Wichtigkeit von Naturwissenschaftsunterricht) mit verschiedenen Wissenstypen in Verbindung gebracht werden (z.B. Alake-Tuenter et al., 2013). Nach Van Merriënboer und Paas (2003) ist wirksame Unterrichtspraxis eine Integration verschiedener Fähigkeiten, Wissensinhalte und Einstellungen und dies, zusammen mit der Durchführung einer Kette von Handlungen (Maiello & Oser, 2009), kann eine situationsspezifische Disposition bzw. ein Kompetenzprofil formen (ibid.). Der „reasoned action“-Ansatz (Fishbein & Ajzen, 2010) bestätigt, dass Einstellungen dem Handeln vorausgehen und Verhalten beeinflussen (Blömeke, Felbrich, & Müller, 2009). Dies legt einen Mediationsansatz zur Erklärung des Zusammenhangs zwischen Wissen und Praxis via Einstellungen nahe, sodass sich Wissen als „nur indirekt handlungsrelevant“ darstellt (Oser, 2013, S. 39).

Forschungsfrage und Hypothesen: Die Forschungsfrage richtet sich auf Einstellungen und PCK/ PK und deren Einfluss auf die konstruktivistisch orientierte Unterrichtspraxis. Gibt es eine direkte Verbindung zwischen diesen Variablen oder ist sie durch Einstellungen mediiert? Die Hypothesen lauten, dass a) eine positive Einstellung zum forschend-entdeckenden Lernen, eine positive Einstellung zur Relevanz von naturwissenschaftlichem Unterricht und die Selbstwirksamkeit Prädiktoren für die Unterrichtspraxis sind, und dass b) PCK und PK einen mediierten Einfluss auf die Unterrichtspraxis haben.

Methoden: 160 Schweizer Lehrpersonen (Chemie, Biologie, Physik und integrierter naturwissenschaftlicher Unterricht; Kindergarten bis Sekundarstufe I) wurden anfangs Schuljahr 2012/2013 mit einem Kompetenztest (Vignetten) und einem Fragebogen evaluiert. PCK und PK wurden online mit einem Vignettentest erfasst (Brovelli et al., 2012), bei dem die Lehrpersonen acht Unterrichtssituationen (Vignetten) schriftlich kommentieren sollten. Konstruktivistische Unterrichtspraxis, Einstellungen und Selbstwirksamkeit wurden auf einer 4-stufigen Likert-Skala bewertet. Generelles konstruktivistisches Unterrichtsverhalten (KU) wurde nach Fishbein und Ajzen (2010) qua Selbsteinschätzung approximiert, die inhaltliche Operationalisierung basiert auf Duit und Wodzinski (2006), Widodo und Duit (2004) sowie Muijs und Reynolds (2011). Die Einstellung zum forschend-entdeckenden Lernen wurde nach van Hooff et al. (2005) konstruiert. Die Einstellung zur Relevanz von naturwissenschaftlichem Unterricht stammt aus Blömeke, Felbrich und Müller (2009), Selbstwirksamkeit von Rakoczy, Buff und Lipowsky (2005). Cronbach Alpha der Skalen lag zwischen .63 und .81.; die Interrater-Reliabilität für PCK und PK zwischen .64 und .86 (Krippendorffs Alpha für ordinale Daten). Lehrpersonen unterschiedlicher Schulstufen unterschieden sich nicht in PCK und PK.

Ergebnisse: Die Regressionsanalyse ( $R^2=.55$ ) verdeutlicht, dass eine positive Einstellung zum forschend-entdeckenden Lernen (EFL) das konstruktivistische Unterrichtsverhalten (KU) vorhersagt ( $\beta=.41$ ,  $p<.001$ ) ebenso wie die Selbstwirksamkeit (SWE) ( $\beta=.36$ ,  $p<.001$ ) und die Einstellung zur Relevanz von naturwissenschaftlichem Unterrichten (RNU) ( $\beta=.21$ ,  $p=.001$ ). PCK und PK hatten keinen signifikanten Effekt. Multiple Mediationsanalysen zeigten indirekte Effekte von PCK via EFL und SWE, aber nicht via RNU. Der direkte Pfad zwischen PCK und KU war nicht signifikant. Die Pfad-Koeffizienten zwischen PCK und EFL (.20) und PCK und SWE (.18) waren signifikant für  $p<.05$ ; zwischen EFL und KU (.43), SWE und KU (.37) und zwischen RNU und KU (.22), alle  $p<.001$ . PK als Prädiktor ( $R^2=.55$ ) zeigte keine signifikanten Effekte. Nur die Mediatoren EFL (.42), RNU (.22) und SWE (.36), alle  $p<.001$ , wirkten auf KU. Das Gesamtmodell zeigt eine gute Modellanpassung ( $X^2(5)=6.310$ ,  $p=.277$ ; CFI=.99; TLI=.98; RMSEA=.04).

Diskussion: Die Resultate belegen einen indirekten Einfluss von PCK auf die konstruktivistische Unterrichtspraxis. PCK konnte die Unterrichtspraxis nicht direkt voraussagen, sondern die Einstellung zur Implementierung von erforschendem Lernen sowie die Selbstwirksamkeit wirkten mediiierend. PK steht lediglich mit PCK in Zusammenhang und weist im Modell keinen Bezug zu konstruktivistischem Handeln im naturwissenschaftlichen Unterricht aus.

Alake-Tuenter, E., Biemans, H. J. A., Tobi, H., & Mulder, M. (2013). Inquiry-based science teaching competence of primary school teachers: A Delphi study. *Teaching and Teacher Education*, 35, 13-24.

Blömeke, S., Felbrich, A., & Müller, C. (2009). Future Teachers' Beliefs on the Nature of Mathematics. In F. Achtenhagen, F. K. Oser & U. Renold (Eds.), *Teachers' professional development : aims, modules, evaluation* (pp. 25-46). Rotterdam [u.a.]: Sense Publ.

Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2012). Capturing professional competencies for teaching science - a vignette test with authentic teaching situations and a complex open-response format. [Article in German]. *Unterrichtswissenschaft*.

Duit, R., & Wodzinski, C. T. (2006). Planning Good teaching. Categories in didactic thinking when planning lessons. *Science in the classroom: physics*. [Article in German]. *Unterricht überdenken, Unterricht entwickeln*(92), 9-11.

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior : the reasoned action approach*. New York, NY [u.a.]: Psychology Press.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional Competence of Teachers: Effects on Instructional Quality and Student Development. *Journal of Educational Psychology*, 105(3), 805-820.

Maiello, C., & Oser, F. K. (2009). Psychological Variables and Evaluation of Distinctive Characteristics of Teaching Standards. In F. Achtenhagen, F. K. Oser & U. Renold (Eds.), *Teachers' professional development : aims, modules, evaluation* (pp. 3-23). Rotterdam [u.a.]: Sense Publ.

Muijs, D., & Reynolds, D. (2011). *Effective teaching : evidence and practice* (2. ed.). London: Sage.

Oser, F. (2013). Kompetenzen der Lehrenden und Auszubildenden: Ein Ressourcenmodell. In F. Oser, T. Bauder, P. Salzmann & S. Heinzer (Eds.), *Ohne Kompetenz keine Qualität. Entwickeln und Einschätzen von Kompetenzprofilen bei Lehrpersonen und Berufsverantwortlichen*. (pp. 29-65). Bad Heilbrunn: Klinkhardt.

Rakoczy, K., Buff, A., & Lipowsky, F. (2005). Documentation of the data collection and evaluation tools in the Swiss-German video study. "Quality of teaching, learning and behavior mathematical understanding." 1st survey instruments. [Report in German]. Frankfurt/ M.: GEBF/ DIFP.

van Hooff, E. A. J., Born, M. P., Taris, T. W., van der Flier, H., & Blonk, R. W. B. (2005). Bridging the gap between intentions and behavior: Implementation intentions, action control, and procrastination. *Journal of Vocational Behavior*, 66(2), 238-256.

Van Merriënboer, J. J. G., & Paas, F. (2003). Powerful learning and the many faces of instructional design: Toward a framework for the design of powerful learning environments. In E. De Corte, L. Verschaffel, N. Entwistle & J. J. G. Van Merriënboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions*. (pp. 3-20). Amsterdam et al.: Eri-Pergamon.

Widodo, A., & Duit, R. (2004). Constructivist perspectives on teaching and learning and the practice of teaching physics. [Article in German]. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 233-255.



Sara Leyener (Universität Basel), Christian Herrmann (Universität Basel), Uwe Pühse (Universität Basel), Erin Gerlach (Universität Potsdam)

## ***Einfluss der Unterrichtsmerkmale auf die Lernleistungen im Sportunterricht. Ergebnisse aus der IMPEQT-Studie.***

Das theoretische Rahmenmodell (Helmke & Weinert, 1997) zeigt verschiedene Bedingungsfaktoren (wie z.B. Klasse, Lehrer, Unterrichtsqualität, etc.) auf, die Einfluss auf Schülerleistungen haben. Der Unterrichtsqualität wird dabei ein besonderer Stellenwert auf die Lern- und Leistungsergebnisse der SchülerInnen zugewiesen (Hattie, 2013). In der empirischen Bildungsforschung finden sich eine Reihe von Studien, die der Frage des Einflusses von Prozessmerkmalen auf Leistungen nachgehen (im Überblick Helmke, 2010). Allerdings ist die Befundlage für das Fach Sport im Vergleich zu den Hauptfächern unbefriedigend. Noch ungeklärt ist, was als Lernleistung eines Schülers im Fach Sport zu begreifen ist - fähigkeitsorientierte Tests reichen nicht aus, da sie die zu erwerbenden Kompetenzen unangemessen reduzieren. Der derzeit einzige Ansatz, der in Ansätzen diesen Forderungen nachkommt, ist u. E. der Ansatz der motorischer Basisqualifikationen (Kurz, Fritz & Tscherpel, 2008). Die zentrale Annahme dabei ist, dass diese komplexen und kontextbezogenen Basisqualifikationen die zentrale Voraussetzung für einen Zugang zur Bewegungs-, Spiel- und Sportkultur darstellen. Dieser Beitrag geht der Fragestellung nach, inwieweit Merkmale guten Unterrichts die MOBAQs beeinflussen.

In der, durch die Eidgenössische Sportkommission finanzierten, IMPEQT-Studie (Implementation of Physical Education and the Quality of Teaching) der Universität Basel, wurden im Rahmen eines einjährigen Längsschnitts zwölf motorische Basisqualifikationen sowie Unterrichtsmerkmale mittels Fragebogen in 42 Klassen an Schweizer Schulen (N = 948) erhoben.

Erste querschnittliche Ergebnisse zeigten, dass vor allem auf Individualebene die Merkmale „Angst vor der Lehrkraft“ und „Motivation durch die Sportlehrkraft“ eine wichtige Rolle spielen. Auf Klassenebene hingegen die Merkmale „Klassenführung“ sowie „Disziplin & Zeitnutzung“ wichtig sind. Die längsschnittlichen Ergebnisse sind demnach mit Spannung zu erwarten, da sie empirisches Hintergrundwissen generieren, das auch für die Lehrerausbildung von hoher Relevanz sein dürfte.

Hattie, J. (2013). Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von Visible Learning von W. Beywl und K. Zierer. Baltmannsweiler: Schneider

Helmke, A. (2010). Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts (3. Aufl.). Seelze-Velber: Kallmeyer.

Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), Psychologie des Unterrichts und der Schule (S. 71-176). Göttingen: Hogrefe.

Kurz, D., Fritz, T. & Tscherpel, R. (2008). Der MOBAQ-Ansatz als Konzept für Mindeststandards für den Sportunterricht? In V. Oesterhelt, J. Hofmann, M. Scholz & H. Altenberger (Hrsg.), Sportpädagogik im Spannungsfeld gesellschaftlicher Erwartungen, wissenschaftlicher Ansprüche und empirischer Befunde (S. 97-106). Hamburg: Czwilina.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## **Individuelle Förderung von Schüler/innen im Klassenkontext: Schülerwahrnehmungen in Schweden und Deutschland**

### Hintergrund

Die Überlegenheit skandinavischer Schüler/innen in Schulleistungsvergleichen wird oft mit Eigenschaften der dortigen Schulsysteme erklärt. Das schwedische Schulsystem beispielsweise setzt auf ein Gesamtschulkonzept, stellt alle Lernmittel frei, vergibt Noten erst ab der achten Klasse und kennt kein Sitzenbleiben. Im Unterrichtsgeschehen zeigen sich individualisierte und unterstützende Lernformen. Diese korrespondieren mit einer Binnendifferenzierung in der Klasse: Schüler/innen verfolgen unterschiedliche Lernziele und haben unterschiedliche Lern tempi. Schulbücher, Arbeitshefte und Wochenpläne sind daran angepasst. Lehrkräfte sind darauf eingestellt, ein individualisiertes, förderndes und unterstützendes Lernklima zu schaffen.

Wird Förderung und Unterstützung im Unterricht von schwedischen Schüler/innen auch anders wahrgenommen als von deutschen Schüler/innen? Die vorliegende Studie untersucht exemplarisch die Wahrnehmung von Merkmalen des deutschen und des schwedischen Mathematikunterrichts durch Schülerinnen und Schüler im direkten Vergleich.

### Methode

An der fragebogenbasierten Studie nahmen N = 88 Schüler/innen im Durchschnittsalter von 14 Jahren teil (n = 41 schwedische Schüler/innen, davon 20 Mädchen; n = 47 deutsche Schüler/innen, davon 24 Mädchen). In Schweden wurden zwei Klassen der 7./8. Jahrgangsstufe einer Schule auf Gotland befragt; in Deutschland wurden zwei Klassen der 7./8. Jahrgangsstufe eines Gymnasiums in Mecklenburg-Vorpommern befragt.

Der selbstentwickelte Fragebogen wurde parallel auf Deutsch und auf Schwedisch erstellt. Im Bereich Klassenklima wurden abgestufte (1-4) Zustimmungen zu generellen Aussagen (wie beispielsweise „Ich fühle mich in meiner Klasse wohl“ oder „Ich kann mich in meiner Klasse gut konzentrieren“) erfragt. Die Fragen zum Unterrichtsgeschehen und zur Förderung bezogen sich stets auf den Mathematikunterricht. Im Bereich Unterrichtsgestaltung und Differenzierung wurden Verhaltensweisen in Bezug auf die unterrichtende Lehrkraft erfragt (z.B. „Wie oft bittest du deinen Mathematiklehrer um Hilfe?“; „Mein Mathematiklehrer nimmt sich ausreichend Zeit, meine Frage zu beantworten“). Im Bereich Gruppenarbeit wurde erfragt, inwiefern diese im Mathematikunterricht eingesetzt wird, aber auch, in welchem Maße die Schüler/innen Gruppenarbeit als sinnvoll und hilfreich für die Förderung ihres Leistungspotentials einschätzen (z.B. „Ich trage mehr zur Lösung der Aufgaben bei als meine Mitschüler/innen“). Weiter wurden Fragen zur Förderung der Schüler/innen außerhalb des regulären Unterrichtes sowie zur materiellen Ausstattung der Schule und dem Einsatz von Medien gestellt. Schließlich wurden Selbsteinschätzungen der Mathematikleistung durch die Schüler/innen erfasst (z.B. „Ich bin gut in Mathematik“; „Denkst du, dass du dich im letzten halben Jahr in Mathematik verbessert hast?“).

Der Fragebogen wurde im Januar 2013 in Deutschland und in Schweden von Lehrkräften der Klassen erhoben.

### Ergebnisse

Die Ergebnisse zeigen, dass sich schwedische Schüler/innen im Vergleich zu deutschen Schüler/innen von ihren Mathematiklehrkräften und im Mathematikunterricht stärker gefördert fühlen. Dazu gehört beispielsweise, dass sich die schwedischen Schüler/innen stärker von ihrem Mathematiklehrer unterstützt und verstanden fühlen, dass die Lehrkraft ihr Lerntempo besser einschätzen und ihr Selbstvertrauen in größerem Maße stärken kann. Achtzig Prozent der schwedischen Schüler/innen sind der Meinung, dass sich ihr Mathematiklehrer ausreichend Zeit für sie nimmt. Im Gegensatz dazu stimmen nur 45 Prozent der deutschen Schüler und Schülerinnen dieser Aussage zu.

Darüber hinaus werden im schwedischen Mathematikunterricht häufiger als im deutschen Mathematikunterricht Gruppenarbeitsformen eingesetzt, welche die schwedischen Schüler/innen im Vergleich zu den deutschen Schüler/innen auch als stärker fördernd einschätzen.

Die Ergebnisse zeigen außerdem, dass Schüler/innen in Schweden in größerem Maße durch die Schule gefördert werden. Dies bezieht sich auf zusätzliche Förderangebote am Nachmittag innerhalb der Schule, die in Schweden durch die Lehrkraft realisiert werden, die in der Klasse auch regulär unterrichtet.

Bezüglich des Klassenklimas (generelle Aussagen ohne Bezug auf unterrichtliches Handeln) wurde kein Unterschied zwischen deutschen und schwedischen Schülerwahrnehmungen festgestellt.

### Diskussion

Die vorliegende Studie konzentrierte sich auf Qualitätsmerkmale des Unterrichts, die mit Blick auf die zunehmende Heterogenität in Deutschland größere Beachtung verdienen. Die Ergebnisse zeigen, dass eine entsprechende Unterrichtskultur auch als fördernd und unterstützend wahrgenommen wird. Umgekehrt wird die Unterrichtskultur im deutschen Mathematikunterricht von deutschen Schüler/innen im Vergleich als weniger fördernd und unterstützend wahrgenommen.



## Wie verändert sich Autonomie aus der Sicht von Schülern im Kontext des produktiven Klassengesprächs?

Theoretischer Hintergrund: Die Forschung zur Lehrer-Schüler Interaktion zeigt deutlich, dass das Klassengespräch die meist verbreitetste Instruktionsform im Unterricht darstellt (u.a. Wells & Arauz, 2006). In Deutschland ist das Klassengespräch darüber hinaus stark lehrerzentriert und wird häufig eng geführt, mit nur wenig Raum für aktive Schülerbeteiligung und Autonomie (Jurik, Gröschner, & Seidel, 2013). Findet demgegenüber eine produktive Klassengesprächsführung statt, zeigen sich jedoch positive Wirkungen auf das Schülerlernen (Sierens, Vansteenkiste, Goossens, Soenens, & Dochy, 2009).

Parallel zu diesem Forschungsstrang hat sich im Umfeld der Forschung zur Selbstbestimmungstheorie (SDT), die Forschung zum Autonomieerleben von Schülern und der Autonomieunterstützung von Lehrkräften intensiviert (u.a. Reeve & Jang, 2006). Autonomieerleben und Autonomieunterstützung sind zwei Facetten des Konstrukts Autonomie im Sinne der SDT (Reeve, Jang, Carrell, Soohyun, & Barch, 2004). Dabei stehen zur Förderung des Autonomieerlebens zwei Wege offen: der Abbau von Kontrolle auf Seiten der Lehrkraft und der Aufbau von autonomen Handlungsspielräumen bei den Schülern.

Ein zentraler Indikator der Autonomieunterstützung von Lehrkräften ist dabei die Verwendung nicht-kontrollierende Sprache (Kaplan & Assor, 2012).

Ziel der Studie ist die Zusammenführung der beiden Forschungsstränge, die bisher noch weitgehend unabhängig voneinander Bedingungen für produktive Lehr-Lern-Situationen untersuchen und dabei den Fokus auf die lernförderliche Gestaltung von Lehrer-Schüler Diskursen legen.

Fragestellungen: Es werden folgende Forschungsfragen untersucht:

1. Wie nehmen Schülerinnen und Schüler Autonomie wahr?
2. Inwieweit hängt das produktive Klassengespräch mit Autonomie zusammen?
3. Inwiefern verändert sich das Klassengespräch hin zu nicht-kontrollierender Sprache (als Indikator für Autonomieunterstützung)?

Methode: In einer einjährigen Interventionsstudie zum produktiven Klassengespräch, die mit dem Dialogischen Videozirkel (DVC) arbeitete, wurden zu insgesamt vier Messzeitpunkten (Pre-Test, DVC1, DVC2, Post-Test) Daten von insgesamt 136 Schülerinnen und Schülern (Alter:  $M = 15.41$ ,  $SD = 0.89$ ; 54 weiblich) der neunten und zehnten Jahrgangsstufe erhoben. Alle Daten entstammen dem mathematisch-naturwissenschaftlichen Unterricht. Zur Datenerhebung wurden Fragebögen sowie Videographien des Unterrichts herangezogen, die im Anschluss mit Hilfe der Software Videograph (Rimmele, 2002) ausgewertet wurden. Hierbei wurden Selbsteinschätzungen der Schüler systematisch mit videographierten Verhaltensdaten verknüpft. Alle Fragebogenskalen (Autonomieunterstützung, Autonomieerleben, wahrgenommene Gesprächsqualität) wurden auf einer 4-stufigen Likert-Skala eingeschätzt, während die Auswertungskategorien (Sprecher, Art der Frage, Art der Rückmeldung und Bezug der Schüleräußerung) von unabhängigen Ratern beurteilt wurden. Alle Kategorien sowie die Fragebogenskalen zeigen zufriedenstellende interne Konsistenzen ( $r = .67 - .87$ ) bis sehr gute Interraterübereinstimmungen ( $\kappa = .69 - 1.9$ ).

Ergebnisse: In Bezug auf Forschungsfrage 1 zeigte sich ( $t(123) = 37.04$ ,  $p < 0.1$ ;  $t(122) = 21.82$ ,  $p < 0.1$ ,  $d = 1.3$ ), dass Schüler die zwei Facetten von Autonomie zu Beginn der Studie (Autonomieerleben, Autonomieunterstützung) signifikant unterschiedlich wahrnehmen. Die Unterstützung der Lehrkraft wurde hierbei positiver eingeschätzt (Mittelwerte und SD zu den einzelnen Zeitpunkten). Zudem fand sich im Verlauf der Intervention ein signifikant positiver Anstieg der beiden Konstrukte statt ( $F(6;86) = 3.74$ ,  $p < .01$ ,  $\eta^2 = .21$ ). In Hinblick auf Forschungsfrage 2 zeigten sich mittlere signifikante Korrelationen zwischen der wahrgenommenen Gesprächsführung und den beiden Facetten von Autonomie ( $r = .31 - .45$ ). Eine MANOVA mit Messwiederholung zeigte auch in diesem Fall, signifikant positive Entwicklungen für alle drei Konstrukte im Verlauf der Intervention ( $F(9;63) = 3.83$ ,  $p < .01$ ,  $\eta^2 = .35$ ).

Die Betrachtung von Aspekten nicht-kontrollierender Sprache zeigte eine Zunahme an sachlich-konstruktiven Rückmeldungen sowie eine Zunahme an Gesprächspausen und eine Abnahme geschlossener Fragen. Die stärksten Veränderungen zeigten sich jedoch bei der Bezugnahme von Schüleräußerungen, wobei der Bezug zu Aussagen von Mitschülern von 3 auf 23% im Verlauf des Schuljahres ansteigt.

Jurik, V., Gröschner, A., & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. *Learning and Instruction*(23), 33-42. doi: 10.1016/j.learninstruc.2012.09.002

Kaplan, H., & Assor, A. (2012). Enhancing autonomy-supportive I-Thou dialogue in schools: conceptualization and socio-emotional effects of an intervention program. *Social Psychology of Education*, 15(2), 251-169. doi: 10.1007/s11218-012-9178-2

Reeve, J. M., & Jang, H. (2006). What Teachers Say and Do to support Students' Autonomy During a Learning Activity. *Journal of Educational Psychology*, 98(1), 209 - 218. doi: 10.1037/0022-0663.98.1.129

Reeve, J. M., Jang, H., Carrell, D., Soohyun, J., & Barch, J. (2004). Enhancing Students' Engagement by Increasing Teachers' Autonomy Support. *Motivation and Emotion*, 28(2), 147-169.

Rimmele, R. (2002). Videograph. Kiel: IPN.

Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology*, 79(1), 57-68. doi: 10.1348/000709908X304398

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *The Journal of the Learning Sciences*, 15(3), 379-428.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823

Chair(s): Heike Itzek-Greulich (Pädagogische Hochschule Ludwigsburg)

Diskutant/in: Markus Vogel (Pädagogische Hochschule Heidelberg)

In diesem Symposium werden vier Studien zusammengeführt, die mit zwei unterschiedlichen Forschungsdesigns auf eine systematische Evaluation und Verbesserung von Lehr-Lernarrangements in der Schul- und Unterrichtspraxis abzielen. Es werden hierbei zum einen zwei randomisierte Interventionsstudien vorgestellt, die innovative Lehr-Lernansätze erprobten und deren Wirksamkeit untersuchten. Randomisierte Interventionsstudien gelten als „gold standard“ in der Entwicklung und Prüfung von Ansätzen zur Verbesserung der Effektivität von Lernumgebungen (vgl. No Child Left Behind Legislation, 2001). Aber auch andere Designs ermöglichen aussagekräftige Bewertungen der Wirksamkeit von Lernumgebungen (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2005) und können u.a. dann verwendet werden, wenn eine Randomisierung bzw. experimentelle Variation im Treatment nicht möglich oder sinnvoll ist. In diesem Sinne werden in dem Symposium auch zwei nicht-experimentelle Studien präsentiert.

Die vier Einzelbeiträge entstammen vier unterschiedlichen Domänen (Chemie, Physik, Englisch und Mathematik) und demonstrieren deutlich die wichtige Rolle, die ein fachspezifischer Zugang in der Forschung zur Effektivität von Lehr-Lernarrangements spielen muss. Auch wird deutlich, dass je nach Fragestellung unterschiedliche methodische Zugriffe (z.B. quasi-experimentelle Designs mit Pre-Post Erhebungen unter Berücksichtigung von Kontrollgruppen) realisiert werden müssen. Gleichzeitig sind alle vier Projekte dadurch charakterisiert, dass die Datengrundlage aufgrund der großen Stichproben (n auf Klassenebene > 50) hohe Power verspricht und die Voraussetzung für mehrebenenanalytische Verfahren (vgl. Raudenbush & Bryk, 2002) schafft.

Die vorgestellte Studie aus der Chemiedidaktik erforscht anhand einer Interventionsstudie, welche Effekte die Einbindung eines Schülerlaborbesuchs in den Schulunterricht im Vergleich zu herkömmlichen Formen (nur Schülerlaborbesuch/nur Schulunterricht) auf die Schülerleistung hat. Die zweite Interventionsstudie stammt aus der Physikdidaktik und untersucht im Kontext der Wärmelehre den Einfluss des world wide webs auf das Hausaufgabenverhalten und dessen sinnvollen Einsatz im Physikunterricht. Der dritte Beitrag befasst sich mit der Wirksamkeit des bilingualen Geschichtsunterrichts, wobei der Geschichtsunterricht in deutscher wie in englischer Sprache verglichen wird. Des Weiteren wird über Interaktionseffekte in Bezug auf Schülermotivation, Emotionen und Kompetenzen zwischen Geschichts- und regulärem Englischunterricht berichtet. Im vierten Beitrag wurde basierend auf der Erwartungs-Wert-Theorie (Eccles, 1983) an neunten Gymnasialklassen untersucht, wie die Relevanz des Gelernten und die wahrgenommene Wertschätzung des Faches Mathematik mit den Wertschätzungen in diesem Fach zusammenhängen.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823

Eccles, Jacquelynne S. (1983): Expectancies, values, and academic behaviors. In: J. T. Spence (Hg.), *Achievement and achievement motives: Psychological and sociological approaches*. San Francisco: Freeman, S. 75–146.

No Child Left Behind Act of 2001, Public Law No. 107–110.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: applications and data analysis methods* (2. ed.). Beverly Hills: Sage.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W.H., & Shavelson, R.J. (2007). *Estimating causal effects using experimental and observational designs* (Report from the Governing Board of the American Educational Research Association Grants Program). Washington, DC: American Educational Research Association.



Heike Itzek-Greulich (Pädagogische Hochschule Ludwigsburg), Barbara Flunger (Eberhard Karls Universität Tübingen), Christian Vollmer (Pädagogische Hochschule Heidelberg), Benjamin Nagengast (Eberhard Karls Universität Tübingen), Markus Rehm (Pädagogische Hochschule Heidelberg), Ulrich Trautwein (Eberhard Karls Universität Tübingen)

## ***Erhöht die Einbindung eines Schülerlaborbesuchs in den naturwissenschaftlichen Unterricht die Effektivität dieses Lehr-Lernarrangements?***

Experimentieren ist essentiell für das Verständnis der naturwissenschaftlichen Denk- und Arbeitsweisen (Prenzel & Parchmann, 2003). Die Ergebnisse internationaler Vergleichsstudien wie TIMSS und PISA bescheinigen deutschen Schülerinnen und Schülern allerdings Defizite in diesem Bereich. Aufgrund dieser Ausgangslage wurden in den letzten Jahren verschiedene Ansätze entwickelt und erprobt. Dazu gehört auch die Einbeziehung außerschulischer Lernorte und die Nutzung speziell für Schulklassen eingerichteter Programme wie bspw. Schülerlabore (Engeln & Vorst, 2007). Ein wichtiges Ziel aller Schülerlabore ist es, die Begeisterung der Heranwachsenden für das Experimentieren sowie das Verständnis für die Naturwissenschaften zu steigern und auf diese Weise den fachlichen Nachwuchs im MINT-Bereich zu fördern. Mittlerweile liegen in Deutschland mehrere Studien (Guderian & Priemer, 2008; Pawek, 2012) vor, die dem Besuch von Schülerlaboren eine positive Wirkung auf Motivation und Lernleistung von Schülerinnen und Schülern im naturwissenschaftlichen Unterricht zusprechen. Weitgehend ungeklärt ist jedoch die Frage, welche Effekte eine Einbindung in den Schulunterricht im Vergleich zu den herkömmlichen Formen (Schülerlabor ohne Einbindung/nur Schulunterricht) hat.

Die vorliegende Studie untersucht die Wirkung der Einbindung eines Schülerlabors in den naturwissenschaftlichen Schulunterricht. Ziel ist es, mithilfe einer umfangreichen Stichprobe die Unterschiede in Motivation und Lernzuwachs von in vier unterschiedlichen Lehr-Lernarrangements unterrichteten Schülerinnen und Schülern zu untersuchen.

Es wurden drei Unterrichtsarrangements (Treatments) zu einem Thema der organischen Chemie entwickelt und im Schuljahr 2012/13 überprüft. In der ersten Treatmentgruppe wurde die Unterrichtseinheit im Schülerlabor ohne Einbindung unterrichtet, in der zweiten Treatmentgruppe wurde die Unterrichtseinheit nur im Schulunterricht behandelt, in der dritten Treatmentgruppe wurde der Schülerlaborbesuch in den Schulunterricht eingebunden. Mithilfe eines Pre-, Post-, Follow-Up-Designs wurden die drei Treatmentgruppen mit einer Wartekontrollgruppe verglichen. Die randomisiert auf die Gruppen aufgeteilte Stichprobe (68 neunte Realschulklassen) wurde mit strukturiertem Fragebögen und mehreren Leistungstests untersucht.

In Voranalysen wurde mittels Intraklassenkorrelationen die Bedeutung der Klassenebene (im verwendeten Design sind Schüler in Klassen genestet) analysiert und Hinweise darauf gefunden, dass ein mehrebenenanalytisches Auswertungsverfahren angezeigt war. Infolgedessen wurden die Unterschiede der vier Gruppen in der Motivation und der Lernleistung mit Mehrebenenregressionsanalysen geprüft.

Die bereits vorliegenden Befunde weisen darauf hin, dass die Einbindung des Schülerlabors in den Schulunterricht mit höherer Schülerleistung assoziiert ist als in den Vergleichsgruppen. Die Treatmentgruppe „Labor“ schnitt in Teilbereichen schlechter ab als die beiden Treatmentgruppen. Die Kontrollgruppe wies im Vergleich zu den Treatmentgruppen in mehreren Leistungstests die geringsten Werte auf. Im Vortrag werden die zentralen Ergebnisse der Studie vor- und zur Diskussion gestellt.

## **Mehrebenenanalysen zum Einfluss internetgestützter Hausaufgaben auf den Wissenserwerb beim Physiklernen**

Einleitung und Forschungsfrage : 50-60% der Schülerinnen und Schüler in der Sekundarstufe nutzen das World Wide Web (WWW) regelmäßig zum außerschulischen Lernen und insbesondere zur Bearbeitung ihrer Hausaufgaben (Feierabend et al., 2012; Allensbach-Studie, 2013). Grundsätzlich kann die Bearbeitung von Hausaufgaben die schulische Leistungsentwicklung unterstützen (vgl. Trautwein, 2008; vgl. Kohler, 2013). Diese Aussagen basieren aber größtenteils auf Untersuchungen in den Hauptfächern, dem Mathematik- und Fremdsprachenunterricht. Für das Fach Physik liegen keine empirische Daten vor; in Standardwerken der deutschsprachigen Physikdidaktik wird das Thema Hausaufgaben im Physikunterricht nicht diskutiert (vgl. Hopf et al., 2011; vgl. Kircher et al., 2009). Somit liegen in der Domäne Physik unter den Aspekten Hausaufgaben und Internet zwei unterschiedliche Fragen vor:

- 1) Haben Hausaufgaben im Physikunterricht – als Typus eines Nebenfachs – einen Einfluss auf den Wissenserwerb?
- 2) Lässt sich das WWW im Physikunterricht so einsetzen, dass Hausaufgaben ‚lernwirksam‘ werden?

Zur Beantwortung der zweiten Frage wurde im Schuljahr 2012/13 an Gymnasien im Bundesland Baden-Württemberg eine Interventionsstudie durchgeführt. Die Schülerinnen und Schüler bearbeiteten in einem Onlineportal Hausaufgaben, die von den Lehrpersonen an ihren regulären Wärmelehreunterricht adaptiert waren. Die Klassen der Treatmentgruppe erhielten als Intervention und Unterstützungsmaßnahme im Onlineportal neben den Aufgaben zusätzlich vorstrukturierte Linklisten. Diese Linklisten bestanden aus Internetverweisen auf Webseiten, die Informationen enthielten, um die Bearbeitung der Hausaufgaben zu unterstützen; Lösungen waren auf diesen Webseiten nicht zu finden. Die Linklisten waren vorstrukturiert: Sie führten zu fachlich richtigen und zu inhaltlich zielführenden Informationen. Damit wird ein Idealfall simuliert. Die Klassen der Kontrollgruppe erhielten die Linklisten nicht. Damit lautet die spezifizierte Forschungsfrage: Führen Physikhausaufgaben in einem Onlineportal, die mit einer strukturierten Verlinkung zum Internet versehen werden, zu einem größeren Wissenserwerb als ohne Verlinkung?

Design : Es liegt eine quasiexperimentelle Treatment-Kontrollgruppenstudie mit einer natürlichen hierarchischen Datenstruktur vor, die mit einer Mehrebenenregressionsanalyse ausgewertet wird. Die Intervention erfolgte in der Wärmelehre in Klassestufe 9 und dauerte durchschnittlich sieben bis acht Unterrichtswochen. Die Messung der abhängigen Variablen erfolgte an zwei Messzeitpunkten mit einem erprobten Fragebogen zu Basiskonzepten der Wärmelehre (vgl. Crossley, 2012). Verwendet wurde eine inhaltsvalide und reliable Skala mit 21 Items ( $\alpha_{PRE}=.712$ ,  $\alpha_{POST}=.739$ ). Die Bestimmung des Wissenszuwachses erfolgte über die Differenzbildung der PRE- und POST-Test-Scores. Zum Ausschluss von Alternativerklärungen wurden neben dem physikspezifischem Vorwissen u. a. folgende Variablen kontrolliert: Alter, Geschlecht, Intelligenz, Selbstkonzept in Physik, Interesse und Motivation an Physik (zsf. Crossley, 2012), sowie Variablen zur Internetnutzung und zum Umgang mit Hausaufgaben. Diese Variablen können in der Mehrebenenanalyse als Prädiktoren getestet werden.

Stichprobe: Die Stichprobe rekrutierte sich aus 14 Gymnasien. Eine Zufallsstichprobe liegt nicht vor. Die Stichprobe ist aber auf Klassenebene randomisiert und setzt sich aus N = 910 Schülerinnen und Schüler und 35 Klassen zusammen.

Ergebnisse : Bereits eine Pilotierung mit N = 113 Lernenden aus fünf Klassen wies auf die Lernwirksamkeit der Verwendung von Linklisten hin. Ein explorativ zu wertender ICCPILOT = 0.132 indiziert die Notwendigkeit der Mehrebenenanalyse. Die Intraklassenkorrelation der Hauptstudie ist mit ICCHAUPT = .075 geringer, aber nicht vernachlässigbar. Die Mehrebenenanalyse bestätigt die Ergebnisse der Pilotierung: Schülerinnen und Schüler, die die Linklisten als Unterstützungsmaßnahme beim Bearbeiten der Hausaufgaben verwendeten, erzielten durchschnittlich einen signifikant höheren Score im Post-Test und im Wissenszuwachs.

Zusammenfassung : Die Bearbeitung von Hausaufgaben im Physikunterricht mit Internetunterstützung kann den Wissenserwerb fördern. Die Ergebnisse der Mehrebenenanalyse zeigen, dass Linklisten als Unterstützungsmaßnahme beim Erlernen von Basiskonzepten in der Wärmelehre die Lernleistung der Schülerinnen und Schüler verbessern kann. Die Erstellung und Verwendung von Linklisten im Unterricht stellt eine Methode dar, die für Lehrkräfte einfach und mit nur geringem Aufwand umsetzbar ist. Zu klären bleibt jedoch, ob sich diese Ergebnisse auf andere Themengebiete der Physik bzw. auf andere Unterrichtsfächer übertragen lassen.

Sara Dallinger (Pädagogische Hochschule Ludwigsburg), Kathrin Jonkmann (Universität Tübingen), Jan Hollm (Pädagogische Hochschule Ludwigsburg)

## **Lernen bilingual unterrichtete Schüler mehr im Geschichtsunterricht als ihre einsprachig unterrichteten Mitschüler? – Ergebnisse der interdisziplinären Längsschnittstudie COMBIH**

Der bilinguale Sachfachunterricht ist bereits seit längerer Zeit ein fester Bestandteil des Unterrichtsangebots. Beispielsweise gibt es in Baden-Württemberg zurzeit 65 Gymnasien, an denen deutsch-englische Züge eingerichtet wurden und auch an Realschulen wird das bilinguale Unterrichtsangebot auf bis zu 20% aller Schulen ausgedehnt. Dennoch besteht nach wie vor Unklarheit über die Effekte bilingualen Lehren und Lernens auf die Kompetenzentwicklung der Schüler (Pérez-Cañado, 2012). Einige, häufig querschnittliche Studien konnten zeigen, dass der bilinguale Unterricht förderlich für die Englischkompetenzen der Lerner ist (Nold et al., 2008; Zydariß, 2007). Dieser Zusammenhang gilt auch dann, wenn berücksichtigt wird, dass es sich bei bilingual unterrichteten Schülern häufig um eine hinsichtlich der individuellen Lernvoraussetzungen positiv selektierte Gruppe handelt (Rumlich, 2013; Köller, Leucht, & Pant, 2012). Mit der vorliegenden Längsschnittstudie soll zum einen ein weiterer Beitrag zu dieser Forschung geleistet werden und zum anderen der Effekt des bilingualen Unterrichts auf die Kompetenzentwicklung in Englisch im Laufe eines Schuljahrs unter Kontrolle der Lernausgangslagen bestimmt werden. Weiterhin liegen zur Kompetenzentwicklung der Schüler in den bilingual unterrichteten Sachfächern kaum wissenschaftliche Erkenntnisse vor (Bonnet, 2012). So ist weitestgehend ungeklärt, welchen Effekt bilinguales Lehren und Lernen auf das Sachfach hat. Auch die Entwicklung der Schülermotivation im Fach Englisch und in den Sachfächern wurde bisher selten im Rahmen quantitativer Studien erhoben (Lasagabaster & Sierra, 2009). Diese Desiderata versucht das vorliegende interdisziplinäre Forschungsprojekt durch die Zusammenarbeit von Englisch- und Geschichtsdidaktikern sowie empirischen Bildungsforschern zu schließen.

Das Untersuchungsdesign besteht in einer Längsschnittstudie mit zwei Messzeitpunkten am Anfang und Ende der 8. Klasse. Die Stichprobe setzt sich aus 1663 Gymnasialschülern der 8. Klassenstufe zusammen. Hiervon wurden 664 Schüler aus 30 Klassen (28 Schulen) in Geschichte bilingual unterrichtet. Die einsprachig unterrichtete Kontrollgruppe besteht aus zwei verschiedenen Populationen. Zum einen sind dies 588 Schüler aus den Parallelklassen der bilingual unterrichteten Gruppen und zum anderen 411 Schüler von 9 Schulen (17 Klassen), die kein bilinguales Unterrichtsangebot besitzen. Die Schüler bearbeiteten Tests zu ihren kognitiven Grundfähigkeiten (figuraler und verbaler KFT), einen Test zur Erfassung der allgemeinen Englischkompetenz (C-Tests), einen Hörverstehenstest sowie einen Wissenstest zu Inhalten aus dem Geschichtsunterricht der Klassen 7 und 8 (Lückentexte, Multiple Choice Aufgaben). Außerdem füllten die Schüler einen umfangreichen Fragebogen aus mit Fragen zur Motivation und Leistungsemotionen in den Fächern Englisch und Geschichte, zu ihrem Unterrichtserleben und der Einschätzung ihrer Lehrkräfte in beiden Fächern sowie Fragen zu ihrer Person und dem Elternhaus. Darüber hinaus nahmen 71 der an der Studie beteiligten Geschichtslehrkräfte an einer schriftlichen Befragung teil. Sie beantworteten Fragen zu methodisch-didaktischen Aspekten des (bilingualen) Geschichtsunterrichts, ihrer Lehrerbiographie und zu arbeitsbezogenen Verhaltens- und Erlebensmustern. Zur Analyse der Daten wurden Mittelwertsvergleiche und Regressionsanalysen unter Berücksichtigung der hierarchischen Datenstruktur berechnet. Außerdem wurden Mehrebenenanalysen zur Bestimmung der Effekte auf Klassenebene durchgeführt.

Erste Analysen bestätigen positive Selektionseffekte zugunsten der bilingual unterrichteten Schüler hinsichtlich des familiären Bildungshintergrunds und der verbalen kognitiven Grundfähigkeiten. Des Weiteren zeigten sich erwartungsgemäß große Leistungsvorsprünge in den Englischkompetenzen zu Beginn der 8. Klasse, die sich unter Kontrolle dieser Selektionseffekte zwar etwas reduzierten, aber noch immer im Bereich einer dreiviertel Standardabweichung lagen. Zudem zeigten sich Vorteile der bilingual unterrichteten Schüler zu Beginn der 8. Klasse bezüglich des Wissenstests zu Geschichtsinhalten aus Klassenstufe 7 (in der Geschichte bei allen Schülern einsprachig Deutsch unterrichtet wurde). Diese Unterschiede waren jedoch unter Berücksichtigung der positiven Selektionseffekte nicht länger statistisch bedeutsam. Es bestanden keine Unterschiede zwischen den Gruppen hinsichtlich des Vorwissenstests zu Geschichtsinhalten der Klasse 8. Wie sich die Kompetenzen und die Motivation der bilingual und einsprachig unterrichteten Schüler in den Fächern Englisch und Geschichte im Verlauf der 8. Klasse unter Kontrolle dieser Eingangsunterschiede entwickeln, wird derzeit analysiert.

**GEBF** TAGUNG  
2014

E P10 c

Poster-

Symposiums-

beitrag

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823

2

Unterricht und schulbezogene Interventionen



Brigitte Schreier (Eberhard Karls Universität Tübingen), Anna-Lena Dicke (Eberhard Karls Universität Tübingen), Hanna Gaspard (Eberhard Karls Universität Tübingen), Isabelle Häfner (Eberhard Karls Universität Tübingen), Barbara Flunger (Eberhard Karls Universität Tübingen), Oliver Lüdtke (Humboldt-Universität zu Berlin), Benjamin Nagengast (Eberhard Karls Universität Tübingen), Ulrich Trautwein (Eberhard Karls Universität Tübingen)

## ***Der Zusammenhang relevanzbezogener Unterrichtsmerkmale mit den Wertüberzeugungen von Schülerinnen und Schülern in Mathematik***

Das Unterrichtsgeschehen ist zentral für die motivationale Entwicklung von Schülerinnen und Schülern (vgl. Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). Überblickbeiträge weisen darauf hin, dass die Lehrkraft Motivation fördern kann, indem sie Bezüge zwischen dem Lernmaterial und der Lebenswelt der Lernenden herstellt (z.B. Fredricks, Blumenfeld, & Paris, 2004; Wigfield et al., 2006), insbesondere im Fach Mathematik (vgl. Turner & Meyer, 2009). Auch die motivationale Haltung der Peers beeinflusst die Schülermotivation (Fredricks et al., 2004; Juvonen, Espinoza, & Knifsend, 2012).

Die Erwartungs-Wert-Theorie wird häufig als Grundlage zur Erfassung von Leistungsmotivation herangezogen (vgl. Eccles & Wigfield, 2002). Das Modell nimmt an, dass leistungsbezogenes Schülerverhalten durch vier subjektive Wertkomponenten—intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung—bestimmt wird. Das Verhalten von Bezugspersonen beeinflusst diese Wertüberzeugungen (Eccles & Wigfield, 2002), die sich laut Forschungsergebnissen im Laufe der Sekundarschulzeit negativ entwickeln (z.B. Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002).

Für das Fach Mathematik wurde gezeigt, dass intrinsischer Wert und Wichtigkeitsüberzeugung bei Sekundarschülerinnen und -schüler zunehmen, wenn die wahrgenommene Bedeutung des Lehrplans und der Unterrichtsmethoden für persönliche Ziele steigt (Wang, 2012). Außerdem verändert sich das Schülerinteresse, ein dem intrinsischen Wert verwandtes Konstrukt, in Abhängigkeit von der wahrgenommenen Wertschätzung des Fachs Mathematik durch die Klassenkameradinnen und -kameraden (Frenzel, Goetz, Pekrun, & Watt, 2010). Unklar bleibt, welche konkreten Unterrichtsstrategien sich zur Vermittlung der Relevanz mathematischer Lerninhalte eignen und welche Rolle die wahrgenommene Werteheftung der Klassenkameradinnen und -kameraden differenziell für die vier Wertüberzeugungen der Schülerinnen und Schüler in Mathematik spielt.

Ziel der vorliegenden Studie war es, unter Berücksichtigung von Schüler- und Lehrerperspektive den Zusammenhang zwischen relevanzbezogenen Unterrichtsmerkmalen und den vier Wertüberzeugungen von Sekundarschülerinnen und -schülern in Mathematik zu untersuchen. In Anlehnung an bisherige Forschung (z.B. Kunter & Baumert, 2006; Rakoczy, Klieme, & Pauli, 2008; Salmela-Aro, 2010) wurden aus Lehrersicht zwei Unterrichtsstrategien—die Themenerarbeitung durch Alltagsbeispiele und die Verbindung mathematischer Inhalte mit Inhalten anderer Fächer—als Prädiktoren herangezogen. Die Praxisorientierung des Mathematikunterrichts (vgl. Wang, 2012) sowie die wahrgenommene Wertschätzung des Fachs Mathematik in der Klasse (vgl. Frenzel et al., 2010) stellten die Prädiktoren aus Schülersicht dar.

An 25 Gymnasien wurden Fragebogendaten von 1868 Schülerinnen und Schülern der neunten Jahrgangsstufe und ihren 73 Mathematiklehrkräften erhoben. Zur detaillierten Erfassung der Wertüberzeugungen intrinsischer Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung wurde ein umfassendes Werterhebungsinstrument eingesetzt. Mithilfe von Mplus wurde für jede der vier Wertüberzeugungen als abhängige Variable ein lineares Regressionsmodell auf zwei Ebenen gerechnet. Die Prädiktoren aus Schülersicht wurden auf der Schülerebene sowie als latente Aggregate auf der Klassenebene ins Modell aufgenommen. Auf der Klassenebene wurden zusätzlich die Prädiktoren aus Lehrersicht ins Modell eingefügt. Es wurde für Schülergeschlecht und Mathematikleistung kontrolliert.

Multiple Regressionsanalysen zeigten, dass die Praxisorientierung im Mathematikunterricht und die wahrgenommene Wertschätzung der Mathematik in der Klasse auf beiden Ebenen positiv mit intrinsischem Wert, Wichtigkeits- und Nützlichkeitsüberzeugung und negativ mit der Kostenüberzeugung assoziiert waren. Auf der Schülerebene hatte die Praxisorientierung für alle Wertüberzeugungen eine größere Bedeutung als die wahrgenommene Wertschätzung der Mathematik in der Klasse. Auf Klassenebene sagte dahingegen die wahrgenommene Wertschätzung der Mathematik in der Klasse intrinsischen Wert, Wichtigkeits- und Kostenüberzeugung stärker vorher als die Praxisorientierung—nicht aber die Nützlichkeitsüberzeugung. Während die aus Lehrersicht berichtete Demonstration von inhaltlichen Verbindungen positiv mit Wichtigkeits- und Nützlichkeitsüberzeugung assoziiert war, fanden sich keine Zusammenhänge zwischen der Themeneinführung mit Alltagsbeispielen und den Wertüberzeugungen der Schülerinnen und Schüler.

Zusammenfassend spielen sowohl die individuelle Schülerwahrnehmung der Praxisorientierung und die Demonstration inhaltlicher Verbindungen im Mathematikunterricht als auch die wahrgenommene Haltung der Klassenkameradinnen und -kameraden zur Mathematik eine wichtige Rolle für intrinsischen Wert, Wichtigkeits-, Nützlichkeits- und Kostenüberzeugung von Schülerinnen und Schülern. Diese Studie liefert somit Anregungen zur Identifikation von Unterrichtsmaßnahmen, die dabei helfen könnten, dem Motivationsabfall von Schülerinnen und Schülern in Mathematik entgegenzuwirken.



Thorsten Luka (Universität Potsdam), Christian Jäntsch (Universität Potsdam), Jana Vogel (Universität Potsdam), Stefanie Bosse (Universität Potsdam), Nadine Spörer (Universität Potsdam)

## ***Beeinflusst die Klassengröße das Sozialverhalten? Eine empirische Überprüfung sozialpsychologischer Modelle***

Obgleich der Einfluss einer reduzierten Klassengröße auf die Leistungen von Schülerinnen und Schülern sowohl in seinem Ausmaß als auch im Zusammenspiel der zugrunde liegenden Prozesse gut erforscht ist (Finn, Pannozzo, & Achilles 2003; Hattie, 2008), ist über den Einfluss auf das Sozialverhalten wenig bekannt (Blatchford, 2012). Bisherige Studien konnten vor allem zeigen, dass eine Reduzierung der Klassengröße mit einem Rückgang antisozialen Verhaltens -- wie Regelverstößen und Störungen des Unterrichts -- in der Grundschule einhergeht (Finn et al. 2003). Einige Studien zeigten darüber hinaus, dass auch ein Anstieg prosozialen Verhaltens -- wie die Bereitschaft zur Kooperation und Verantwortungsübernahme -- zu beobachten ist (Finn et al. 2003).

Derzeit mangelt es den Studien zum Zusammenhang von Klassengröße und Sozialverhalten -- ebenso wie Untersuchungen zur Leistungsentwicklung -- an einer theoretischen Fundierung. Insbesondere fehlt es an Rahmenmodellen, die Vorhersagen über die differentiellen Wirkungen von kleinen und großen Klassen auf das Sozialverhalten zulassen (Blatchford, 2012). Zwar beziehen sich Blatchford (2012) sowie Finn et al. (2003) auf sozialpsychologische Theorien zur diffusion of responsibility und group cohesiveness, um differentieller Effekte kleiner Klassen auf das Lernen und das Sozialverhalten von Schülern zu erklären. Eine valide empirische Absicherung dieser Konzepte an großen Stichproben und auf Grundlage von Schülerbefragungen steht jedoch noch aus. Der hiesige Beitrag greift zwei dieser eng aufeinander bezogenen Theorien auf und untersucht folgende Fragestellungen: (a) Nimmt die Bereitschaft, andere in der Klasse zu unterstützen, mit sinkender Klassengröße zu (diffusion of responsibility) und (b) Wirkt sich eine Verkleinerung der Klassen positiv auf den Zusammenhalt der Klasse aus (group cohesiveness)?

Grundlage der Studie bilden die Untersuchungen zu Schüler-, Lehrer- und Unterrichtsmerkmalen im Rahmen des Pilotprojekts „Inklusive Grundschule“ des Landes Brandenburg, das seit dem Schuljahr 2012/13 von der Universität Potsdam wissenschaftlich begleitet wird. Ziel der als Längsschnittstudie angelegten Begleitforschung ist die Identifizierung von Gelingensbedingungen für gemeinsamen Unterricht von Kindern mit und ohne sonderpädagogischen Förderbedarf. Die untersuchte Stichprobe umfasst N = 1465 Schülerinnen (50%) und Schüler (50%) in 72 Klassen (53% 2. Klasse; 47% 3. Klasse). Hiervon haben 85 (6%) Kinder einen festgestellten sonderpädagogischen Förderbedarf.

Die Bereitschaft der Schülerinnen und Schüler einander zu unterstützen wurde mithilfe standardisierter Befragungen aus Lehrerperspektive (Skala Kooperation  $\alpha = .87$ , LSL, Petermann & Petermann, 2006) sowie aus Schülerperspektive (Skala Klassenklima  $\alpha = .74$ , FESS 3-4, Rauer & Schuck, 2003) erhoben. Der Zusammenhalt der Klasse wurde über einen Soziometriebogen zu Freundschaftsnetzwerken innerhalb der Klasse gemessen (HIFTO, Agard, Veldman, Kaufman & Semmel, 1978).

Die Daten werden zur Zeit mehrebenenanalytisch ausgewertet. Im Falle des Soziometriefragebogens wird zudem auf Verfahren der Netzwerkanalyse zurückgegriffen.

Agard, J. A., Veldman, D. J., Kaufman, M. J., & Semmel, M. I. (1978). How I Feel Towards Others. An instrument of the PRIME instrument battery. Unpublished manuscript, U.S. Department of Education, Office of Special Education.

Blatchford, P. (2012). Three generations of research on class-size effects. In K. R. Harris, S. Graham. & T. Urdan (Hrsg.), *The American Psychological Association (APA) Educational Psychology Handbook* (S. 529 - 554). Washington DC, USA: APA.

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The "why's" of class size. *Student behavior in small classes. Review of Educational Research*, 73 (3), 321-368.

Hattie, J. (2008). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, New York : Taylor & Francis.

Petermann, U., & Petermann, F. (2006). *Lehrereinschätzliste für Sozial- und Lernverhalten*. Göttingen: Hogrefe.

Rauer, W., & Schuck, K.-D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen*. Göttingen: Hogrefe.



## **Implementierung eines mobilen Lernarrangements zum Thema Fliegen im Sachunterricht – schulische Rahmenbedingungen für Nachhaltigkeit**

### 1. Begründung und Fragestellung

Die naturwissenschaftlich-technische Interessen- und Wissensbildung von Kindern und Lehrkräften im Primarbereich soll durch ein Angebot zum Thema Fliegen im Sachunterricht unterstützt werden. Das internationale Forschungsprojekt Innovation naturwissenschaftlich-technischer Bildung in Grundschulen der Region Bodensee (INTeB) beforscht den Einsatz eines sogenannten „mobilen Lernarrangements zum Thema Fliegen“ (finanziert von der Internationalen Bodensee Hochschule (IBH)). Das pädagogische Konzept besteht aus einer Fortbildung mit Coaching-Elementen zu einem mobilen Lernarrangement (Lernkisten für den Offenen Unterricht). Um eine nachhaltige Wirkung des Konzepts zu erkunden, wird in diesem Forschungsvorhaben folgenden Fragenstellungen nachgegangen:

- Einsatz des mobilen Lernarrangements in den Klassen: Wie sind die schulischen Rahmenbedingungen vor Ort?
- Aktuelle und weitere Unterrichtsgestaltung an den Schulen: Inwiefern fördern bzw. behindern Faktoren des Systems Schule den Einsatz und die Nachhaltigkeit des mobilen Lernarrangements?

Im Rahmen dieses Forschungsvorhabens werden personal-systemtheoretische Überlegungen mit einbezogen. Das Ziel ist die nachhaltige Wirkung des Forschungsprojektes zu unterstützen und die Forschungsergebnisse in Kooperation mit der Schulverwaltung, beispielsweise in Fortbildungsangeboten der Hochschulen und der Schulverwaltungen weiter zu verwenden.

### 2. Theoretischer Hintergrund

Die insgesamt vier Teilprojekte des Gesamtprojekts INTeB sind im adaptierten Angebot-Nutzungsmodell nach Helmke (2006) verortet. Dieses Forschungsvorhaben fokussiert den Bereich der Kontextbedingungen, da vom vorgefundenen Kontext die Unterrichtsgestaltung sowie der Unterrichtserfolg beeinflusst wird (vgl. Helmke, 2006). Ein Forschungsdefizit besteht derzeit in diesem Bereich, da Kontexte einen Einfluss haben, aber ihr Einfluss bisher kaum erforscht ist. Aus diesem Grund sind sie für Interventionsüberlegungen zu überprüfen. Die von Helmke eher allgemein formulierten Bereiche werden mittels der Personal Systemtheorie (König & Volmer, 2005) aufgeschlüsselt und theoretisch fundiert. Die Personale Systemtheorie sieht Schule als soziales System das von folgenden Faktoren (Kontextbedingungen) bestimmt wird: den relevanten Personen des sozialen Systems, den subjektiven Deutungen der betreffenden Personen, den sozialen Regeln, Regelkreisen, der Systemumwelt und der Entwicklung des Systems. Die Hauptthese der Personal Systemtheorie ist, dass die „Veränderung sozialer Systeme immer nur bedeuten [kann], dass sich Menschen Gedanken über ihre Situation machen, auf Basis dieser Deutungen handeln und damit das System verändern“ (König & Volmer, 2005, S. 32). Die schulischen Rahmenbedingungen werden von den Lehrkräften in spezifischer Weise wahrgenommen, beispielsweise durch ihre Sicht auf die Ausstattung der Schulen mit Materialien und Räumen.

### 3. Beschreibung des methodischen Vorgehens

Die personal-systemtheoretischen Bedingungen werden mittels Fragebogen und Interviews erfasst. Die Stichprobe besteht aus insgesamt 45 Lehrkräften, die in den drei beteiligten Ländern Deutschland, Österreich und Schweiz unterrichten. Auf der Grundlage von Fragebögen wurden leitfadengestützte Interviews, angelehnt an das Konstruktinterview nach König & Volmer, erstellt. Die Daten wurden an drei verschiedenen Zeitpunkten der Intervention erhoben. In den Interviews werden die Fortbildung sowie das Coaching ebenso thematisiert wie die Bedingungen für Innovationen an Primarschulen in den beteiligten Ländern. Die gewonnenen Daten aus den Interviews sind die Basis für eine Gruppendiskussion zum Ländervergleich. Die Auswertung der Interviews erfolgt unter Nutzung des Programms MAXQDA vor allem nach der Qualitativen Inhaltsanalyse nach Mayring (2010). Auf Basis der Personal Systemtheorie wurde ein Kategoriensystem deduktiv erstellt, welches durch induktiv gewonnene Kategorien bzw. Ausdifferenzierung vorhandener Kategorien ergänzt wurde. Nach den ersten Auswertungen zeigten sich bereits Tendenzen. Es scheint so zu sein, dass aus Sicht der Lehrkräfte beim Einsatz eines mobilen Lernarrangements im Unterricht u.a. die materielle Umwelt (bspw. die räumlichen Bedingungen) eine wesentliche Rolle spielt. In einer weiteren Detailanalyse wurden diese Einflussfaktoren weiter aufgeschlüsselt. Im Vortrag werden die Ergebnisse vorgestellt.

Helmke, A. (2006): Unterrichtsqualität und Lehrerprofessionalität. Seelze: Kallmeyer

König, E./ Volmer, G./ Bentler, A./ Bührmann, T. (2005): Systemisch denken und handeln. Personale Systemtheorie in Erwachsenenbildung und Organisationsberatung. 1. Auflage. Weinheim, Basel: Beltz

Mayring, P. (2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 9. Auflage. Weinheim: Deutscher Studien Verlag



Chair(s): Katrin Rakoczy (DIPF), Birgit Harks (DIPF)

Diskutant/in: Elmar Souvignier (Westfälische Wilhelms-Universität Münster)

Das geplante Symposium widmet sich zum einen der systematischen Aufarbeitung bisheriger Forschung zu formativem Assessment. Zum anderen werden eigene empirische Befunde zu folgenden Fragestellungen vorgestellt: Welche Faktoren tragen dazu bei, dass formatives Assessment von den Lehrkräften angemessen in den Unterricht implementiert wird? Wie wirken sich verschiedene Methoden formativen Assessments auf den Lehr-Lern-Prozess aus? Die vier Beiträge untersuchen die genannten Fragestellungen aus psychologischer (Hondrich et al. & Rakoczy et al.), erziehungswissenschaftlicher (Smit & Birri) und fachdidaktischer (Ropohl & Rönnebeck) Perspektive, wobei alle vier Beiträge auf interdisziplinär angelegten Projekten basieren. Elmar Souvignier wird als Experte für prozessorientierte Diagnostik und Förderung die Beiträge diskutieren.

Im Einzelnen befassen sich die vier Beiträge mit folgenden Fragestellungen:

Mathias Ropohl und Silke Rönnebeck stellen ein Literaturreview zu der Frage vor, welche Methoden formativen Assessments im Zusammenhang mit forschendem Lernen im naturwissenschaftlichen und technischen Unterricht empirisch untersucht wurden und wie diese Methoden eingesetzt wurden. Sie kommen zu dem Schluss, dass es kaum Studien gibt, die formatives Assessment in Kombination mit forschendem Lernen untersuchen und skizzieren anschließend ein entsprechendes Forschungsprojekt.

Lena Hondrich und Kolleg/innen untersuchen, inwiefern Lehrkräften die Umsetzung einer Intervention zum formativen Assessment im Sachunterricht der 3. Jahrgangsstufe gelingt und wovon das Gelingen abhängt. Sie zeigen, dass die Umsetzung durch geeignete Unterstützung durch Materialien gelingen kann und dass eine positive Einschätzung der Methode wichtig für erfolgreichen Transfer ist.

Robbert Smit und Thomas Birri untersuchen, inwiefern Beurteilungsraster als Teil einer formativen Beurteilung im Mathematikunterricht der 5. Jahrgangsstufe sowohl Lehrkräfte als auch Lernende unterstützen. Die Ergebnisse der vorzustellenden Pilotstudie zeigen, dass sowohl Lehrende als auch Lernende berichten, von den Beurteilungsrastern zu profitieren.

Katrin Rakoczy und Kolleg/innen untersuchen die Wirkung zweier Interventionen zum formativen Assessment im Mathematikunterricht der 9. Klasse. Sie finden, dass die Interventionen vermittelt darüber, wie Lernende die Interventionen wahrnehmen, positiv auf deren Interessenentwicklung wirken.

Die vier Beiträge fokussieren somit unterschiedliche Merkmale und Bedingungen, die für die Gestaltung, erfolgreiche Implementation und positive Wirkung von formativem Assessment berücksichtigt werden müssen.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823

## **Formatives Assessment beim forschenden Lernen in den Naturwissenschaften und in Technik – Ergebnisse einer Literaturrecherche**

### Theoretischer Hintergrund

Das forschende Lernen (inquiry-based learning) ist ein erprobter Unterrichtsansatz in den naturwissenschaftlichen Fächern (Furtak, Shavelson, Shemwell, & Figueroa, 2012). Im Rahmen des forschenden Lernens sollen Schülerinnen und Schüler in den Naturwissenschaften möglichst selbstständig Probleme diagnostizieren und diese mit naturwissenschaftlichen Methoden untersuchen (Linn, Davis, & Bell, 2004). Das Ziel ist das evidenzbasierte Ableiten von neuen Erkenntnissen und neuem Wissen aus Schülerperspektive (National Research Council, 1996). In Technik steht beim forschenden Lernen hingegen die Entwicklung eines gegenständlichen Produkts sowie dessen Funktionsprüfung im Vordergrund (International Technology Education Association, 1996).

Empirische Befunde belegen Vorzüge dieses Ansatzes gegenüber anderen Unterrichtskonzepten (Furtak, Seidel, Iverson, & Briggs, 2012). Gleichwohl benötigen Schülerinnen und Schüler Unterstützung beim forschenden Lernen, um einen zunehmenden Grad an Selbstständigkeit zu erlangen (Hmelo-Silver, Duncan, & Chinn, 2007; Martin-Hansen, 2002). Insbesondere der prozesshafte Charakter, der vielfältige fachspezifische sowie fachübergreifende Kompetenzen erfordert, gilt als Herausforderung. Eine methodische Möglichkeit, selbstständiges Lernen zu unterstützen, bietet formatives Assessment (z. B. Black & Wiliam, 1998; Kingston & Brooke, 2011).

Fragestellung: Im Rahmen des von der EU geförderten Projekts ‚Assess Inquiry in Science, Technology, and Mathematics Education‘ (ASSIST-ME) werden daher Methoden des formativen Assessments entwickelt und in Unterrichtseinheiten, die dem Ansatz des forschenden Lernens folgen, implementiert. Die Entwicklung der Methoden soll den internationalen Forschungsstand in diesem Bereich berücksichtigen. Daher war das erste Ziel anhand einer umfassenden Literaturrecherche einen Überblick über bisherige empirische Befunde zum formativen Assessment beim forschenden Lernen in der Primarstufe sowie den Sekundarstufen I und II zu erlangen. Drei Fragen waren Ausgangspunkte der Bestandsaufnahme:

- 1) Welche Aspekte von forschendem Lernen sind empirisch untersucht worden?
- 2) Welche Methoden formativen Assessments sind im Zusammenhang mit forschendem Lernen empirisch untersucht worden?
- 3) Wie wurden diese Methoden eingesetzt?

Methode: Zur Beantwortung der drei Fragen wurde eine Literaturrecherche anhand von drei Suchstrategien durchgeführt. Zunächst wurden anhand von Schlüsselwörtern Suchen in den Datenbanken ‚Web of Science‘ und ‚ERIC‘ realisiert. Anschließend wurde zusätzlich in relevanten Zeitschriften (z. B. Journal of Research in Science Teaching, International Journal of Technology and Design Education) nach weiteren themenbezogenen Publikationen recherchiert. Dritte Suchstrategie war die Suche von Literatur in Literaturverzeichnissen nach dem Schneeballprinzip. Ferner wurde zur Validierung und Ergänzung der durch die drei Suchstrategien erzielten Ergebnisse eine Expertenbefragung durchgeführt. Die gefundenen Publikationen wurden nach dem Fokus der Untersuchung sortiert und ausgewählt. Publikationen, die nicht den Zielen der Recherche entsprechen, wurden ausgeschlossen. Dazu zählen zum Beispiel rein theoretische Publikationen und Publikationen mit Bezug zum Universitätsniveau. Auf diese Weise konnten 161 relevante Publikationen für die Naturwissenschaften und Technik identifiziert werden. Diese wurden anhand eines Kategoriensystems analysiert. Dabei standen die Aspekte des forschenden Lernens sowie die untersuchten und eingesetzten Methoden des formativen Assessments im Vordergrund.

Ergebnisse: Gerade in den naturwissenschaftlichen Fächern konnte mit N = 148 eine zufriedenstellende Anzahl an Publikationen gefunden werden. In Technik ist die gefundene Anzahl an Publikationen mit N = 13 hingegen sehr gering. Insgesamt konnten anhand der Publikationen 15 verschiedene Aspekte von forschendem Lernen identifiziert werden (z. B. Suchen nach Informationen, Planen einer Untersuchung). Ein Schwerpunkt lag auf den beiden Aspekten ‚Argumentieren‘ und ‚Kommunizieren/Debattieren‘. Im Hinblick auf die verwendeten Methoden kann festgehalten werden, dass vor allem schriftlichen Aufgaben, Mapping-Verfahren und Laborjournale eingesetzt bzw. untersucht wurden. Diese Methoden wurden in der Regel summativ eingesetzt, so dass trotz der intendierten Kombination von forschendem Lernen und formativem Assessment kaum Studien mit Fokus auf dieser Kombination analysiert werden konnten.

Aufbauend auf den Ergebnissen wird im Vortrag abschließend ein Forschungsprojekt skizziert, das den Zusammenhang zwischen einer Lehrerfortbildung und deren Wirkung auf den Einsatz von Methoden des formativen Assessments im Unterricht untersucht.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Review of Educational Research*, 82(3), 300-329.

Furtak, E. M., Shavelson, R. J., Shemwell, J. T., & Figueroa, M. (2012). To teach or not to teach through inquiry: Is that the question? In S. M. Carver & J. Shrager (Eds.), *The journey from child to scientist. Integrating cognitive development and the education sciences* (1st ed., pp. 227-244). Washington, D.C.: American Psychological Association.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.

International Technology Education Association. (1996). *Technology for all Americans: A Rationale and Structure for the Study of Technology*. Verfügbar unter [http://www.iteea.org/TAA/PDFs/Taa\\_RandS.pdf](http://www.iteea.org/TAA/PDFs/Taa_RandS.pdf)

Kingston, N., & Brooke, N. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28-37.

Linn, M. C., Davis, E. A., & Bell, P. (Eds.). (2004). *Internet environments for science education*. Mahwah: Lawrence Erlbaum Associates Publishers.

Martin-Hansen, L. (2002). Defining Inquiry: Exploring the many types of inquiry in the science classroom. *The Science Teacher*, 69(2), 34-37.

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: The National Academies Press.





## Rubrics als Werkzeug für eine standardorientierte und formative Beurteilung

Einführung: Ziel der präsentierten Studie ist die Untersuchung der unterstützenden Wirkung eines Beurteilungsrasters (Rubric) als Teil einer formativen Beurteilung auf den Lernerfolg von Schülerinnen und Schülern. Rubrics sind definiert als Dokumente, welche die Erwartungen bei einer Beurteilung anhand von Kriterien oder Merkmalen auflisten und Stufen von exzellent bis ungenügend beschreiben (Andrade, Du, & Wang, 2008). Sie sind ein nützliches Werkzeug um Unterrichtslektionen an Standards auszurichten (Burke, 2006). Das in diesem Projekt eingesetzte Beurteilungsraster gehört zum Inhalt „Argumentieren und Begründen in der Mathematik“ und passt zu den Schweizer Bildungsstandards für Mathematik (Schweizerische Konferenz der Erziehungsdirektoren (EDK), 2011).

Theoretischer Hintergrund: In allen drei deutschsprachigen Ländern wurden im letzten Jahrzehnt nationale Bildungsstandards eingeführt. Nebst der Funktion des Bildungsmonitoring sollen die Standards auch zu „einem Motor der pädagogischen Entwicklung“ (Klieme et al., 2003) in den Schulen werden. Offen bleibt jedoch in der Expertise von Klieme et al., wie Lehrpersonen ihren Unterricht angepasst an die nationalen Standards gestalten sollen. Während grossflächige Leistungstests zwar hilfreich für die Überprüfung des Bildungssystems sind, bringen sie wenig für den Unterricht in der Klasse (Popham, 2003; Smit, 2008). Ursache ist der eher generelle Charakter solcher Tests, die viele Themen abdecken müssen. Beurteilungsraster sind viel näher am Unterrichtsgeschehen angesiedelt. Sie helfen den Lehrpersonen komplexere Kompetenzen in authentischen, handlungsorientierten Formen, wie Aufsätzen, mündlichen Präsentationen oder naturwissenschaftlichen Experimenten zu beurteilen (Popham, 1997). Andrade, Du & Wang (2008) konnten nachweisen, dass Schüler und Schülerinnen, welche einen Rubric für die Selbstbeurteilung von ersten Entwürfen eines Aufsatzes benutzten, die Qualität ihrer Arbeit wirksam steigern konnten (partiell  $\eta^2 = .15$ ). Ein Beurteilungsraster unterstützt sowohl das Feedback der Lehrperson, wie auch die Peer- und Selbstbeurteilung durch die Lernenden (Saddler & Andrade, 2004). Standardbasierte Rubrics helfen den Lehrpersonen den Lernenden die Kriterien für die Kompetenzerreichung transparent zu machen (Ross, McDougall, Hogaboam-Gray, & LeSage, 2003). Forschungsergebnisse liegen vor allem für die Sprachkompetenz vor. Aus diesem Grunde liegt der Fokus unserer Studie im Bereich Mathematik und auf der handlungsnahen Kompetenz „Argumentieren und Begründen“.

Unsere Forschungsfragen für die Pilotstudie lauteten folgendermassen:

- Unterstützen Rubrics die Lehrperson beim Planen, Umsetzen und Evaluieren eines kompetenzorientierten Unterrichts?
- Helfen Rubrics den Schüler/-innen kompetenzorientierte Lernziele und Aufgaben zu verstehen?
- Helfen Rubrics der Lehrperson und den Schüler/-innen den Kompetenzstand und Lernfortschritt zu bestimmen?
- Unterstützen Rubrics das Erteilen von formativem Feedback durch die Lehrperson?

Methode, Ergebnisse und Diskussion: Da unser Forschungsprojekt einen Pilotcharakter aufwies, beschränkten wir uns auf eine kleine Stichprobe von vier fünften Klassen ( $n = 67$ ) aus zwei Schweizer Kantonen. Wir benutzten einerseits quantitative Methoden um den Kompetenzstand der Schüler/-innen im Bereich mathematisches Argumentieren am Ende des Projektes zu messen. Andererseits setzten wir Interviews ein um die vier Lehrpersonen und je eine kleine Schülergruppe zu befragen hinsichtlich des Erfolgs unseres didaktischen Settings mit integrierter formativer Beurteilung und einem standardorientierten Rubric als ein unterstützendes Werkzeug. Der Leistungstest bestand aus 13 offenen Items, welche alle aus dem HarmoS-Mathematiktest (Schweizerische Konferenz der Erziehungsdirektoren (EDK), 2004) stammen und teilweise adaptiert wurden. Mittels Rasch-Modell konnten wir die vier Stufen des Rubrics validieren. Da die Schüler/-innen aus zeitlichen Gründen nur jeweils fünf offene Items bearbeiten konnten, ist die Reliabilität der Personen-Kompetenzwerte dagegen eher tief. Die Ergebnisse der qualitativen Inhaltsanalyse weisen auf eine eher differentielle Entwicklung der Argumentationskompetenz im Verlauf der Projektzeit von einem Monat hin. Ein grundlegendes Verständnis, dass Argumentieren aus mehr als richtigem Rechnen besteht, zeigte sich bei allen Schüler/-innen. Sie lernten sich mittels Rubric selbst zu beurteilen und die Lehrpersonen gaben an, das Werkzeug bei Rückmeldungen als unterstützend zu empfinden. Wichtig ist, so die Lehrpersonen, eine Illustration der Kompetenzstufen anhand von Beispielen. Im Hauptprojekt ist geplant, die Wirkung des Rubrics innerhalb des didaktischen Settings mittels eines Kontrollgruppendesigns zu überprüfen.

Andrade, H., Du, Y., & Wang, X. (2008). Putting Rubrics to the Test: The Effect of a Model, Criteria Generation, and Rubric-Referenced Self-Assessment on Elementary School Students' Writing. *Educational Measurement: Issues and Practice*, 3-13.

Burke, K. (2006). *From standards to rubrics in 6 steps*. Heatherton, Victoria: Hawker Brownlow Education.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., . . . Vollmer, H.J. (2003). *Zur Entwicklung nationaler Bildungsstandards : eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).

Popham, W.J. (1997). What's Wrong - and What's Right - with Rubrics. *Educational Leadership*, 10(10), 72-75.

Popham, W.J. (2003). The seductive allure of data. *Educational Leadership*, 60, 48-52.

Ross, J.A., McDougall, D., Hogaboam-Gray, A., & LeSage, A. (2003). A Survey Measuring Elementary Teachers' Implementation of Standards-Based Mathematics Teaching. *Journal for Research in Mathematics Education*, 34(4), 344-363.

Saddler, B., & Andrade, H. (2004). The Writing Rubric. *Educational Leadership*, 62(2), 48-52.

Schweizerische Konferenz der Erziehungsdirektoren (EDK). (2004). *Harmos – Zielsetzungen und Konzeption des Projektes*. Bern: EDK.

Schweizerische Konferenz der Erziehungsdirektoren (EDK). (2011). *Grundkompetenzen für die Mathematik*. Nationale Bildungsstandards. Retrieved 23.10.13 [http://educod.ch/record/96784/files/grundkomp\\_math\\_d.pdf](http://educod.ch/record/96784/files/grundkomp_math_d.pdf)

Smit, R. (2008). Formative Beurteilung im kompetenz- und standardorientierten Unterricht. *Beiträge zur Lehrerbildung*, 26(3), 383-392.



## **Formatives Assessment im Mathematikunterricht: Effekte auf Wahrnehmung und Interesse von Lernenden**

Theoretischer Hintergrund & Fragestellung: Formatives Assessment gilt als eine der effektivsten Ansätze, um Unterricht und Lernen zu verbessern (z.B. Black & Wiliam, 1998). Methoden der Leistungsbeurteilung sind nach Black und Wiliam (2009, S. 6) dann formativ, wenn „... evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited“. Damit Informationen über Schülerleistungen, wie bei Black und Wiliam beschrieben, von Lernenden interpretiert und genutzt werden können, um ihren weiteren Lernprozess zu optimieren, müssen diese Informationen den Lernenden zur Verfügung gestellt werden, d.h. an sie zurückgemeldet werden. Somit stellt Leistungsrückmeldung neben Leistungsbeurteilung ein zentrales Element von formativem Assessment dar (Sadler, 2011). Rückmeldung hat sich in bisheriger Forschung insbesondere dann als interessen- und leistungsförderlich erwiesen, wenn sie Lernende über ihren Lernstand im Verhältnis zum Lernziel informiert und ggf. die zur Aufgabenlösung notwendigen kognitiven Prozesse bzw. entsprechende Strategien rückmeldet (Hattie & Timperley, 2007). Dabei sollte sie sich auf konkrete Aufgaben und Lösungsprozesse beziehen und einen individuellen Bezugsrahmen anlegen. Eine schriftliche Rückmeldevariante für Mathematikaufgaben, die diesen Kriterien genügt (lösungsprozessbezogene Rückmeldung), wurde im DFG-Projekt Co2CA (Conditions and Consequences of Classroom Assessment) entwickelt. Sie wird von Lernenden unter Laborbedingungen als kompetenzunterstützender und nützlicher wahrgenommen als die im Unterrichtsalltag weit verbreitete Rückmeldung durch Noten und zeigt über diese Wahrnehmung vermittelt indirekte Wirkung auf die Interessen- und Leistungsentwicklung von Lernenden (Harks, Rakoczy, Hattie, Besser, & Klieme, 2013; Rakoczy, Harks, Klieme, Blum, & Hochweber, 2013). Anknüpfend an diese Befunde wurde im Projekt Co2CA lösungsprozessbezogene Rückmeldung im Rahmen zweier Interventionen zum formativen Assessment in den Mathematikunterricht implementiert. Ziel des vorliegenden Beitrags ist es, die Wirksamkeit dieser Interventionen im Hinblick auf die Wahrnehmung der Lernenden und ihre Interessenentwicklung zu prüfen.

Methode: Es handelt sich um eine Interventionsstudie mit Prä- und Posttest, die im Schuljahr 2010/11 in der Unterrichteinheit „Satz des Pythagoras“ durchgeführt wurde. Es wurden zwei unterschiedliche Varianten formativen Assessments mit einer Kontrollgruppe verglichen. Die verschiedenen Bedingungen wurden durch Fortbildungen der Lehrkräfte realisiert. In der Kontrollgruppe wurden den Lehrkräften die fachlichen Unterrichtsinhalte und zu behandelnden Aufgaben vermittelt. In der Interventionsgruppe 1, der „lösungsprozessbezogenen“ Bedingung, wurden die Lehrkräfte darüber hinaus darin geschult, zu bestimmten festgelegten Zeitpunkten die Leistung der Lernenden anhand von Diagnosebögen zu erfassen und lösungsprozessbezogene schriftliche Rückmeldung zu geben. In der Interventionsbedingung 2, der „lernprozessbegleitenden“ Bedingung, wurden die Lehrkräfte darüber hinaus darin trainiert, verbale Rückmeldung, die adaptiv auf die individuellen Bedürfnisse der Lernenden zugeschnitten ist, zu geben. An der Studie nahmen 41 Lehrkräfte mit ihren Klassen (ca. 1200 Lernende) teil und wurden zufällig auf die drei Bedingungen verteilt. Direkt vor und nach den Unterrichtsstunden, in denen die Interventionen durchgeführt wurden, wurde die Wahrnehmung und das Interesse der Lernenden per Fragebogen erhoben und die Mathematikleistung getestet.

Ergebnisse & Diskussion: Mehrebenenpfadanalysen zeigen, dass Lernende die Rückmeldung im Unterricht in beiden Interventionsbedingungen als kompetenzunterstützender, in der lösungsprozessbezogenen Bedingung darüber hinaus als nützlicher wahrnehmen, und dass diese Wahrnehmungen zum Teil positiv mit der Interessenentwicklung zusammenhängen. Die lösungsprozessbezogene Intervention wirkt sich indirekt, d.h. vermittelt über die wahrgenommene Nützlichkeit, positiv auf die Interessenentwicklung aus. Die Ergebnisse zeigen somit, dass Fortbildungen zu formativem Assessment auf Basis schriftlicher lösungsprozessbezogener Rückmeldung dazu führen können, dass Lernende die Rückmeldepraxis im Unterricht positiver beurteilen, was zum Teil zu einer positiveren Interessenentwicklung führt. Inwiefern die Interventionen für Leistungsbeurteilung und Rückmeldung im Unterricht nutzbar gemacht werden können, wird im geplanten Beitrag diskutiert.

Black, P. & Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-148.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31. doi:10.1007/s11092-008-9068-5

Harks, B., Rakoczy, K., Hattie, J., Besser, M. & Klieme, E. (2013). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology*, 1-22. doi:10.1080/01443410.2013.785384

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81e112. doi:10.3102/003465430298487 .

Rakoczy, K., Harks, B., Klieme, E., Blum, W. & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction*, 27, 63-73. doi:10.1016/j.learninstruc.2013.03.002

Sadler, D.R. (2010). Beyond feedback: developing student capability in complex appraisal, *Assessment & Evaluation in Higher Education*, 35, 535-550. doi: 10.1080/02602930903541015



## ***The interplay between individual school self-concept, socio-motivational support and achievement motivation amongst adolescent students***

Theoretical Background: Works of several authors state that adolescence is frequently marked by a decline in students' achievement motivation (Ladd, Herald-Brown & Kochel, 2009; Dohn, 1991; Finn, 1989, 2006), which in turn is often associated with a decline in school self-concept (Areepattamannil, 2012). The development of academic skills and thus the individual school self-concept is bolstered by achievement motivation, which can be defined as an aggregate of achievement drive, perseverance and effort and fear of success. However, less is known about the mediating role of socio-motivational support in the association between achievement motivation and individual school self-concept. According to Wentzel and colleagues (Wentzel, Battle, Russell & Looney, 2010; Wentzel, 1998), positive social relationships at school support achievement motivation. Based on the increasingly complex nature of social relationships during adolescence (Bukowski, Simard, Dubois & Lopez, 2011) both teacher-student and student-student relationships become essential for students' achievement motivation (Wentzel, 2009; Wentzel et al., 2010; Harter, 1996; Flanagan, Erath & Bierman, 2008; Raufelder & Mohr, 2011; Wentzel, 1998). In order to enhance our understanding of the abovementioned constructs (school self-concept, socio-motivational support and achievement motivation) and the ways in which they are related, we have examined whether socio-motivational support can be an effective starting point to strengthen students' motivation and individual school self-concept.

Research Questions: Are the individual school self-concept, variables of socio-motivational support and achievement motivation associated and, if yes, does socio-motivational support from peers and teachers function as a mediator in the association between individual school self-concept and achievement motivation?

Method: The current study examined the interplay of individual school self-concept, socio-motivational support (teacher-student relationship, student-student relationship, teachers as positive motivators, peers as positive motivators) and achievement motivation in a large sample of 7th and 8th grade students (N = 1088; MAge = 13.7) in secondary schools in Brandenburg, Germany. The data was based on self-report questionnaires. To test for associations between individual school self-concept, social support and achievement motivation structural equation modelling (SEM) was applied.

Results: The results showed that the teacher-student relationship as well as "teachers as positive motivators" partially mediated the association between individual school self-concept and achievement motivation. In contrast, neither "peers as positive motivators" nor the student-student relationship mediated this association. These results support the notion that maintaining a positive teacher-student relationship as well as encouraging teachers in the role of positive motivators could be an effective starting point for prevention and intervention programs aimed at offsetting the decline in individual school self-concept and achievement motivation during adolescence. While the teacher's role ("teacher-student relationship" and "teacher as positive motivators") seems to be more institutionalized and uniform and therefore connected to achievement motivation and the individual school self-concept, the role of peers should be understood in a more differentiated way: while a high quality of the student-student relationship might be good for students' well-being (Wentzel 2009) and reduce fear of success, it can also have a negative impact on achievement drive. However, when peers serve a more institutional and academic role as positive motivators, adolescents benefit from these relationships through increased achievement motivation (Raufelder, Drury, Jagenow, Hoferichter & Bukowski, 2013). In summary, our findings extend current research by distinguishing between types of students' relationships that can differentially affect their achievement motivation.

- Areepattamannil, S. (2012). Mediation Role of Academic Motivation in the Association between School Self-Concept and School Achievement among Indian Adolescents in Canada and India *Social Psychology of Education: An International Journal*, 15, 367–386.
- Dohn, H. (1991). "Drop-out" in the Danish high school (gymnasium): An investigation of psychological, sociological and pedagogical factors. *International Review of Education*, 37, 415–428.
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59, 117–142. doi: 10.3102/00346543059002117
- Finn, J. D., (2006). The adult lives of at-risk students: The roles of attainment and engagement in high school. Report to National Center of Educational Statistics, 328, Washington, DC: U. S.
- Flanagan, K. S., Erath, S. A., & Bierman, K. B. (2008). Unique Associations between Peer Relationships and Social Anxiety in Early Adolescence. *Journal of Clinical Child and Adolescent Psychology*, 37, 759–769. doi: 10.1080/15374410802359700
- Harter, S. (1996). Teacher and classmate influences on scholastic motivation, self-esteem, and level of voice in adolescents. In J. Juvonen & K. Wentzel (Eds.), *Social Motivation – Understanding children's school adjustment*, (pp. 11–42). Cambridge: University Press.
- Ladd, G. W., Herald-Brown, S. L., & Kochel, K. P. (2009). Peers and motivation. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 323–348). New York, NY: Routledge.
- Raufelder, D. & Mohr, S. (2011). Zur Bedeutung sozio-emotionaler Faktoren im Kontext Schule unter Berücksichtigung neurowissenschaftlicher Aspekte. In Ittel, A., Merkmens, H., Stecher, L., & Zinnecker, J. (Hrsg.), *Jahrbuch Jugendforschung* (pp. 74–96). Wiesbaden: VS Verlag.
- Raufelder, D., Drury, K., Jagenow, D., Hoferichter, F. & Bukowski, W. (2013). Development and Validation of the Relationship and Motivation (REMO) scales to assess students' perceptions of peers and teachers as motivators in adolescence. *Learning and Individual Differences* 23, 182–189. doi: 10.1016/j.lindif.2013.01.001
- Wentzel, K. R. (1998). Social relationships and motivation in middle school: the role of parents, teachers and peers. *Journal of Educational Psychology*, 90, 202–209. doi: 10.1037/0022-0663.90.2.202
- Wentzel, K. R. (2009). Students' Relationships with Teachers as Motivational Contexts. In K. R. Wentzel, & A. Wigfield (Eds.), *Handbook of Motivation at School* (pp. 301–322). New York, NY: Routledge.
- Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35 193–202. doi: 10.1016/j.cedpsych.2010.03.002



## ***Einfluss einer bedeutsamen Schülerwahl auf die Motivation***

Theoretischer Hintergrund: Gemäß der Selbstbestimmungstheorie der Motivation von Deci und Ryan (2000) ist die Entstehung motivierten Verhaltens auf die Erfüllung der Grundbedürfnisse nach sozialer Einbindung, Kompetenz und Selbstbestimmung zurückzuführen. Autonomie bildet das Kernkonzept in der Selbstbestimmungstheorie (Ryan & Deci, 2006). Die Bedeutung des Autonomieerlebens für den Menschen wurde in zahlreichen Studien gezeigt (Ryan & Deci, 2006), auch für den schulischen Kontext (Reeve, 2002). Reeve, Hamm und Nix (2002) haben drei Aspekte von Selbstbestimmung identifiziert: Ort der Handlungsverursachung (perceived locus of causality), Volition (volition) und wahrgenommene Wahl (perceived choice). Insbesondere der Aspekt perceived choice wurde in vielen Studien untersucht, dabei zeigt sich häufig ein uneinheitliches Bild (Ryan & Deci, 2006). Mit den Auswirkungen von Choice im schulischen Kontext haben sich Katz und Assor (2007) befasst. Die von ihnen begutachteten Studien zeigen ebenfalls widersprüchliche Ergebnisse. Unter dem Blickwinkel der Selbstbestimmungstheorie betrachtet, zeigen sich allerdings einheitliche Ergebnisse: Choice hat positive Auswirkungen, wenn die drei Grundbedürfnisse des Menschen (Autonomie, Kompetenz, soziale Eingebundenheit) berücksichtigt werden. Optionen sollten bedeutsam für die Schülerinnen und Schüler sein (Autonomie), nicht zu zahlreich und unverständlich (Kompetenz) und nicht im Widerspruch zum sozialen Umfeld (soziale Eingebundenheit) stehen (Katz & Assor, 2007). Damit das Autonomieerleben durch eine Wahl positiv beeinflusst wird, sollte wenigstens eine Option als relevant oder interessant wahrgenommen werden (Katz & Assor, 2007). Können Wählende nur zwischen negativ besetzten Optionen eine Entscheidung treffen, werden sie ihre Wahlfreiheit als geringer einschätzen im Vergleich zu Wählenden mit positiv wahrgenommenen Optionen (Williams, 1998).

Fragestellung: Die Studie soll die Frage klären, ob Wahlmöglichkeiten in der Schule eine positive Auswirkung auf die Motivation haben auch wenn die Optionen als wenig interessant eingeschätzt werden oder ob wenigstens eine Option als interessant erachtet werden muss. Gleichzeitig wird überprüft, ob die Interessantheit des Themas allein, ohne Schülerwahl, zu einer erhöhten intrinsischen Motivation führen kann.

Methode: Untersucht wurden 12 Klassen der 5. Jahrgangsstufe mit einem Durchschnittsalter von 11.07 (SD: 0.42) Jahren. Die zwölf Klassen wurden gleichmäßig auf vier Gruppen verteilt, die sich durch zwei Faktoren unterschieden (1. Wahl: Wahl, keine Wahl und 2. Thema: sehr interessantes Thema, wenig interessantes Thema). Zu Beginn der Studie wählten die Klassen der Wahlbedingung das Unterrichtsthema. Es folgten drei Biologiestunden zu dem gewählten Thema. Nach der letzten Stunde wurde die intrinsische Motivation mit einer übersetzten und adaptierten Form des IMI (Intrinsic Motivation Inventory) von Deci und Ryan (2011) gemessen.

Ergebnisse: Es zeigte sich, dass eine Wahl zwischen Themen unterschiedlicher Interessantheit den Aufbau intrinsischer Motivation im Unterricht deutlich förderte. Unterricht zu einem interessanten Thema ohne Wahlmöglichkeit hatte marginal positive Effekte. Die Wahl zwischen gleich wenig interessanten Themen hatte keinen positiven Einfluss auf die Motivation im Unterricht. Ein Vergleich der Gruppe mit bedeutsamer Wahlmöglichkeit (W+/I+) mit der Kontrollgruppe (W-/I+), die denselben Unterricht ohne vorherige Wahlmöglichkeit erhielt zeigt signifikant höhere Werte in den Subskalen Interesse/Vergnügen, wahrgenommene Kompetenz und wahrgenommene Wahlfreiheit. Ein Vergleich der Gruppe „W+/I+“ (Wahl mit unterschiedlich interessanten Themen) und Gruppe „W+/I-“ (Wahl mit gleich wenig interessanten Themen) zeigte statistisch bedeutsame Unterschiede in den Skalen Interesse/Vergnügen, wahrgenommene Kompetenz und wahrgenommene Wahlfreiheit. Ein Vergleich der Gruppen „W+/I-“ (Wahl mit gleich wenig interessanten Optionen) und „W-/I-“ (keine Wahl), die Unterricht zu demselben wenig interessanten Thema hatten, zeigte keine Unterschiede in den Subskalen des Intrinsic Motivation Inventory. Hier zeigt sich, dass Schülerwahl ohne attraktive Wahlmöglichkeiten keinen positiven Effekt auf die Schülermotivation hatte. Ein Vergleich der Gruppe „W-/I+“ (keine Wahl, sehr interessantes Thema) mit der Gruppe „W-/I-“ (keine Wahl, wenig interessantes Thema) zeigte lediglich eine Tendenz in der Subskala Interesse/Vergnügen. Im vorliegenden Design hatte Unterricht zu einem interessanten Thema allein im Vergleich zu Unterricht zu einem wenig interessanten Thema kaum positive Auswirkungen auf die Motivation.

Deci, E. L. & Ryan, R. M. (2000): The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behaviour. *Psychological Inquiry*, 11(4), 227-268.

Deci, E. L. & Ryan, R. M. (2011). Intrinsic Motivation Inventory (IMI). Verfügbar unter: <http://www.psych.rochester.edu/SDT/measures/intrins.html> [01.01.2011].

Katz, I. & Assor, A. (2007). When Choice Motivates and When It Does Not. *Educational Psychology Review*, 19(4), 429-442.

Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (S. 183-203). Rochester, NY: University Of Rochester Press.

Reeve, J., Hamm, D. & Nix, G. (2003). Testing Models of the Experience of Self-Determination in Intrinsic Motivation and the Conundrum of Choice. *Journal of educational Psychology*, 95(2), 375-392.

Ryan, R. M. & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25, 54-67.

Ryan, R. M. & Deci, E. L. (2006). Self-Regulation and the Problem of Human Autonomy: Does Psychology Need Choice, Self-Determination, and Will? *Journal of Personality* 74(6), 1557-1586.

Williams, S. (1998). An Organizational Model of Choice: A Theoretical Analysis Differentiating Choice, Personal Control, and Self-Determination. *Genetic, Social, and General Psychology Monographs*, 124 (4), 465-491.



**Selbstregulation und Lernerfolg Chemiestudierender**

## Postergruppe

Theoretischer Hintergrund: Studierende der Naturwissenschaften tendieren besonders stark dazu, ihr Studium vorzeitig abzubrechen. Solche Studierende, die ein Hochschulstudium beginnen und die Hochschule wieder verlassen, bevor sie einen Abschluss erreicht haben, werden laut Heublein, Richter, Schmelzer und Sommer (2012) als Studienabbrecher bezeichnet. Die Quote an Studienabbrechern ist in Chemiestudiengängen mit 43 % besonders hoch. Der von Chemiestudierenden am häufigsten genannte Grund das Studium abzubrechen, ist mit 38 % ein Leistungsproblem. Welche Variablen prädiktiven Charakter bezüglich des Studienerfolgs haben, wurde in verschiedenen Studien untersucht. So zeigte Freyer (2013), dass die Abiturgesamtnote, Wunschfach und Vorwissen bezogen auf den Studienerfolg prädiktive Eigenschaften haben. Weiterhin gibt es Hinweise auf die Lernförderlichkeit selbstregulierten Lernens (SRL) (z. B. Hattie, 2009), was von Freyer (2013) allerdings nicht untersucht wurde. SRL kann nach Schmitz (2001) als ein zirkulärer, metakognitiver Prozess gesehen werden, bei dem Lernende ihr Lernen planen, überwachen und regulieren. Diese Fähigkeit kann trainiert werden (Perels, Landmann & Schmitz, 2007). Zur Erhebung von Strategien des selbstregulierten Lernens werden überwiegend Fragebögen eingesetzt. Der durch Lernstrategiefragebögen erfasste Strategieeinsatz korreliert jedoch häufig nicht mit dem Lernerfolg (Artelt, 2000). Dies wird z. B. von Artelt (2000) auf eine handlungsferne, also allgemein gehaltene Formulierung der Items zurückgeführt. Dabei wird davon ausgegangen, dass Lernende ihre Lernhandlungen nicht abstrahieren und bei einem allgemein gehaltenen Item auf konkrete Lernhandlungen beziehen können. Somit könnten Lernende angeben, dass sie oft bestimmte Strategien anwenden, obwohl sie das tatsächlich nicht tun. Leopold (2009) konnte zeigen, dass mit handlungsnahen, also auf eine bestimmte Lernhandlung bezogenen Items der Lernerfolg besser vorausgesagt werden kann. Fragestellung: In dieser Arbeit wird der Frage nachgegangen, inwiefern SRL einen Einfluss auf den Lernerfolg Chemiestudierender im ersten Semester hat. Zudem wird, aufgrund der problematischen Erfassung von Lernstrategien durch handlungsferne Fragebögen, ein Fragebogen konzipiert, der SRL direkt bezogen auf eine bestimmte Lernhandlung erfassen soll.

Methode: Um den Fragestellungen nachzugehen, wird in den Übungen zur Allgemeinen Chemie eine Interventionsstudie durchgeführt. Zunächst wurde im Zeitraum 10/12 – 02/13 eine Pilotierung mit  $N \approx 100$  Erstsemesterstudierenden des Lehramts im Fach Chemie an der Universität Duisburg-Essen (UDE) durchgeführt. Diese wurden auf zwei Treatments aufgeteilt. Die Kontrollgruppe (srl0) erhielt nur das um fachliche Beispielaufgaben ergänzte übliche Übungsmaterial. Die Interventionsgruppe (srl+) erhielt neben diesen Materialien noch ein SRL-Training. In der Hauptstudie (10/13 – 02/14) werden alle Studierenden des ersten Semesters ( $N \approx 60$ ) trainiert. Als Kontrollgruppe dient die srl0-Gruppe der Pilotstudie. In den Treatmentgruppen werden in einem Prä-Post-Testdesign verschiedene Variablen erhoben. Zu beiden Testzeitpunkten werden das Fachwissen (Freyer, 2013), das Fachinteresse (Freyer, 2013) und der Lernstrategieeinsatz (Wild & Schiefele, 1994) erhoben. Im Prä-Test werden zudem die Fähigkeit zum schlussfolgernden Denken (Wilhelm, Schroeders & Schipolowski, 2009) sowie verschiedene demographische Variablen (Wunschstudiengang, Oberstufenkurs Chemie, Abiturgesamtnote) erfasst. Im Post-Test wird außerdem nach dem aktuellen Kompetenzerleben (Williams & Deci, 1996) gefragt. Semesterbegleitend kommt ein neu entwickelter SRL-Fragebogen AlChEMiE (Allgemeine Chemie Metakognitionsinventar Essen) zum Einsatz. Dieser enthält 18 Items, die auf Basis bestehender Fragebögen (z. B. Wild & Schiefele, 1994) adaptiert wurden. Außerdem wird eine Substichprobe ( $n \approx 10$ ) mit Interviews zum Lernverhalten befragt. Der Interviewleitfaden beruht dabei auf dem SRLIS (Self-Regulated Learning Interview Schedule) von Zimmerman und Martinez-Pons (1986) und einer deutschen Version von Spörer (2004).

Ergebnisse: In der Pilotstudie zeigte der Lernstrategie-Fragebogen AlChEMiE eine gute innere Konsistenz ( $\alpha > 0.8$ ) bei jedem Messzeitpunkt. Weiterhin deuteten sich positive Effekte des SRL-Trainings auf den Lernerfolg an.

Artelt, C. (2000). Strategisches Lernen. Pädagogische Psychologie und Entwicklungspsychologie. Münster: Waxmann.

Freyer, K. (2013). Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie (Diss., Universität Duisburg-Essen, Essen).

Hattie, J. A. C. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. London: Routledge.

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D. & Besuch, G. (2010). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen: Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08. Hannover. Zugriff unter [http://www.his.de/pdf/pub\\_fh/fh-201002.pdf](http://www.his.de/pdf/pub_fh/fh-201002.pdf)

Heublein, U., Richter, J., Schmelzer, R. & Sommer, D. (2012). Die Entwicklung der Schwund und Studienabbruchquoten an den deutschen Hochschulen: Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010. Hannover. Zugriff unter [www.his.de/pdf/pub\\_fh/fh-201203.pdf](http://www.his.de/pdf/pub_fh/fh-201203.pdf)

Leopold, C. (2009). Lernstrategien und Textverstehen: Spontaner Einsatz und Förderung von Lernstrategien. Pädagogische Psychologie und Entwicklungspsychologie. Münster: Waxmann.

Perels, F., Landmann, M. & Schmitz, B. (2007). Trainingskonzeption und Selbstregulation. In M. Landmann & B. Schmitz (Hrsg.), Selbstregulation erfolgreich fördern (S. 19–31). Stuttgart: Kohlhammer.

Schmitz, B. (2001). Self-Monitoring zur Unterstützung des Transfers einer Schulung in Selbstregulation für Studierende: Eine prozessanalytische Untersuchung. Zeitschrift für Pädagogische Psychologie, 15(3), 181–197. doi:10.1024//1010-0652.15.34.181

Spörer, N. (2004). Strategie und Lernerfolg: Validierung eines Interviews zum selbstgesteuertem Lernen (Diss., Universität Potsdam, Potsdam). Zugriff unter <http://opus.kobv.de/ubp/volltexte/2005/150/>

Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktoren- struktur und Reliabilität eines neuen Fragebogens. Zeitschrift für Differentielle und Diagnostische Psychologie, 15(4), 185–200. Zugriff unter <http://opus.kobv.de/ubp/volltexte/2009/3363/>

Wilhelm, O., Schroeders, U. & Schipolowski, S. (2009). BEFKI: Berliner Test zur Erfassung fluider und kristalliner Intelligenz: Unveröffentlichtes Manuskript. Unveröffentlichtes Manuskript.

Williams, G. C. & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. Journal of Personality and Social Psychology, 70(4), 767–779. doi:10.1037/0022-3514.70.4.767

Zimmerman, B. J. & Martinez-Pons, M. (1986). Development of a Structured Interview for Assessing Student Use of Self-Regulated Learning Strategies. American Educational Research Journal, 23(4), 614–628. doi:10.3102/00028312023004614

Montag,  
03.03.2014,  
17:30 - 18:30,

Casino-Gebäude  
Raum 823





## **Wie gut können volitionale Kompetenzen von Schülern/-innen das Auftreten von Problemen beim selbstregulierten Lernen und die Bewertung der eigenen Arbeitsergebnisse vorhersagen?**

Jeder Lernende wird mit Lernsituationen konfrontiert, in welchen seine Motivation für eine Aufgabe fehlt oder nachlässt. Volitionale Strategien der Ressourcenaktivierung stellen einen Teil des Modells motivationsbezogener Kompetenzen dar (Schaller & Spinath, submitted). Sie sollen dazu dienen, die eigene Motivation in solchen Situationen regulieren zu können und umfassen sowohl volitionale Strategien (sensu Wolters, 2003) als auch Strategien der Motivationsregulation (sensu Schwinger, von der Laden, & Spinath, 2007). Es liegen bisher unterschiedliche Sammlungen solcher Strategien vor (z. B. McCann & Garcia, 1999; Schwinger, von der Laden, & Spinath, 2007; in Teilen Wild & Schiefele, 1994). Die Skalen zur Erfassung motivationsbezogener Kompetenzen (MOBEKO; Schaller & Spinath, submitted) integrieren und erfassen neben weiteren Kompetenzen ökonomisch ein breites Spektrum solcher Strategien. Studien konnten zeigen, dass motivationsregulierende Strategien positiv mit Anstrengung und Leistung von Schülern/-innen und Student/-innen zusammenhängen (bspw. Schwinger, Steinmayr, & Spinath, 2012). Wir vermuteten, dass sich die Verwendung solcher ressourcenaktivierender Strategien nicht nur positiv auf Anstrengungsbereitschaft, sondern auch auf das Vermeiden oder Überwinden von Arbeitsproblemen und Prokrastination beim selbstregulierten Lernen auswirkt. Nach einem erfolgreichen Abschluss der Aufgabe sollte sich dies ebenfalls in einer positiveren Bewertung der eigenen Handlungsergebnisse niederschlagen. Durch das Empfinden, Willenskraft aufgewandt und Ressourcen aktiviert zu haben und so zum Erfolg gekommen zu sein, sollte die Verwendung solcher Strategien mit positiven Emotionen und internaler Attribution verknüpft sein.

Zur Untersuchung dieser Fragestellungen wurden die neu entwickelten Skalen zur Erfassung motivationsbezogener Kompetenzen bei Studierenden auf den Schulkontext angepasst und eingesetzt. Kompetenzen der Ressourcenaktivierung wurden zu einem ersten Messzeitpunkt bei  $N = 128$  Gymnasialschülern/-innen der Klassenstufen 11 und 12 erfasst. In einer Nachbefragung bearbeiteten die Schülern/-innen retrospektiv Fragen zu notwendiger Willenskraft bei der Zielverfolgung (Notwendige Volition; Schaller & Spinath, 2013), Anstrengungsbereitschaft (Skala *Anstrengung* aus dem Inventar zur Erfassung von Lernstrategien im Studium; Wild & Schiefele, 1994), Prokrastination (Kurzskala der General Procrastination Scale; Klingsieck & Fries, 2012) und Problemen beim selbstregulierten Lernen (Fragebogen zu Problemen beim Selbstregulierten Lernen; Holz-Ebeling, 1997). Sie wurden außerdem zur emotionalen Bewertung (in Anlehnung an Pekrun, Goetz, Titz, & Perry, 2002) und Attribution bezüglich der Ergebnisse ihrer schulischen Aufgaben der vergangenen Woche befragt.

Es zeigte sich, dass durch die erfassten Kompetenzen der Ressourcenaktivierung sowohl Anstrengungsbereitschaft als auch Prokrastination und das Auftreten von Problemen beim selbstregulierten Lernen vorgesagt werden konnten. Schüler/-innen mit höheren ressourcenaktivierenden Kompetenzen zeigten höhere Anstrengungsbereitschaft, weniger Aufschiebeverhalten und weniger Arbeitszeitprobleme, Arbeitseffektivitätsprobleme und Befindlichkeitsstörungen. Sie gaben außerdem mehr positive und weniger negative Emotionen bei der Ergebnisbewertung an und sahen Erfolge stärker in ihrem eigenen Dazutun begründet. Auch bei Kontrolle der notwendigen Willenskraft blieben diese Zusammenhänge bestehen.

Die Ergebnisse liefern Unterstützung für die Validität der Skalen zur Erfassung ressourcenaktivierenden Kompetenzen im schulischen Kontext. Sie zeigen außerdem die Bedeutung solcher Strategien und deren Förderung für die Selbstregulation bei Schülern/-innen. Die Ergebnisse werden in einem Rahmenmodell motivationsbezogener Kompetenzen eingeordnet und diskutiert.

Holz-Ebeling, F. (1997). Arbeitsprobleme oder das Misslingen von Selbstregulation - Pädagogisch-psychologische Untersuchungen zum Lernen im Studium. Unveröffentlichte Arbeit.

Klingsieck, K. B., & Fries, S. (2012). Allgemeine Prokrastination: Entwicklung und Validierung einer deutschsprachigen Kurzskala der General Procrastination Scale (Lay, 1986). *Diagnostica*, 58, 182-193.

McCann, E. J., & Garcia, T. (1999). Maintaining motivation and regulating emotion: Measuring individual differences in academic volitional strategies. *Learning and Individual Differences*, 11, 259-279.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91-106.

Schaller, P., & Spinath, B. (submitted). Entwicklung und Validierung eines Instruments zur Erfassung motivationsbezogener Kompetenzen (MOBEKO).

Schaller, P., & Spinath, B. (2013). Notwendige Volition. Unveröffentlichte Arbeit.

Schwinger, M., von der Laden, T., & Spinath, B. (2007). Strategien zur Motivationsregulation und ihre Erfassung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39, 57-69.

Schwinger, M., Steinmayr, R., & Spinath, B. (2012). Not all roads lead to Rome - Comparing different types of motivational regulation profiles. *Learning and Individual Differences*, 22, 269-279.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, 38, 189-205.



## Die Genese des Fähigkeitsselbstkonzepts – Identifizierung kausaler Wirkungsmuster?

Aus den Erkenntnissen der empirischen Selbstkonzeptforschung kann bilanziert werden, dass Fähigkeitsselbstkonzepte als „die Gesamtheit der kognitiven Repräsentation eigener Fähigkeiten in akademischen Leistungssituationen“ (Schöne et al., 2002, S. 7) auf verschiedene Weise Lernprozesse begünstigen oder auch erschweren können. Daraus wird geschlussfolgert, dass das Fähigkeitsselbstkonzept ein wichtiger Prädiktor für das Verhalten in Lern- und Leistungssituationen sei und somit einen entscheidenden Anteil am Schulerfolg habe (vgl. Köller & Baumert, 2001; Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, 2003). Aus diesem Grund ist die Entwicklung eines positiven akademischen Fähigkeitsselbstkonzepts von Schülerinnen und Schülern als ein zentrales Bildungsziel anzusehen. Dickhäuser (2006) stellt fest, dass in der pädagogisch-psychologischen Forschung mittlerweile Einigkeit herrscht hinsichtlich des Zusammenhangs zwischen dem Fähigkeitsselbstkonzept und der Leistung. Auf der einen Seite wird angenommen, dass die vorangegangene Leistung einer Person sein Fähigkeitsselbstkonzept determiniert. Diese Annahme wird in der Literatur als „skill development“-Ansatz bezeichnet (vgl. Filipp, 2006). Demzufolge entsteht das Fähigkeitsselbstkonzept teilweise durch kumulierte Erfahrungen (vgl. Dickhäuser, 2006). Auf der anderen Seite wird ebenso davon ausgegangen, dass das Fähigkeitsselbstkonzept die nachfolgende Leistung beeinflusst. Der Effekt des Fähigkeitsselbstkonzepts auf Leistung wird als Prozess des „self enhancement“ bezeichnet. (vgl. ebd.; Filipp, 2006). Insgesamt können beide vorgestellten theoretischen Annahmen, dass sich ein hohes Fähigkeitsselbstkonzept günstig auf die Leistungsentwicklung auswirkt („self enhancement“-Ansatz) und wiederum erzielte Leistungen auf das Fähigkeitsselbstkonzept zurückwirken („skill development“-Ansatz), als bestätigt bezeichnet werden (vgl. Marsh, 1988). Jedoch hebt Dickhäuser (2006) hervor, dass die wesentlichen Mechanismen von skill development- und self enhancement-Prozessen unklar sind. Das Zusammenwirken beider genannten Modelle ist jedoch in der pädagogischen Praxis ein häufiger Fall und wird als „reciprocal effect“-Modell bezeichnet (vgl. Marsh & Yeung, 1997).

Auf Grundlage dieser theoretischen Annahmen verfolgt der Beitrag das Ziel die genauen kausalen Wirkungsmuster zu identifizieren und Erkenntnisse über die Interdependenz zwischen dem Fähigkeitsselbstkonzept und der Schülerleistung zu gewinnen. Somit wird der Frage nachgegangen, welche der theoretischen Ansätze des „reciprocal effect“-Modells sich bei der Genese des Fähigkeitsselbstkonzepts von Schülerinnen und Schülern empirisch absichern lassen?

Mit einem Cross-Lagged-Panel Design werden die angenommenen Kausalzusammenhänge zwischen den beiden Konstrukten Fähigkeitsselbstkonzept und Schülerleistung ermittelt. Als Datengrundlage dienen dafür zwei Erhebungen im Schulentwicklungsprojekt „Ganz In – Mit Ganzttag mehr Zukunft. Das neue Ganzttagsgymnasium NRW“. Etwa 3000 SuS haben im Schuljahr 2010/11 erstmals und im Schuljahr 2012/13 erneut an einer Leistungsmessung und einer Schülerbefragung teilgenommen. Schließlich vermögen die Aussagen über Zusammenhänge zwischen dem Fähigkeitsselbstkonzept und der Schülerleistung Aufschluss über die genauen Mechanismen zur Genese des akademischen Selbstkonzepts zu geben.

Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte: Entstehung, Auswirkung, Förderung. Zeitschrift für Pädagogische Psychologie, 20(1-2), 5-8.

Filipp, S.H. (2006). Entwicklung von Fähigkeitsselbstkonzepten. Zeitschrift für Pädagogische Psychologie, 20(1-2), 65-72.

Köller, O., & Baumert, J. (2001). Leistungsgruppierungen in der Sekundarstufe I: Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. Zeitschrift für Pädagogische Psychologie, 15(2), 99-110.

Marsh, H. W. (1988). Causal effect of academic self-concept on academic achievement: A reanalysis of Newman (1984). Journal of Experimental Education, 56, 100-104.

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. Journal of Educational Psychology, 89(1), 41-54.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). Skalen zur Erfassung des schulischen Selbstkonzepts – SESSKO. Göttingen: Hogrefe.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2003). Das Fähigkeitsselbstkonzept und seine Erfassung. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), Diagnostik von Motivation und Selbstkonzept: Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik, Band 2 (S. 3-14). Göttingen: Hogrefe.



## **Motivationsregulation in unterschiedlichen Lernsituationen im Studium**

Im tertiären Bereich wird dem selbstregulierten Lernen hohe Relevanz zugeschrieben, da sich Studierende umfangreiche Wissensbestände selbst erarbeiten müssen (vgl. Wild, 2000). Boekaerts (1996) beschreibt dabei die aktive Kontrolle von motivationalen Prozessen als einen wesentlichen Aspekt. Bisher wurde die Motivationsregulation bei Studierenden nur in wenigen Studien adressiert (vgl. Schwinger, Steinmayr, & Spinath, 2009) und dabei auf eine systematische Erfassung von typischen Lernsituationen verzichtet, welche diesen Aspekt des selbstregulierten Lernens in besonderem Maße erfordern. Insbesondere die Frage, ob sich Lernsituationen im Hinblick auf die Wichtigkeit unterscheiden, Strategien der Motivationsregulation einzusetzen wurde bisher nicht empirisch überprüft. Ziel dieses Beitrags ist einerseits die Identifikation von Lernsituationen im Studium, zu deren erfolgreichen Bewältigung Kompetenzen des selbstregulierten Lernens und insbesondere Kompetenzen der Motivationsregulation erforderlich sind. Andererseits soll die Frage untersucht werden, ob und inwieweit sich die verschiedenen Situationen im Hinblick auf die Relevanz der Motivationsregulation unterscheiden.

Mit N=39 Dozierenden sowie N=69 leistungsstarken Studierenden der Fächer Elektrotechnik, Lehrer(innen)bildung in den MINT-Fächern, Psychologie und Wirtschaftswissenschaften wurden an verschiedenen Universitätsstandorten (Universität Augsburg, Technische Universität Darmstadt, Universität Erlangen-Nürnberg und Universität Wien) halb-standardisierte Expert(inn)eninterviews geführt. Dabei wurde nach zentralen Lernsituationen im jeweiligen Studienfach, sowie nach Strategien des selbstregulierten Lernens gefragt, die für diese Situationen als besonders zielführend angesehen wurden. Die von den Expert(inn)en beschriebenen Lernsituationen wurden kategorisiert, wobei sich eine gute Inter-Rater-Übereinstimmung ergab. Zur Kategorisierung der genannten Strategien diente ein Kategoriensystem, das sowohl theoriegeleitet als auch induktiv aus dem Datenmaterial abgeleitet wurde. Hier war die Inter-Rater-Übereinstimmung akzeptabel bis gut. Auf Basis der aus den Interviews gewonnenen prototypischen Lernsituationen und der kategorisierten Regulationsstrategien wurde eine Onlinebefragung konzipiert. In diesem Format wurde die Wichtigkeit des Einsatzes unterschiedlicher Regulationsstrategien über die verschiedenen prototypischen Lernsituationen hinweg von N=141 Dozierenden und N=162 leistungsstarken Studierenden eingeschätzt.

Die Ergebnisse der Interviews zeigten, dass aus Expert(inn)ensicht über alle Fächer hinweg eine stabile Kerngruppe prototypischer Lernsituationen vorhanden ist, die Kompetenzen des selbstregulierten Lernens erfordern. Die in diesen Situationen als bedeutsam erachteten Strategien des selbstregulierten Lernens fielen in mehr als 19% der Fälle in die Kategorie „Motivationsregulation“, was die Wichtigkeit von Kompetenzen zur Regulation der Motivation im Studium unterstreicht. Weiterhin zeigte sich, dass Expert(inn)en den Einsatz von Motivationsregulationsstrategien abhängig von der jeweiligen Lernsituation unterschiedlich häufig als bedeutsame Strategie des selbstregulierten Lernens benannten. Dieses Ergebnismuster konnte durch die Befunde der Onlinebefragung validiert und weiter präzisiert werden: Abhängig von der jeweiligen Anforderungssituation schätzten die Expert(inn)en die Wichtigkeit der Anwendung von Motivationsregulationsstrategien signifikant unterschiedlich ein. So wurde beispielsweise beim „Verfassen einer Abschlussarbeit“ die Wichtigkeit, Strategien der Motivationsregulation einzusetzen als sehr viel höher eingeschätzt, als etwa bei der „Referatsvorbereitung“. Im Beitrag werden die sich daraus ergebenden Implikationen für die Diagnose der Motivationsregulation von Studierenden diskutiert.

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *Educational Psychologist*, 1, 100-112.

Schwinger, M., Steinmayr, R., & Spinath, B. (2009). How do motivational regulation strategies affect achievement: Mediated by effort management and moderated by intelligence. *Learning and Individual Differences*, 19, 621-627.

Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster: Waxmann.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## Zur Rolle des Wissens über Strategien zur Motivationsregulation beim selbstregulierten Lernen

Theoretischer Hintergrund: Erst seit kurzem befasst sich die Forschung zum selbstregulierten Lernen (SRL) auch mit der Regulation der Motivation (vgl. Leutner, Barthel & Schreiber, 2001; Schwinger, von der Laden & Spinath, 2007; Schwinger & Stiensmeier-Pelster, 2012; Winne & Hadwin, 2008; Wolters, 1998, 2003). Der Fokus dieser Forschung liegt auf der Untersuchung der Nutzung von Motivationsregulationsstrategien. Wenn es allerdings um die Frage geht, unter welchen Bedingungen Lernende ihre eigene Motivation regulieren, sollte der Fokus stärker auf die lernerseitigen Voraussetzungen für die Nutzung von Motivationsregulationsstrategien gelegt werden. Nach dem Modell des good strategy user (Pressley, Borkowski & Schneider, 1989) stellt sich daher die Frage, inwiefern das Wissen über Strategien zur Motivationsregulation die Motivation im Lernprozess und den resultierenden Lernerfolg beeinflusst. Daraus ergaben sich folgende Hypothesen: H1: Es besteht ein positiver Zusammenhang zwischen dem Wissen über Strategien zur Motivationsregulation und der aktuellen Motivation nach einer Phase der Selbstmotivierung. H2: Es besteht ein positiver indirekter Zusammenhang zwischen dem Wissen über Strategien zur Motivationsregulation und dem Lernerfolg beim SRL, vermittelt über die aktuelle Motivation im Lernprozess.

Methode: Stichprobe: N=99 Schülerinnen und Schüler (SuS) aus vier Klassen der Sekundarstufe I haben an dieser Studie teilgenommen. 40 Schüler waren männlich, das Alter betrug M=14.04 Jahre (SD=0.96 Jahre).

Instrumente: Das Wissen über Strategien zur Motivationsregulation wurde anhand eines Strategiewissenstests erfasst, welcher analog zum WLST (Schlagmüller & Schneider, 2007) konzipiert ist. Eingeschätzt wird die Adäquatheit vorgegebener Strategien zur Regulation der Motivation in einer motivational schwierigen Lernsituation. Der Score berechnet sich aus der Übereinstimmung des individuellen Urteils mit einem aggregierten Expertenurteil. Der Test ist inhaltlich unterteilt in eine Skala zur Initiation der Motivation zu Beginn des Lernens und eine Skala zur Persistenz der Motivation während des Lernens. Die Motivation zu lernen wurde anhand des Fragebogens zur Aktuellen Motivation (Rheinberg, Vollmeyer & Burns, 2001) erfasst. Der Lernerfolg wurde anhand von Multiple-Choice-Wissenstests zu zwei computerbasierten Lernumgebungen "Auftrieb in Flüssigkeiten" und "Säuren und Basen" erfasst (vgl. Marschner, 2011; Thillmann, 2008).

Durchführung: Zunächst füllten die SuS den Test zum Wissen über Motivationsregulationsstrategien aus. Dann erhielten sie Screenshots der beiden Lernumgebungen und schätzten jeweils ihre Motivation ein, damit zu lernen. Analog zu der Studie von Leutner et al. (2001) wurden die SuS im Anschluss dazu aufgefordert, sich selbst zum Lernen mit der für sie weniger motivierenden Lernumgebung zu motivieren. Im Anschluss wurde erneut ihre Motivation erfasst. Ebenso wurde ihre Motivation unmittelbar vor und nach dem Lernen erfasst. Zudem wurde vor und nach dem Lernen das inhaltspezifische Wissen erfasst.

Ergebnisse: Die Reliabilitäten der eingesetzten Maße erwiesen sich als zufriedenstellend. Zur Überprüfung von H1 wurden Korrelationen zwischen dem Motivationsregulationsstrategiewissen und der Motivation nach der Selbstmotivierung berechnet. Lediglich zwischen der Motivation nach der Selbstmotivierung und der Strategiewissensskala Persistenz zeigt sich ein positiver Zusammenhang ( $r=.30$ ;  $p<.05$ ). Somit kann Hypothese 1 nur in Bezug auf das Strategiewissen zur Persistenz bestätigt werden. Zur Überprüfung von H2 wurden Pfadmodelle gerechnet. Hierbei zeigt sich ein positiver Effekt des Strategiewissens zur Persistenz auf die Motivation vor dem Lernen ( $\beta=.32$ ;  $p<.05$ ), welche wiederum einen positiven Effekt auf die Motivation nach dem Lernen ( $\beta=.72$ ;  $p<.05$ ) und diese auf den Lernerfolg hat. Das berichtete Modell zeigt einen akzeptablen Fit ( $\chi^2=1.058$ ;  $df=1$ ;  $p=.304$ ;  $CFI=.999$ ;  $RMSEA=.024$ ). Somit kann Hypothese 2 in Bezug auf das Strategiewissen zur Persistenz bestätigt werden.

Diskussion: Die Ergebnisse dieser Studie sprechen einerseits für die Relevanz des Strategiewissens zur Motivationsregulation für die Motivation und den Lernerfolg beim SRL. Andererseits konnte anhand der Skalenaufteilung gezeigt werden, dass insbesondere der Fähigkeit zur Aufrechterhaltung der Motivation im Lernprozess eine zentrale Rolle zukommt (vgl. Vollmeyer & Rheinberg, 2000). Damit ergeben sich neue empirische Hinweise für eine gezielte Förderung der Fähigkeit zum Selbstregulation der Motivation beim Lernen.

Leutner, D., Barthel, A. & Schreiber, B. (2001). Studierende können lernen, sich selbst zum Lernen zu motivieren: Ein Trainingsexperiment. Zeitschrift für Pädagogische Psychologie, 15, 155-167.

Marschner, J. (2011). Adaptive Feedback zur Unterstützung des selbstregulierten Lernens durch Experimentieren. Dissertation. Universität Duisburg-Essen.

Rheinberg, F., Vollmeyer, R. & Burns, B.D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. Diagnostica, 47, 57-66.

Pressley, M., Borkowski, J.G. & Schneider, W. (1987). Cognitive strategy users coordinate metacognition, and knowledge. In R. Vasta & G. Whitehurst (Eds.), Annals of child development (pp. 89-129). New York: JAI Press.

Schlagmüller, M. & Schneider, W. (2007). WLST 7-12. Würzburger Lesestrategie-Wissenstest für die Klassen 7-12. Göttingen: Hogrefe.

Schwinger, M. & Stiensmeier-Pelster, J. (2012). Effects of motivational regulation on effort and achievement: A mediation model. International Journal of Educational Research, 56, 35-47.

Schwinger, M., von der Laden, T. & Spinath, B. (2007). Strategien zur Motivationsregulation und ihre Erfassung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39, 57-69.

Thillmann, H. (2008). Selbstreguliertes Lernen durch Experimentieren: Von der Erfassung zur Förderung. Dissertation. Universität Duisburg-Essen.

Vollmeyer, R. & Rheinberg, F. (2000). Does motivation affect performance via persistence? Learning and Instruction, 10, 293-309.

Winne, P.H. & Hadwin, A.F. (2008). The wave of motivation and self-regulated learning. In D.H. Schunk & B.J. Zimmerman (Eds.), Motivation and self-regulated learning: Theory, research, and applications (297-314). Mahwah, NJ: Erlbaum.

Wolters, C. A. (1998). Self-regulated learning and college students' regulation of motivation. Journal of Educational Psychology, 90, 224-235.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. Educational Psychologist, 38, 189-205.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823





Melanie Vogelsberg (Rheinische-Friedrich-Wilhelms Universität), Udo Käser (Rheinische-Friedrich-Wilhelms Universität),  
Una Maria Röhr-Sendlmeier (Rheinische-Friedrich-Wilhelms Universität)

## ***Entwicklungsförderung metamemorialer Gedächtnisstrategien: Inwiefern können Schulanfänger spielerisch an die Strategie des kategorialen Organisierens herangeführt werden?***

Für den Schulerfolg sind metakognitive Fähigkeiten von besonderer Wichtigkeit (Hasselhorn & Labuhn, 2008). Das Beherrschen von Gedächtnisstrategien steht beispielsweise eng mit schulischem Lernerfolg in Zusammenhang. Hierbei entwickeln sich Fähigkeiten, Lernsituationen strategisch zu meistern, bei Heranwachsenden im Wesentlichen in der Grundschulzeit (Schneider & Büttner, 2008). Insofern ist es eine zentrale Aufgabe der Grundschule, den Erwerb solcher Kompetenzen durch geeignete Lernumgebungen und Fördermaßnahmen zu unterstützen. Entsprechend beinhalten deutsche Curricula auch die Forderung nach der Vermittlung metakognitiver Kompetenzen.

Ob dies im deutschen Schulsystem gut gelingt, ist allerdings fraglich. So zeigen zum Beispiel Käser und Cummings (2012), dass deutsche Schülerinnen und Schüler der dritten bis fünften Schulklasse nur in geringem Maße über die Fähigkeit verfügen, Gedächtnisaufgaben mittels geeigneter Strategien effektiv zu lösen. Insbesondere zeigen sich erhebliche Defizite hinsichtlich der Fähigkeit des kategorialen Organisierens. Dies ist überaus problematisch, da gerade diese Strategie für die Aneignung von Welt und das Lernen zu lernen eine entscheidende Rolle spielt, weswegen sie von Perleth (1992) auch als mächtige Strategie bezeichnet wird.

Dies macht deutlich, wie wichtig die Förderung der kategorialen Organisationsfähigkeit in der Grundschule ist. Weitgehend offen ist jedoch die Frage, in welchem Alter und in welcher Weise mit einer derartigen Förderung sinnvoll begonnen werden kann. Einerseits weist etwa Hasselhorn (1996) darauf hin, dass es nicht sinnvoll sei, diese Strategie vor dem neunten Lebensjahr zu üben, da sich das notwendige Vorwissen hinsichtlich der Typizität von Kategorien und der Assoziativität ihrer Elemente erst im Alter von neun bis zehn Jahren konsolidiert. Andererseits wird aber zum Beispiel im Rahmen von mathematischer Frühförderung in der Vorschule bzw. im Anfangsunterricht in Mathematik der Grundschule schon sinnvoll mit Aufgaben zum Kategorisieren und Abzählen elementarer Objekte gearbeitet, um die Verbindung zwischen (An-) Zahl- und Mengenkonzept zu fördern (Krajewski & Schneider, 2006).

Vor dem Hintergrund dieser Überlegungen wird ein Kurskonzept für Erst- und Zweitklässler präsentiert, welches das Ziel verfolgt, über einen Zeitraum von drei Monaten den Entwicklungsprozess von Gedächtnisstrategien spielerisch zu fördern. Der Kurs umfasst neun Trainingseinheiten, in denen die Kinder unterschiedliche Spiele, die spezifisch zur Förderung der kategorialen Organisationskompetenz konzipiert worden sind, kennenlernen und in Kleingruppen erproben. Hierzu werden Ergebnisse einer Evaluationsstudie präsentiert (N = 272), in der die Wirksamkeit dieses Konzepts unter den variierenden Bedingungen einer expliziten, impliziten oder inzidentellen Lernumgebung (Röhr-Sendlmeier & Käser, 2012) in einem quasiexperimentellen Kontrollgruppendesign mit einer Kontroll- und drei Trainingsgruppen über vier Messzeitpunkte im Längsschnitt untersucht wurde.

Dabei zeigt sich, dass im Wesentlichen nur ältere Kinder unter der expliziten Trainingsbedingung im Verlauf der Untersuchung verbesserte Gedächtnisleistungen erzielen, ohne dass diese Entwicklung jedoch auf eine Verbesserung metamemorialer Fähigkeiten zurückgeführt werden kann. Insofern legen die Ergebnisse der Studie nahe, dass der Einsatz des Trainings essentielle Determinanten für die Verbesserung von Gedächtnisleistungen, wie Aufmerksamkeitsprozesse und kapazitative Aspekte des Arbeitsgedächtnisses, fördert, wenngleich die Schülerinnen und Schüler nicht auf der Ebene des Metagedächtnisses profitieren.

Diskutiert wird zum einen die Frage, inwiefern dieses Ergebnis auf die Struktur des Trainings oder auf Entwicklungsfaktoren der Kinder zurückgeführt werden kann, sowie zum anderen, ob das Training möglicherweise einen vorbereitenden Charakter im Sinne eines Primings oder eines Sleeper-Effektes besitzt, der sich ggf. erst im Verlauf der späteren Entwicklung während der dritten oder vierten Klasse niederschlägt. Zugleich wird der Frage nachgegangen, welche Empfehlungen für die notwendige Förderung metakognitiver Fähigkeiten von Schülerinnen und Schülern in der Grundschule angesichts der vorliegenden Befunde ausgesprochen werden können.

Hasselhorn, M. (1996). Kategoriales Organisieren bei Kindern. Zur Entwicklung einer Gedächtnisstrategie. Göttingen: Hogrefe.

Hasselhorn, M. & Labuhn, A. S. (2008). Metakognition und selbstreguliertes Lernen – Metacognition and Self-regulated Learning. In W. Schneider & M. Hasselhorn (Hrsg.), Handbuch der Pädagogischen Psychologie. Band X (S. 28-37). Göttingen: Hogrefe.

Käser, U. & Cummings, A. V. (2012). Gedächtnisleistung und metamemorale Organisationsleistung in Abhängigkeit vom Modus des Lernens. In U. M. Röhr-Sendlmeier (Hrsg.), Inzidentelles Lernen – wie wir beiläufig Wissen erwerben (S. 147-194). Berlin: Logos.

Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. Psychologie in Erziehung und Unterricht, 53, 246-262.

Perleth, C. (1992). Strategienutzung, Metagedächtnis und intellektuelle Begabung. München: LMU.

Röhr-Sendlmeier, U. M. & Käser, U. (2012). Das Lernen komplexer sprachlicher Stimuli – Wissenserwerb nach unterschiedlichen Lernmodi. In U. M. Röhr-Sendlmeier (Hrsg.), Inzidentelles Lernen – wie wir beiläufig Wissen erwerben (S. 43-86). Berlin: Logos.

Schneider, W. & Büttner, G. (2008). Entwicklung des Gedächtnisses bei Kindern und Jugendlichen. In R. Oerter & L. Montada (Hrsg.), Entwicklungspsychologie (S. 480-501). 6. Auflage. Weinheim, Basel, Berlin: PVU.





Timo Reuter (Universität Koblenz-Landau), Wolfgang Schnotz (Universität Koblenz-Landau), Renate Rasch (Universität Koblenz-Landau)

## **Externe Repräsentationen als kognitive Werkzeuge beim Bearbeiten von problemhaltigen Textaufgaben in der Grundschule**

Wie können Grundschüler beim Lösen problemhaltiger Textaufgaben kognitiv unterstützt werden? Problemaufgaben stellen hohe Anforderungen an Schüler, da sie nicht mit Routinen zu lösen sind. Das Problem muss adäquat repräsentiert, ein mentales Modell entworfen und in einem dynamisch iterativen Prozess mit den in der Repräsentation externalisierten Informationen abgeglichen werden (Schnotz, Baadte, Müller und Rasch, 2011). Dies geschieht häufig nur unzureichend, da viele Schüler keine eigenen externen Repräsentationen erstellen (Elia, Van den Heuvel-Panhuizen und Kolovou, 2009). Werden aber ausgearbeitete externe Repräsentationen als Hilfsmittel vorgegeben, können u.a. kognitive Selbsterklärungsprozesse und die Entlastung des Arbeitsgedächtnisses zu kurz kommen (Cox, 1997).

In diesem Spannungsfeld geht die Studie erstens der Frage nach, ob und in welcher Form vorgegebene externe Repräsentationen in Abhängigkeit von Aufgaben- und Schülermerkmalen den Lösungsprozess bei problemhaltigen Textaufgaben unterstützen. Zweitens soll untersucht werden, ob durch eine wiederholte Verwendung vorgegebener externen Repräsentationen ein Lerneffekt hinsichtlich der selbstständigen Erstellung von Repräsentationen festzustellen ist.

In einem multifaktoriellen experimentellen Design bearbeiten Viertklässler (N≈200) zu mehreren Messzeitpunkten jeweils drei unterschiedliche Problemaufgaben: eine Vergleichs-, eine Bewegungs- und eine Kombinatorik-Aufgabe (Rasch, 2003). Zu den Messzeitpunkten 1 (Prä-Messung) und 4 (Post-Messung) werden keine Repräsentationen als Hilfsmittel vorgelegt. Zu den Messzeitpunkten 2 und 3 erhalten die Probanden der Experimentalgruppe das eine Mal zu jeder Aufgabe eine Zeichnung als Hilfsmittel, das andere Mal zu jeder Aufgabe eine Tabelle. Die Probanden der Kontrollgruppe bearbeiten die Aufgaben ohne vorgegebenes Hilfsmittel. In der Experimentalgruppe wird zudem der Ausarbeitungsgrad der Tabellen und Zeichnungen variiert: Je höher der Ausarbeitungsgrad, desto mehr Informationen enthält die Repräsentation bereits und desto geringer ist der Raum für eigene Ausarbeitungen durch die Probanden. Die Durchführung erfolgt im Klassenraum, wobei die Treatments ad Random auf die einzelnen Schüler verteilt werden. Abhängige Variablen sind die Lösungsqualität und die Bearbeitungsdauer. Die Lösungsqualität wird inhaltsanalytisch anhand der Aufgabenblätter, die Bearbeitungsdauer mittels digitaler Zeitmessung erhoben. Schülermerkmale wie Leseverständnis, kognitive Fähigkeiten, Rechenfertigkeit, Meta-Kognition und Einstellungen zum Problemlösen werden als intervenierende Variablen betrachtet und mit standardisierten Tests sowie schriftlichen Fragebögen gemessen. Die Daten werden mit multivariaten statistischen Verfahren ausgewertet.

Das Poster präsentiert das methodische Design der Studie, gibt einen Überblick der zentralen Hypothesen und stellt erste Ergebnisse sowie weitere geplante Auswertungsstrategien vor. Die Untersuchung möchte Ansätze zeigen, wie der Umgang mit Repräsentationen als kognitive Werkzeuge beim Lösen problemhaltiger Textaufgaben in der Grundschule verbessert werden kann, um das von PISA und TIMMS attestierte kognitive Potenzial deutscher Schüler mehr als bisher im Unterricht ausnutzen zu können (Leutner, Fleischer, Wirth, Greif und Funke, 2012).

Cox, R. (1997). Representation interpretation versus representation construction: a controlled study using switchERII. In: du Boulay, B. und Mizoguchi, R. (Hrsg.), *Artificial intelligence in education: Knowledge and media in learning systems* (Proceedings of the 8th World Conference of the Artificial Intelligence in Education Society), S 434–444, Amsterdam: IOS.

Elia, I., van den Heuvel-Panhuizen, M. und Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM Mathematics Education (ZDM)*, 41 (5), S. 605–618.

Leutner, D., Fleischer, J., Wirth, J., Greiff, S., & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien. *Untersuchungen zur Dimensionalität. Psychologische Rundschau*, 63 (1), S. 34–42.

Rasch, R. (2003). 42 Denk- und Sachaufgaben. *Wie Kinder mathematische Aufgaben lösen und diskutieren*. Seelze-Velber: Kallmeyer.

Schnotz, W., Baadte, C., Müller, A., & Rasch, R. (2011). Kreatives Problemlösen mit bildlichen und beschreibenden Repräsentationen. In: Sachs-Hombach, K., Totzke, R. (Hrsg.), *Bilder – Sehen – Denken. Zum Verhältnis von begrifflichen-philosophischen und empirisch-psychologischen Ansätzen in der bildwissenschaftlichen Forschung* (S. 204-252). Köln: Halem.



## **Formatgebundene Verarbeitungspräferenzen: Gibt es Verbalisierer und Visualisierer wirklich?**

Es existieren zahlreiche empirische Belege, dass individuelle Merkmale des Lernenden den Wissenserwerb mit Texten und Bildern entscheidend beeinflussen können (Mayer, 2005). Neben dem domänenspezifischen Vorwissen und der allgemeinen Intelligenz, scheinen vor allem räumliche Fähigkeiten positiv mit einer effektiven Verarbeitung von piktorialen und verbalen Repräsentationen assoziiert zu sein (Höfler, 2010). Darüber hinaus existiert jedoch die weitverbreitete Annahme, dass auch formatgebundene Verarbeitungspräferenzen beim Wissenserwerb mit multimedialen Dokumenten berücksichtigt werden sollten. Diese Vorstellung geht auf die Vermutung zurück, dass Individuen im Lauf ihrer kognitiven Entwicklung stabile Vorlieben, Gewohnheiten oder Präferenzen entwickeln, die letztlich mitbestimmen, wie effektiv Informationen aufgenommen und verarbeitet werden. Etwas überspritzt formuliert entstehen auf diese Weise unterschiedliche Arten zu denken, die in der theoretischen Rahmenkonzeption der cognitive styles zusammengefasst und beschrieben werden. Ein vor allem in der Multimediaforschung bekanntes Beispiel für einen cognitive style ist die Unterscheidung zwischen sogenannten Verbalisierern und Visualisierern. Demnach sollen Verbalizer besonders gut mit gesprochenen oder geschriebenen Texten Wissen erwerben können, während Visualizer erfolgreicher mit piktorialen Repräsentationsformaten (schematische Abbildungen, Schaubilder) lernen. Obwohl diese theoretische Konzeption im ersten Moment sehr eingängig erscheinen mag, lassen sich nur wenige empirische Belege zugunsten der Verbalizer-Visualizer-Hypothese finden (Kollöffel, 2012). Hinzu kommen grundlegende, theoretische Probleme, die eine Gültigkeit der VV-Hypothese nur wenig plausibel erscheinen lassen. Zum einen ist die postulierte Bipolarität des Konstrukts eine viel zu strenge Annahme, die im Grunde nur ein entweder oder zulässt. Zum anderen ist es nicht sehr wahrscheinlich, dass formatgebundene Verarbeitungspräferenzen auf globaler Ebene kognitive Verarbeitungsprozesse beeinflussen, ohne dass domänenspezifische Wissensinhalte oder Kompetenzen eine Rolle spielen. Da die Vorstellung unterschiedlicher Verarbeitungspräferenzen aber gerade bei Lehramtsstudierenden weit verbreitet ist, konzipierten wir in Zusammenarbeit mit einer hessischen Realschule eine empirische Untersuchung, um die VV-Hypothese zu falsifizieren. Zu diesem Zweck bearbeiteten 142 Schülerinnen und Schüler einer sechsten Jahrgangsstufe mehrere Testaufgaben bzw. Fragebögen. Unter anderem wurde den Kindern eine übersetzte Version des Verbalizer-Visualizer-Questionnaires (VVQ) nach Kirby, Moore und Schofield (1988) vorgelegt. Zudem bearbeiteten Sie einen Test zur Bild-Text-Integration (Ullrich et al., 2012) und einen Test zur Lesekompetenz, der keine Bilder enthielt. Wenn die VV-Hypothese tatsächlich zutreffen würde, dann sollten die als Verbalisierer kategorisierte Kinder besser im Lesetest abschneiden und die als Visualizer kategorisierten Schülerinnen und Schüler höhere Punktwerte bei den Text-Bild-Aufgaben erreichen. Ursprünglich war geplant worden, anhand des VVQs Extremgruppen zu bilden und miteinander zu vergleichen, doch diese erwiesen sich als zu klein, weshalb eine Trennung entlang des Medians vorgenommen wurde. Zu unserer großen Überraschung zeigte sich dabei, dass die Ergebnisse der Kinder in den beiden Leistungstests der VV-Hypothese entsprachen. So lösten die Schülerinnen und Schüler, die anhand des VVQs als Visualisierer kategorisiert worden waren, signifikant mehr Items bei den Text-Bild-Aufgaben ( $t=2.19$ ;  $df=110$ ;  $p=.03$ ) als Kinder der Verbalisierer-Kategorie. Umgekehrt lösten Schülerinnen und Schüler, die als Verbalisierer kategorisiert worden waren, signifikant mehr Items beim Lesekompetenztest ( $t=2.10$ ;  $df=110$ ;  $p=.04$ ) als Kinder der Visualisierer-Kategorie. Mit einem  $d=.42$  bzw.  $d=.40$  (nach Cohen) konnte somit kleine bis mittlere Effekte nachgewiesen werden, die offenbar auf formatspezifische Verarbeitungspräferenzen zurückzuführen sind. Obwohl die Ergebnisse der Untersuchung ziemlich eindeutig für die VV-Hypothese sprechen, sollte erwähnt werden, dass der VVQ ein unzuverlässiges Instrument darstellt. Demnach mussten wir einzelne Items trotz korrekter Polung entfernen, da sie schlichtweg nicht trennscharf waren, um überhaupt eine sinnvolle Auswertung zu ermöglichen. Die unzureichende interne Konsistenz des Fragebogens ( $\alpha = .38$ ), lässt daher den Verdacht aufkommen, dass möglicherweise nicht nur die formatgebundene Verarbeitungspräferenz gemessen wurde, die zur Kategorisierung der Schüler und damit zu den signifikanten Ergebnissen geführt hat.

Höfler, T. N. (2010). Spatial ability: Its influence on Learning with visualizations - a meta-analytic review. *Educational Psychology Review*, 22, 245-269.

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology*, 13, 169-184.

Kollöffel, B. (2012). Exploring the relation between visualizer-verbalizer cognitive styles and performance with visual or verbal learning material. *Computers & Education*, 58, 697 - 706.

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge, England: Cambridge University Press.

Ullrich, M., Schnotz, W., Horz, H., McElvany, N., Schroeder, S., Baumert, J. (2012). Kognitionspsychologische Aspekte der Bild-Text-Integration. *Psychologische Rundschau*, 63,11-17.

## **Kognitive Ressourcen beim Problemlösen: Strategische Allokation und Automatisierung**

Theoretischer Hintergrund: Aufgaben aus der Domäne Problemlösen zeichnen sich durch ihre Komplexität und ihre Neuartigkeit aus (Klieme, Leutner, & Wirth, 2005; Mayer, 1994). In diesem Beitrag wird Problemlösen vor allem in Bezug auf Technologiereiche Umgebungen (Technology-Rich Environments; kurz TRE) und auf die alltäglichen, beruflichen und soziale Anforderungen von Erwachsenen (OECD, 2009) betrachtet. Generell – und auch bezogen auf TRE – erfordert Problemlösen einen hohen Aufwand kognitiver Ressourcen, weil unter anderem abstraktes Denken und beispielsweise der Umgang mit einer dynamischen Umgebung gefordert wird (Klieme, 2004; Mayer, 1994). Entsprechend findet sich vor allem bei schweren Problemlöseaufgaben und schwachen Problemlösern ein enger positiver Zusammenhang zwischen der insgesamt in die Lösung der Aufgabe investierten Zeit und dem Lösungserfolg (Goldhammer, Naumann, Stelter, Tóth, Rölke, & Klieme, in press). Dabei ist allerdings – wenn man unterschiedliche Teilschritte des Problemlöseprozesses betrachtet – nicht von einer uniformen Beziehung zwischen Zeitverbrauch und Lösungswahrscheinlichkeit auszugehen. Ein positiver Zusammenhang, der aus der gezielten Allokation kognitiver Ressourcen resultiert, ist vor allem bei solchen Teilschritten zu erwarten, die bewusste Verarbeitung erfordern und für die Aufgabenlösung unmittelbar relevant sind. Damit überhaupt ausreichend kognitive Ressourcen vorhanden sind, ist es andererseits hilfreich, wenn basale Teilschritte der Aufgabenbearbeitung weitgehend automatisiert sind. Für solche Teilschritte ist folglich eine negative Beziehung zwischen Zeitverbrauch und Lösungswahrscheinlichkeit zu erwarten.

Fragestellung: Ziel dieser Arbeit ist, die These zu erhärten, dass sowohl die Allokation von Ressourcen bei bewusst bearbeiteten Teilschritten einer Problemlösesequenz als auch die Automatisierung basaler Teilschritte positiv zur Bearbeitung von Problemlöseaufgaben beiträgt.

Methode: Die hier verwendeten Daten stammen aus dem deutschen Feldtest des 'Programme for the International Assessment of Adult Competencies' (PIAAC-Studie; N=1020), in der der Bildungsstand von Erwachsenen unter anderem in der Domäne 'Problem Solving in Technology-Rich Environments' computerbasiert erhoben wurde. Für die hiesige Analyse wurde die Aufgabe 'Job Search' ausgewählt. Bei dieser Aufgabe müssen die Testpersonen ausgehend von einer Suchmaschinen-Ergebnisseite anhand bestimmter Kriterien Internetseiten zur Jobsuche auswählen und bookmarken. Dabei lassen sich sowohl Teilschritte identifizieren, die kontrolliert erfolgen müssen (die Auswahl und Lektüre von relevanten Suchergebnissen), als auch solche, die potenziell automatisierbar sind (das Setzen von Bookmarks).

Ergebnisse und Diskussion:

Allokation von Ressourcen: Zur Bestimmung des Zusammenhangs zwischen der Zeit, die die Testpersonen in der Aufgabe 'Job Search' auf relevanten Ergebnisseiten verbringen und der Wahrscheinlichkeit einer korrekten Lösung wurden logistische Regressionen gerechnet. Die Ergebnisse stützen die Hypothese: Für die auf relevanten Ergebnisseiten verbrachte Zeit ergab sich ein positiver Zusammenhang mit der Lösungswahrscheinlichkeit,  $BETA = 0.96$  ( $z = 2.53$ ,  $p < .01$ ). Zu Kontrollzwecken wurde außerdem die Zeit bestimmt, die auf irrelevanten Ergebnisseiten verbracht wurde. Für diese ergab sich wie erwartet kein Zusammenhang mit der Lösungswahrscheinlichkeit,  $BETA = 0.13$  ( $z = 0.23$ ,  $p = .82$ ).

Automatisierung: Zur Bestimmung des Zusammenhangs zwischen der Dauer automatisierbarer Teilschritte und der Lösungswahrscheinlichkeit für die Aufgabe 'Job Search' wurden ebenfalls logistische Regressionsanalysen durchgeführt. Dabei wurde der Mittelwert für alle von einer Testperson innerhalb der Aufgabeneinheit vorgenommenen Bookmarking-Vorgänge als Prädiktor verwendet. Die Ergebnisse zeigen den erwarteten negativen Zusammenhang zwischen der Geschwindigkeit, mit der Bookmarks gesetzt werden und dem Lösungserfolg ( $BETA = -0.43$ ,  $z = 1.87$ ,  $p < .05$ ). Zur Kontrolle wurde außerdem die Zeit bestimmt, die insgesamt mit der Lösung der Aufgabe verbracht wurde. Für diese ergab sich kein Zusammenhang mit der Lösungswahrscheinlichkeit.

Insgesamt stützen die Ergebnisse die Annahme, dass bestimmte Teilschritte bei der Bewältigung von Problemlöseaufgaben kontrolliert erfolgen (müssen), und die Problemlösung von zusätzlich eingesetzten kognitiven Ressourcen profitiert. Andere Teilschritte dagegen sind Automatisierungsprozessen zugänglich. Wenn für solche Teilschritte kognitive Ressourcen mobilisiert werden (müssen), scheint dies dem Bearbeitungsergebnis eher abträglich. Implikationen dieser Befunde für die Struktur und Messung des Konstrukts 'Problemlösekompetenz in technologiereichen Umgebungen' werden diskutiert.

Fitts, P.M., & Posner, M.I. (1967). Learning and skilled performance in human performance. Belmont CA: Brock-Cole.

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (in press). The time on task effect in reading and problem solving is moderated by item difficulty and ability: Insights from computer-based large-scale assessment. Journal of Educational Psychology

Klieme, E. (2004). Assessment of cross-curricular problem-solving competencies. In J.H. Moskowitz & M. Stephens (eds.). Comparing Learning Outcomes. International assessments and education policy (pp. 81-107). London: Routledge Falmer.

Klieme, E., Leutner, D., & Wirth, J. (Hrsg.). (2005). Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie. Wiesbaden: VS Verlag für Sozialwissenschaften.

Mayer, R. E. (1994). Problem solving, teaching and testing for. In T. Husén & T. N. Postlethwaite (Eds.), The international encyclopedia of education (Vol. 8, 2n. ed., pp. 4728-4731). Oxford: Pergamon.

OECD. (2009). PIAAC problem solving in technology-rich environments: A conceptual framework. OECD Education Working Paper No. 36. Paris: OECD.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823

Esther Beltermann (Klinikum der LMU München, LMU München), Insa Wessels (Klinikum der LMU München, LMU München), Ingo Kollar (LMU München, Universität Augsburg), Martin R. Fischer (Klinikum der LMU München, LMU München)

## **Wodurch unterscheidet sich das Wissen von medizinischen Experten und Medizinstudierenden hinsichtlich des Ablaufs von Stationsvisiten: Eine Analyse internaler Skripts**

Einleitung: Die Durchführung von Stationsvisiten stellt eine zentrale, aber komplexe Tätigkeit des ärztlichen Arbeitsalltags im Krankenhaus dar. Dementsprechend ist von zentraler Bedeutung, dass Medizinstudierende Fertigkeiten hinsichtlich der kompetenten Durchführung von Stationsvisiten erwerben. Gerade junge Ärzte berichten allerdings Schwierigkeiten im Verständnis von und Handeln in Stationsvisiten (Norgaard et al., 2004). Es wird angenommen, dass sich mit zunehmender Expertise die Art und Organisation des Wissens sowie der Abruf von Informationen verändern und sich positiv auf das Handeln auswirken (Schmidt & Boshuizen, 1993). Ein viel versprechendes theoretisches Konstrukt, mit dessen Hilfe Expertiseunterschiede hinsichtlich der kompetenten Durchführung von Stationsvisiten sichtbar gemacht werden können, ist das Skriptkonzept (Schank, 1999). Skripts sind kognitive Gedächtnisstrukturen, die das Verstehen von und Handeln in komplexen Situationen erleichtern. Skripts beinhalten vier Gedächtniskomponenten (Fischer et al., 2013). Diese umfassen Wissen über: (a) die Situation selbst (Play), (b) die für die Situation typischen Phasen (Szenen), (c) die Handlungssequenzen innerhalb dieser Phasen (Scriptlets), sowie (d) die beteiligten Personen (Rollen). In dieser Studie sollen die visitenbezogenen Skripts von Mediznern unterschiedlichen Expertisestufen analysiert werden. Ziel ist, dabei Unterschiede zwischen den internalen Skripts von Medizinstudierenden und Mediznern auf höheren Expertisestufen zu identifizieren. Diese Erkenntnisse sollen in das Design potentiell effektiver Instruktionen zur Förderung der Visitenkompetenz von Medizinstudierenden eingehen.

Methode: Die Stichprobe umfasst 40 Personen unterschiedlicher Expertisestufen (N Student=14 N PJ=6, N Assistenzarzt=11, N Stationsarzt=5, N Oberarzt=4; 22 männlich, 18 weiblich; Alter: M=31.1, SD=10.12) der Inneren Medizin am Klinikum der LMU München. Auf Grundlage der Struktur-Ge-Technik von Scheele und Groeben (1988) wurden standardisierte Interviews hinsichtlich der bei den Teilnehmern vorhandenen Skriptkomponenten im Visitenkontext durchgeführt. Die im Interview genannten Komponenten wurden auf Karten notiert und unmittelbar im Anschluss an das Interview durch den Interviewten validiert. Die Akzeptanz der Interviewtechnik wurde mit einem Paper-Pencil-Fragebogen geprüft. Jede während des Interviews beschriftete Karte wurde auf Basis eines induktiv-deduktiv entwickelten Kodierschemas von zwei unabhängigen Ratern der passenden Skriptkomponente zugeordnet. Die Interraterübereinstimmung war über alle (Cohens Kappa =.93) und einzelne Skriptkomponenten hinweg (Cohens Kappa: Rolle: 1, Szene: .94, Scriptlet: .93, Operation: .92) sehr zufriedenstellend (Bortz & Döring, 2006). Gruppenunterschiede wurden mittels einfaktoriellen ANOVA nach Kruskal-Wallis, und Mann-Whitney Tests statistisch analysiert.

Erste Ergebnisse: Die Interviewtechnik wurde insgesamt sehr gut akzeptiert und von den Interviewees als geeignetes Verfahren zur Abbildung ihres Visitenverständnisses bezeichnet. Von den Interviewten wurden durchschnittlich 3.97 (SD=1.46, min :1, max :7) beteiligte Rollen genannt. Interviewte Studierende und PJ-Studierende nannten hierbei signifikant häufiger die Rolle Studierende als die anderen Befragten ( p=.026).

Insgesamt wurden im Mittel 35.72 Skriptkomponenten (SD=16.62) je Interview genannt. Durchschnittlich 4.18 (SD=1.73) davon entfallen auf die Szenenkomponente. Stationsärzte nannten signifikant ( $X^2=12.02$ ,  $df=4$ ,  $p=.027$ ) mehr Szenen (M=6, SD= 1.82) als Studierende (M=3.71, SD= 1.54), was entsprechend den Befunden der Expertiseforschung auf eine hierarchischere Wissensrepräsentation bei Experten gegenüber Novizen hinweist (Schmidt & Boshuizen, 1993). Zwischen allen anderen Gruppen wurden keine signifikanten Unterschiede hinsichtlich der Anzahl der genannten Szenen gefunden. Erste Analysen auf Scriptletebene verwiesen auf die Notwendigkeit, Operationen als eine weitere Analyseebene zu berücksichtigen, die eine Differenzierung von Handlungen unterschiedlicher Abstraktionsniveaus ermöglicht. Sowohl auf Scriptlet-, als auch auf Operationsebene wurden leichte bis moderate Unterschiede in der (absoluten und relativen) Anzahl genannter Komponenten gefunden. Dabei nannten Studierende sowie Ärzte höherer Ausbildungsstufen mehr Komponenten als PJ-Studierende und Assistenzärzte. Dieser u-förmige Zusammenhang zwischen Expertisegrad und der Ausdifferenzierung des Wissens wurde bereits in früheren Studien gefunden (Gruber, 2007). Expertenwissen zeichne sich jedoch im Gegensatz zu Novizen durch eine hohe Abstraktion und höhere Qualität der Inhalte aus (Gruber, 2007; Nievelstein et al., 2008). Tiefergehende Analysen, die sich insbesondere auf diese inhaltliche Ebene sowie die Scriptlets und Operationen unterschiedlicher Rollen beziehen, sind notwendig.

Bortz, J. & Döring, N. (2006). Forschungsmethoden und Evaluation. Heidelberg: Springer Verlag.

Fischer, F., Kollar, I., Stegmann, K. & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56.

Gruber, H. (2007). Bedingungen von Expertise. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (S. 93-112). Münster: Lit.

Nievelstein, F., Van Gog, T., Boshuizen, H. P. A. & Prins, F. J. (2008). Expertise-related differences in conceptual and ontological knowledge in the legal domain. *European Journal of Cognitive Psychology*, 20, 1043-1064.

Norgaard, K., Ringsted, C. & Dolmans, D. (2004). Validation of a checklist to assess ward round performance in internal medicine. *Medical Education*, 38(7), 700-707.

Schank, R. C. (1999). *Dynamic memory revisited*. Cambridge: Cambridge University Press.

Scheele, B. & Groeben, N. (1988). Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien: die Heidelberger Struktur-Ge-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen. Tübingen: Francke.

Schmidt, H. G. & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational psychology review*, 5(3), 205-221.



Johanna Hömberg (Rheinische Friedrich-Wilhelms-Universität Bonn), Sarah Potthoff (Rheinische Friedrich-Wilhelms-Universität Bonn), Ronja Torspecken (Rheinische Friedrich-Wilhelms-Universität Bonn), Udo Käser (Rheinische Friedrich-Wilhelms-Universität Bonn)

## **„Piratenreise mit Kapitän Katego“: Evaluation eines Konzepts zur Förderung der kategorialen Organisationskompetenz für Schülerinnen und Schüler der dritten Klasse**

Aktuelle Befunde der empirischen Bildungsforschung machen deutlich, dass es um die Vermittlung metamemorialer Kompetenzen in der Grundschule nicht gut bestellt ist. So konnten Käser und Cummings (2012) in einer experimentellen Studie zeigen, dass überhaupt nur etwa zwei von drei Schülerinnen und Schüler der dritten bis fünften Klasse in expliziten Lernsituationen angeben, beim Lernen von Stimuli strategisch vorzugehen. Einen automatisierten Einsatz von Gedächtnisstrategien zeigt nur etwa jede fünfte Schülerin bzw. Schüler. Die Strategie des kategorialen Organisierens wird selbst in instruktionalen Lernsituationen sogar nur von ca. jeder zehnten Schülerin / Schüler angewendet. Gleichzeitig sind jedoch diejenigen Schülerinnen und Schüler, welche metamemorale Strategien wie Naming, Rehearsal oder kategoriales Organisieren nutzen, im Lernprozess besonders erfolgreich. Vor allem ermöglicht der Einsatz der Gedächtnisstrategie des kategorialen Organisierens eine deutlich verbesserte Lernleistung.

Diese Ergebnisse, welche die Bedeutung des Beherrschens der kategorialen Organisationsstrategie für erfolgreiches Lernen unterstreichen, stimmen gut mit kognitionspsychologischen Befunden überein (zum Beispiel Hasselhorn & Labhuhn, 2008). Überraschend und aus pädagogisch psychologischer Perspektive ernüchternd ist jedoch der Befund, dass Kinder selbst unter expliziten Lernbedingungen wirksame Organisationsstrategien nur selten verwenden, da bereits Kinder im Alter von ca. neun Jahren über Gedächtnisstrategien wie das kategoriale Organisieren verfügen und frühere Defizite wie das Mediations-, Produktions- oder Nutzungsdefizit überwunden haben können (Krajewski, Kron & Schneider, 2004, Lockl & Schneider, 2007).

Insofern sind Maßnahmen zur Förderung metamemorialer Kompetenzen in der Grundschule unverzichtbar. Um ein solches Förderkonzept handelt es sich bei dem Programm „Piratenreise mit Kapitän Katego“ für Schülerinnen und Schüler der dritten Klasse, das vorgestellt wird und für das erste Ergebnisse hinsichtlich der Wirksamkeit des Programms auf Basis eines Pre-Post-Vergleichs der Gedächtnis- und der kategorialen Organisationsleistung für eine Fördergruppe von Drittklässlern vorgestellt und diskutiert werden.

Hierbei begeben sich die Teilnehmer bei diesem Programm auf eine Phantasie- und spielerische Reise und begleiten spielerisch die pädagogischen Agenten Kapitän Katego, das Piratenmädchen Mema und den Papageien Repetito auf ihrer Piratenfahrt. In sechs Sitzungen, in denen die Teilnehmer individuell an Lernmappen mit Übungen zur Typizität unterschiedlicher Kategorien und zur Assoziativität ihrer Elemente arbeiten, wird die Strategie des kategorialen Organisierens systematisch trainiert.

Die bisherigen Erfahrungen sind sehr positiv und weisen auf eine hohe Wirksamkeit der Fördermaßnahme hin: Einerseits verdoppelt sich die Gedächtnisleistung der Teilnehmer vom Pre- zum Posttest nahezu. Andererseits zeigt sich für den Zusammenhang zwischen Gedächtnisleistung und kategorialer Organisationsleistung, für den zum ersten Messzeitpunkt eine Nullkorrelation vorliegt, zum zweiten Messzeitpunkt ein überzufälliger Zusammenhang ( $r = .602$ ). Der Unterschied zwischen den Korrelationen für die Pre- und die Postmessung wird signifikant.

Diskutiert werden Möglichkeiten, das Förderkonzept im Grundschulunterricht nutzbar zu machen. Insbesondere werden Perspektiven beleuchtet, die Maßnahme konzeptionell als Baustein im Regelunterricht zu verwenden.

Hasselhorn, M. & Labhuhn, A. S. (2008). Metakognition und selbstreguliertes Lernen – Metacognition and Self-regulated Learning. In W. Schneider & M. Hasselhorn (Hrsg.), Handbuch der Pädagogischen Psychologie. Band X (S. 28-37). Göttingen: Hogrefe.

Käser, U. & Cummings, A. V. (2012). Gedächtnisleistung und metamemorale Organisationsleistung in Abhängigkeit vom Modus des Lernens. In U. M. Röhr-Sendlmeier (Hrsg.), Inzidentelles Lernen – wie wir beiläufig Wissen erwerben (S. 147-194). Berlin: Logos.

Krajewski, K., Kron, V. & Schneider, W. (2004). Entwicklungsveränderungen des strategischen Gedächtnisses beim Übergang vom Kindergarten in die Grundschule. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 36, 47–58.

Lockl, K. & Schneider, W. (2007). Entwicklung von Metakognition. In M. Hasselhorn (Hrsg.), Handbuch der Entwicklungspsychologie. Band 7 (S. 255–265). Göttingen: Hogrefe.



Sielle Phelan (TU München), Inga Specht (TU München), Rahel Grüninger (Universität Koblenz-Landau), Wolfgang Schnotz (Universität Koblenz-Landau), Doris Lewalter (TU München)

## **Konfligierende Wissenschaftsthemen im Museum: Einfluss von Präsentationsmodalität und Besuchermerkmalen auf die Verarbeitung**

Was im Museumskontext mit passiver Bewunderung für naturwissenschaftliche und technologische Erfolge begann und sich ab den 60er Jahren hin zu einem aktiven Lernen entsprechender Prinzipien veränderte, hat sich in den letzten Jahren zunehmend zu einer kritischen Diskussion gegenwärtiger Forschungsprozesse und -befunde weiterentwickelt (Durant, 2004). Diese Entwicklung stellt sowohl Museen als auch Museumsbesucher vor eine Herausforderung, denn nicht mehr gesichertes Wissen sondern fragile und oft konfligierende Erkenntnisse stehen im Zentrum der Präsentation sowie der Diskussion. Die vorliegende Studie soll dazu beitragen genauere Einblicke in den Einfluss der Präsentation entsprechender Information anhand von Ausstellungstexten auf die Konfliktverarbeitung zu gewinnen. Die zentralen Fragen der Studie beziehen sich dabei auf die Auswirkung der Präsentationsmodalität und personale Merkmale auf die kognitive Konfliktverarbeitung und die Einstellung der Besucher.

Die quasi-experimentelle Feldstudie mit Prä- und Postfragebogen wurde in zwei technisch-naturwissenschaftlichen Museen mit deutschsprachigen Besuchern ab einem Alter von 15 Jahren durchgeführt (n= 553). Für die Studie wurden zwei unterschiedliche Ausstellungstexte zum Thema Bionik mit vergleichbarem Komplexitätsgrad und Länge entwickelt und in einer bestehenden Ausstellung integriert. Jeder Text enthielt zwei konfligierende Positionen zum jeweiligen Thema. Die Modalität der Informationspräsentation wurde variiert indem visuell präsentierte Ausstellungstexte akustisch präsentierten Audioguides gegenüber gestellt wurden.

Im Präfragebogen wurden für die Verarbeitung konfligierende Informationen als relevant erachtete personale Merkmale wie z.B. epistemologische Überzeugungen (Conley, Pintrich, Vekiri & Harrison, 2004) Ambiguitätstoleranz (Kruglanski, Webster & Klem, 1993) und need for cognition (Caccioppo, Petty & Kao, 1984) erfasst. Als Kontrollvariablen wurden Vorwissen und Verbale Intelligenz (Anger, Mertesdorf, Wegner & Wülfing, 1980) aufgenommen. Die Indikatoren der Konfliktverarbeitung wurde unter anderem auf der Basis theoretischer Überlegungen und empirischer Befunde zur Forschung zu Textverstehen und Inferenzbildung (McNamara & Kintsch, 1996; Schnotz, 1994; Rickheit & Strohner, 1995) und in Anlehnung an Schmalhofer und Glavanov (1986) sowie Stadtler, Scharrer und Bromme (2011) entwickelt. Außerdem wurde sowohl im Prä- als im Postfragebogen die Einstellung zum Thema erfasst. Der Beitrag stellt zum einen Befunde zum Einfluss der Modalität der Informationsdarbietung auf die kognitive Konfliktverarbeitung der Museumsbesucher dar und behandelt zum anderen die Wirkung der genannten Personenmerkmale auf die Verarbeitung konflikthafter Information. Ebenso werden mögliche Einflüsse der Modalität und der Personenmerkmale auf die Einstellung der Museumsbesucher zur dargestellten Thematik unter Kontrolle des Vorwissens überprüft. Die Befunde werden in Hinblick auf ihre theoretische Bedeutung und praktische Implikationen diskutiert.

- Anger, H., Mertesdorf, F., Wegner, R. & Wülfing, G. (1980). Verbaler Kurz-Intelligenztest (VKI / Form A). Beltz Test.
- Durant, J. (2004). The challenge and the opportunity of presenting 'unfinished science'. D. Chittenden et al, 47-60.
- Caccioppo, J.T., Petty, R.E., & Kao, C.F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48, 306-307.
- Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186-204.
- Kruglanski, A.W., Webster, D.M., & Klem, A. (1993). Motivated Resistance and Openness to Persuasion in the Presence or Absence of Prior Information. *Journal of Personality and Social Psychology*, 65(5), 861-876.
- McNamara, D.S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, 22(3), 247-288.
- Rickheit, G., & Strohner, H. (Eds.) (1995). *Inferences in text processing*. Amsterdam: Elsevier
- Schmalhofer, F., & Glavanov, D. (1986). Three Components of Understanding a Programmer's Manual: Verbatim, Propositional, and Situational Representations. *Journal of Memory and Language*, 25, 279- 294.
- Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten*. Weinheim: Beltz.
- Stadtler, M., Scharrer, L., & Bromme, R. (2011). How Reading Goals and Rhetorical Signals Influence Recipients' Recognition of Intertextual Conflicts. In L. Carlson, C. Hoelscher & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1346-1351). Austin, TX: Cognitive Science Society

Katrin Böhme (IQB, Humboldt-Universität zu Berlin), Sebastian Weirich (IQB, Humboldt-Universität zu Berlin), Nicole Haag (IQB, Humboldt-Universität zu Berlin), Hans Anand Pant (IQB, Humboldt-Universität zu Berlin)

## **Wie wirken sich Veränderungen in der Aufgabengestaltung auf die Itemkennwerte aus? Layoutmanipulationen in bildungsstandardbasierten Lesekompetenzaufgaben**

Die Durchführung regelmäßig administrierter, großer Schulleistungsstudien wie bspw. der Vergleichsarbeiten in der dritten Jahrgangsstufe (VERA-3) basiert nach wie vor auf Paper-Pencil-Tests. Da an diesen flächendeckenden Studien bundesweit jährlich etwa 650.000 Schülerinnen und Schüler teilnehmen, bedeutet dies einen extremen Papierverbrauch und enorme Druck- sowie Kopierkosten. Auf Seiten der Bildungsadministration besteht daher der Wunsch, diese Kosten so weit wie möglich zu reduzieren. Dies wäre bspw. durch ein platzsparendes Layout der Aufgaben möglich. Für VERA-3 2014 wurde ein solches platzsparendes Layout für Aufgaben des Kompetenzbereichs Lesen entwickelt.

Im Rahmen anderer großer Schulleistungsstudien hat sich allerdings gezeigt, dass bereits geringfügige Veränderungen der Aufgabengestaltung, -reihenfolge, -kombination oder -position Auswirkungen auf die Itemkennwerte und somit auch auf die Fähigkeitsschätzungen der Schülerinnen und Schüler haben. So wurden bspw. im Rahmen der sogenannten 1985-86 NAEP Reading Anomaly bei Long Term Trend (LTT)-Analysen unplausible Veränderungen der Leseleistungen von Kindern und Jugendlichen identifiziert (Beaton, 1988; Zwick, 1992). Hierbei handelte es sich um Effekte von bis zu  $d = .22$ , die keine wahren Fähigkeitsunterschiede abbildeten, sondern ein Artefakt aufgrund von Veränderungen im Testdesign und der Testadministration darstellten (Mazzeo & von Davier, 2008). Als ein ursächlicher Faktor wurden Veränderungen im Layout der eingesetzten Lesetexte sowie im Antwortformat geschlossener Items identifiziert (Beaton, 1988; Mazzeo & von Davier, 2008). Die aus diesem Befund abgeleitete Empfehlung lautete daher: „When measuring change, do not change the measure“ (Beaton, 1990, S. 10). Dies gilt insbesondere für Verlinkungskontexte, in denen nach Möglichkeit vollständig identische Testhefte eingesetzt werden sollten (Mazzeo & von Davier, 2008).

Um die im Kontext von VERA-3 erfolgende Anbindung an die Bildungsstandardmetrik sowie unverfälschte Fähigkeitsschätzungen zu gewährleisten, wurde daher vor dem Einsatz des platzsparenden Layouts in der VERA-Testung 2014 geprüft, ob die Veränderung des Layouts zu systematischen Veränderungen der Itemkennwerte führt. Hierfür wurde das platzsparende Layout der Leseaufgaben im Jahr 2013 in einer großen Pilotierungsstudie an 3680 Schülerinnen und Schülern der dritten Jahrgangsstufe erprobt. In dieser Studie wurden in rotierten Testheften 8 Leseaufgaben mit insgesamt 94 Items sowohl im konventionellen als auch im platzsparenden Layout eingesetzt, um Veränderungen in den Itemkennwerten inhaltlich identischer Items untersuchen zu können. Die Layoutveränderungen betrafen konkret eine Verringerung der Schriftgröße, Zeilen- und Absatzabstände sowie die platzsparendere, horizontale Anordnung der Optionen in MC-Items. Um mögliche Effekte zu identifizieren, wurden verschiedene deskriptive und explanatorische IRT-Modelle im Programm R berechnet.

Es zeigte sich kein Haupteffekt des verkürzten Layouts, allerdings konnte ein spezifischer Effekt der Anordnung der Optionen in MC-Items identifiziert werden: MC-Items, bei denen die MC-Optionen platzsparend horizontal angeordnet sind, sind schwerer als inhaltlich identische MC-Items mit konventionell vertikal angeordneten Optionen. Die Trennschärfe der Items wurde durch die Variation des Layouts nicht signifikant beeinträchtigt.

Diese Befunde sprechen dafür, dass sich die veränderte Anordnung der Optionen in MC-Items auf die Itemkennwerte und damit auf die resultierenden Fähigkeitsschätzungen bzw. die Anbindung der Items an die Metrik der Bildungsstandards auswirken kann. Der Kostenvorteil eines platzsparenden Layouts muss daher gegen die aus den Veränderungen resultierenden inhaltlichen Konsequenzen abgewogen werden.

Beaton, A. E. (1988). The NAEP 1985-86 Reading Anomaly: A Technical Report. Princeton, NJ: Educational Testing Service

Beaton, A. E. (1990). Introduction. In A. E. Beaton & R. Zwick, The effect of changes in the national assessment: Disentangling the NAEP 1985-86 reading anomaly (Report No. 17-TR-21). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mazzeo, J. & von Davier, M. (2008). Review of the Programme for International Student Assessment (PISA) test design: Recommendations for fostering stability in assessment results. Zugriff am 31.10.2013 unter [https://edsurveys.rti.org/pisa/documents/mazzeopisa\\_test\\_designreview\\_6\\_1\\_09.pdf](https://edsurveys.rti.org/pisa/documents/mazzeopisa_test_designreview_6_1_09.pdf)

Zwick, R. (1992). Statistical and psychometric issues in the measurement of educational achievement trends: Examples from the National Assessment of Educational Progress. Journal of Educational and Behavioral Statistics, 17, 205-218.



## **Automatische Auswertung offener, natürlichsprachlicher Antworten in PISA-Testaufgaben**

Natürlichsprachliches, offenes Antwortformat in Testaufgaben kann in einigen Fällen gegenüber geschlossenen Formaten erstrebenswert für die Konstruktvalidität sein (vgl. Bridgeman, 1993; Kubinger & Gottschall, 2007; Millis, Magliano, Wiemer-Hastings, Todaro, & McNamara, 2011; Rauch & Hartig, 2010; Rost & Sparfeldt, 2007; Rupp, Ferne, & Choi, 2006). Jedoch bringt das offene Format beim Einsatz menschlicher Kodierer und Kodierinnen gleichzeitig eine Reduktion der Objektivität sowie einen hohen Aufwand mit sich. Auch etwa eine große Errungenschaft der Psychodiagnostik, das adaptive Testen, lässt sich in Kombination mit offenem Antwortformat aktuell schwierig umsetzen. In den meisten Fällen kann es nur mit menschlichen Testleitern und als branched statt tailored testing stattfinden.

Daher ist die nächste zu überwindende Hürde für den Fortschritt dieses Bereichs die Entwicklung einer Methodik zur automatischen Kodierung offenen, natürlichsprachlichen Antwortformats. Für lange Testpersonen- bzw. Schülerantworten wurden, speziell zur Aufsatzbeurteilung, in den letzten drei Jahrzehnten Methoden und Technologien vorgeschlagen (etwa die Latente Semantische Analyse; Deerwester, Dumais, Furnas, & Landauer, 1990), die bereits teilweise Anwendung in der Praxis gefunden haben. Auch für Kurztextantworten liegen inzwischen erste Ansätze und Systeme vor (z. B. c-rater[TM]; Leacock & Chodorow, 2003), die allerdings noch weniger etabliert sind.

Das zu präsentierende Projekt soll durch interdisziplinäre Synergie aktueller Technologien und Methoden aus dem Natural Language Processing-Bereich, aber auch aus der Psychometrie eine Methodik entwickeln, mittels welcher kurze, offene, natürlichsprachliche Antworten der Schülerinnen und Schüler in Deutschland aus der PISA-Stichprobe 2012 automatisch kodiert werden können. Ein Teil der für den gesamten Datensatz vorliegenden menschlichen Kodierungen dient dabei in einem ersten Schritt dem Training des entwickelten Systems, der andere Teil anschließend dessen Validierung im Sinne der menschlichen Kodierung als Außenkriterium (welche den gold standard darstellt). Die Studie verfolgt an dieser Stelle also die Frage, wie akkurat es gelingt, die durch Menschen vorgenommene Kodierung am Computer zu modellieren. Als weitergehende Fragestellung soll überprüft werden, inwieweit andere Materialien wie die Kodier- Richtlinien für die menschlichen Kodierinnen und Kodierer genutzt werden können, um das System zu trainieren, ohne die menschlichen Kodierungen als Trainingsdaten zu verwenden. Auf diese Weise würde ein rein kriteriumsbasierter Ansatz gegenüber dem vielfältig beeinflussten menschlichen Kodierprozess als Beurteilungs-Definition gewählt werden, was als richtig und was als falsch betrachtet wird.

Auf dem Poster werden der Entwicklungsstand sowie erste empirische Befunde präsentiert. Konkurrierende Methoden wie etwa die Latente Semantische Analyse (Deerwester et al., 1990) und die Explizite Semantische Analyse (Gabrilovich & Markovitch, 2009) werden anhand der empirischen Daten auf ihren Beitrag zur Klassifizierungsgüte verglichen. Als Grundlage der Sprachverarbeitung dienen für diese Methoden aus dem Internet gewonnene Textkorpora, u. a. die deutsche Wikipedia. Es werden außerdem die entstehenden Herausforderungen aufgezeigt, die vom System gelöst werden müssen; so etwa die Unreinheit des textlichen Materials (z. B. Rechtschreibfehler) und die Quantifizierung des semantischen Konzepts einer Antwort. Es werden vorläufige Koeffizienten zur Klassifizierungsgüte des Systems präsentiert. Für die präsentierten Befunde dienen Antworten auf 18 papierbasiert administrierte Lese-Items als Basis mit  $n = 4223$ ,  $4234$  oder  $4152$ , je nach Einsatz des entsprechenden Items im von PISA verwendeten Booklet-Design. Antworten werden als entweder richtig oder falsch klassifiziert, bei einem Item auch als teilweise richtig.

- Bridgeman, B. (1993). A comparison of open-ended and multiple-choice question formats for the quantitative section of the Graduate Record Examinations general test. Princeton and N.J.: Educational Testing Service.
- Deerwester, S., Dumais, S. T., Furnas, G. W., & Landauer, T. K. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41 (6), 391–407.
- Gabrilovich, E., & Markovitch, S. (2009). Wikipedia-based semantic interpretation for natural language processing. *Journal of Artificial Intelligence Research*, 34, 443–498.
- Kubinger, K. D., & Gottschall, C. H. (2007). Item difficulty of multiple choice tests dependant on different item response formats – An experiment in fundamental research on psychological assessment. *Psychology Science*, 49 (4), 361–374.
- Leacock, C., & Chodorow, M. (2003). C-rater: Automated scoring of short-answer questions. *Computers and the Humanities*, 37, 389–405.
- Millis, K., Magliano, J., Wiemer-Hastings, K., Todaro, S., & McNamara, D. S. (2011). Assessing and improving comprehension with Latent Semantic Analysis. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Hrsg.), *Handbook of Latent Semantic Analysis* (S. 207–225). Erlbaum: Mahwah.
- Rauch, D. P., & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items: A two-dimensional IRT analysis. *Psychological Test and Assessment Modeling*, 52 (4), 354–379.
- Rost, D. H., & Sparfeldt, J. R. (2007). Leseverständnis ohne Lesen? *Zeitschrift für Pädagogische Psychologie*, 21 (3), 305–314. doi: 10.1024/1010-0652.21.3.305
- Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiplechoice questions shapes the construct: a cognitive processing perspective. *Language Testing*, 23 (4), 441–474. doi: 10.1191/0265532206lt3370a



## **Wie stabil sind Kompetenzmessung bei Vergleichsarbeiten?**

Als eine der vier Säulen der Gesamtstrategie zum Bildungsmonitoring der KMK werden seit einigen Jahren in fast allen Ländern Deutschlands Vergleichsarbeiten in der 3. und 8. Jahrgangsstufe durchgeführt. Speziell für die 8. Jahrgangsstufe werden vom IQB seit 2009 Testhefte in drei schwierigkeitsgestuften Versionen angeboten, die speziell auf die Leistungserwartung von Schülerinnen und Schüler ausgerichtet sind, die a) einen Hauptschulabschluss, b) einen Mittleren Schulabschluss oder c) das Abitur anstreben. Die Ergebnisse sollen die Leistungen einzelner Schulen und Klassen widerspiegeln, sich dabei an einem kriteriellen Maß – den Kompetenz(stufen)modellen der Bildungsstandards – orientieren und im Vergleich über mehrere Jahre Entwicklungen aufzeigen. Durch die verwendete Rasch-Skalierung und die konsequente Verlinkung der Tests mit der Metrik der Bildungsstandards durch das IQB, wird unterstellt, dass die Ergebnisse über Testheftversionen wie Jahre verglichen werden können (Köller, 2008). In Berlin und Brandenburg wird jährlich auf der Basis der Vorjahresergebnisse und der Ergebnisse der Pilotierung diskutiert und zentral festgelegt, welche Testversionen den unterschiedlichen Schulformen, ggf. differenziert nach Unterrichtsniveaus vorgelegt werden. Erste Gegenüberstellungen deuten an, dass sich die erwartete Stabilität nur unzureichend einstellt, was zu konkreten Fragestellungen führt: Wie stabil sind die Messungen über die Jahre? Stellen sich diese Ergebnisse für unterschiedliche Schulformen identisch dar? Ist das festgestellte Leistungsniveau tatsächlich unabhängig von der Schwierigkeit des Testhefts? Wie wirken sich Abweichungen der Schwierigkeit einzelner Items in der Haupterhebung von jener der Pilotierung aus? Verbessert die freie Schätzung einzelner von den Pilotierungswerten abweichenden Items die Messung? Es wird untersucht, in wieweit Variationen bei der Skalierung zu stabileren Ergebnissen führen. Für das Fach Mathematik liegen aus Berlin verpflichtende Vollerhebungen aller Achtklässler aus fünf aufeinanderfolgenden Jahren von 2009 bis 2013 vor.

Hambleton, R. K., & Swaminathan, H. (1989). Item Response Theory. Principles and applications. Boston, MASS: Kluwer Nijhoff Publishing.

Köller, Olaf. (2008). Bildungsstandards in einem Gesamtsystem der Qualitätssicherung im allgemeinbildenden Schulsystem Deutschlands. In E. Klieme & R. Tippelt (Eds.), Qualitätssicherung im Bildungswesen. (pp. 59-75). Weinheim u.a.: Beltz.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## **Goal-Framing in der pädagogisch-psychologischen Diagnostik: Eine systematische Untersuchung von Instruktionseffekten**

Theoretischer Hintergrund: Die klassischen Gütekriterien für Tests in der pädagogisch-psychologischen Diagnostik (insb. Objektivität, Reliabilität, Validität) verlangen unter anderem, dass die Performanz eines Probanden nicht von äußeren Bedingungen der Testdurchführung sondern nur von der individuellen Merkmalsausprägung abhängt (Durchführungsobjektivität; Moosbrugger & Kelava, 2008). Eine Manipulation des Framings (Rahmung) einer Testdurchführung kann folglich als Verletzung der Durchführungsobjektivität betrachtet werden (Brandstätter & Mücke, 2009).

In der Framing-Theorie wird zwischen risky choice-, attribute- und goal-Framing unterschieden, wobei bei letzterem ein übereinstimmendes Ziel (z. B. das Erreichen eines bestimmten Testscores) unterschiedlich, z. B. negativ (Versuche, nicht mehr als X% der Aufgaben falsch zu beantworten.) oder positiv (Versuche, mindestens Y% der Aufgaben richtig zu beantworten.), geframet wird (Levin, Schneider, & Gaeth, 1998). Roney und Lehman (2008) haben gezeigt, dass negatives goal-Framing zu einer geringeren Testleistung führen kann als positives.

Fragestellung: In dieser Studie wird untersucht, inwieweit eine systematische Manipulation der Testdurchführung (negatives, positives und neutrales goal-Framing nach Roney & Lehman, 2008) einen Effekt auf die Leistung von Schülerinnen und Schülern bei Tests zum Fachwissen in der Biologie hat:

Inwieweit treten bei Tests für den Kompetenzbereich Fachwissen im Fach Biologie Effekte des Goal-Framings auf?

Hypothese: Probanden in einer negativ geframeten Testsituation erzielen einen geringeren Testscore als Probanden in einer neutral oder positiv geframeten Testsituation (Roney & Lehman, 2008).

Methode: Es wurden insgesamt 20 MC-Aufgaben zum biologischen Fachwissen der American Association for the Advancement of Science (AAAS, 1993) übersetzt und in 20 Testheften angeordnet (Repeated Treatment Design; Frey, Hartig, & Rupp, 2009).

In einem between subject-Design werden Schülerinnen und Schüler der Klassen 7 bis 10 jeweils mit einem positiven, negativen oder neutralen Frame instruiert. Als Kontrollvariablen werden die Schulnoten in den naturwissenschaftlichen Fächern, das Geschlecht der Probanden und der Cognitive Load (CL) während der Testbearbeitung erhoben (5 Items mit 5-stufiger Rating-Skala; vgl. Nehring, Nowak, Upmeyer zu Belzen, & Tiemann, 2012).

Ergebnisse: Bisherige Befunde (N = 175) legen nahe, dass sich die Lösungswahrscheinlichkeit zwischen der positiven (MS = .43), der negativen (MS = .42) und der neutralen (MS = .45) Framing-Gruppe nicht signifikant unterscheidet. Allerdings geben die Probanden bei positivem Framing einen (marginal) signifikant höheren CL an (MS = 15.54; SE = 0.42) als bei negativem Framing (MS = 14.07; SE = 0.62;  $t(98) = 1.99$ ;  $p < .10$ ). Der CL zeigt sich in einer linearen Regression als signifikanter Prädiktor des Testscores, der etwa 6 % der Varianz erklärt ( $\eta^2 = .06$ ).

Fazit: Die Befunde sprechen insgesamt dafür, dass sich das Framing nicht direkt aber indirekt über den CL auf den Testscore der Probanden auswirkt. Es ergibt sich ein höherer CL in der positiv geframeten Gruppe. Dies widerspricht tendenziell den Befunden von Roney et al. (2008).

Aktuell wird die Studie fortgesetzt, um auf der GEBF-Tagung Ergebnisse einer größeren Stichprobe präsentieren zu können. Die Leistung der Probanden wird dann in einem eindimensionalen Rasch-Modell geschätzt (Wu, Adams, Wilson, & Haldane, 2007). Es wird untersucht, inwieweit die Framing-Manipulationen spezifische Effekte bei Jungen und Mädchen (vgl. Hannover, 2004) und bei Schülerinnen und Schülern mit unterschiedlichen Leistungen in den naturwissenschaftlichen Fächern (Schulnoten) zeigen (vgl. Locke & Latham, 2002). Die Befunde tragen dazu bei, Instruktionseffekte in der pädagogisch-psychologischen Diagnostik theoretisch erklären und weitergehend untersuchen zu können. Es deutet sich an, dass bei einer Testdurchführung nicht nur auf eine konstante Instruktion geachtet werden sollte (Durchführungsobjektivität). Auch die Qualität der Instruktion (z. B. positives, negatives, neutrales Framing) sollte hinsichtlich ihres Effekts auf die kognitive Belastung sowie die Testleistung von Schülerinnen und Schülern berücksichtigt werden.

AAAS (1993). Benchmarks for Science Literacy, Project 2061. New York, NY: Oxford U.P.;

Brandstätter, E., & Mücke, R. (2009). Interpreting test results. *Personality and Individual Differences*, 46, 183–186.;

Frey, A., Hartig, J., & Rupp, A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice*, 28, 39–53.;

Hannover, B. (2004). Gender revisited: Konsequenzen aus PISA für die Geschlechterforschung. In D. Lenzen, J. Baumert, R. Waternann, & U. Trautwein (Eds.), *PISA und die Konsequenzen für die erziehungswissenschaftliche Forschung* (pp. 81–100). Wiesbaden: VS.;

Levin, I., Schneider, S., & Gaeth, G. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76, 149–188.;

Locke, E., & Latham, G. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57, 705–717.;

Moosbrugger, H., & Kelava, A. (2008). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 7–26). Heidelberg: Springer.;

Nehring, A., Nowak, K., Upmeyer zu Belzen, A., & Tiemann, R. (2012). Doing inquiry in chemistry and biology: The context's influence on the students' cognitive load. *La Chimica nella Scuola*, XXXIV, 253–258.;

Roney, C., & Lehman, D. (2008). Self-regulation in goal striving: Individual differences and situational moderators of the goal-framing/performance link. *Journal of Applied Social Psychology*, 38, 2691–2709.;

Wu, M. L., Adams, R., Wilson, M., & Haldane, S. (2007). *ACER ConQuest*. Camberwell, Vic: ACER Press.





## **Erfassung von motorischen Basisqualifikationen (MOBAQ) zur Analyse von Wirkungen im Sportunterricht – Ergebnisse der IMPEQT-Studie**

Theoretischer Hintergrund: Im Rahmen der nationalen „Evaluationskultur im Bildungswesen“ (EDK, 2008) wird allen Schulfächern die Frage gestellt, welche fachspezifischen Kompetenzen aufgebaut werden. Die Sportwissenschaft muss demnach empirische Befunde für das Erreichen der Bildungsziele im Schulfach Sport liefern sowie sich um die Erfassung und Modellierung von sportspezifischen Kompetenzen bemühen. Zentrales Bildungsziel des Schulsports ist die Vermittlung von Motivation und Kompetenz zum lebenslangen Sporttreiben (im Überblick: Balz, 2009). Durch die Vermittlung von motorischen Basisqualifikationen sollen die Schüler/-innen zu einer kritisch-konstruktiven Teilhabe am Kulturgut Bewegung, Spiel und Sport befähigt und der Zugang zu weiteren Bildungsgegenständen wahrscheinlicher gemacht werden.

Methode: Das einzige Testinstrument, das bildungstheoretisch hergeleitet, sich in Kompetenzen beschreiben lässt und in Form von empirisch prüfbaren Testaufgaben einsetzbar ist, ist der MOBAQ-Ansatz (Kurz, Fritz & Tscherbel, 2007). Die Testaufgaben erfassen motorische Basisqualifikationen und beinhalten normativ festgelegte Mindeststandards. Die zentralen Fragen lauten: (1) Welche Struktur bilden diese Testaufgaben ab? (2) Kann die Struktur konfirmativ bestätigt werden? (3) Ist die Struktur des MOBAQ-Testinstruments über die Zeit stabil?

Im Zuge der von der Eidgenössischen Sportkommission (ESK) geförderten IMPEQT-Studie (Implementation of Physical Education and the Quality of Teaching) wurden 890 bzw. 837 Schülern/-innen (Alter t1: M = 13.2 Jahre) zu zwei Messzeitpunkten mittels zwölf ordinal skalierten MOBAQ-Testaufgaben in den Bewegungsbereichen „Laufen, Springen, Werfen“, „Bewegungskünste“ und „Ballspiele“ erfasst. Aufgrund der Skalierung wurden die explorativen, konfirmatorischen und autoregressiven Strukturgleichungsmodelle in Mplus mittels WLSM-Schätzer berechnet:

(1) Mit der t1-Stichprobe wurde eine explorative Faktorenanalyse durchgeführt, um die Struktur der Testaufgaben abzubilden. (2) Mit den Gesamtstichproben zu beiden Messzeitpunkten wurden konfirmatorische Faktorenanalysen berechnet, um die Modellpassung und die psychometrischen Qualitätskriterien zu kontrollieren. (3) In einem autoregressiven Strukturgleichungsmodell wurde die zeitliche Invarianz überprüft.

Ergebnisse: (1) Nach Ausschluss einer Testaufgabe zeigt die explorative Faktorenanalyse eine zweifaktorielle Struktur. Sieben Testaufgaben laden auf den ersten Faktor und beinhalten Ganzkörperbewegungen (z.B. Hangeln, Hindernislauf), vier Testaufgaben laden auf den zweiten Faktor und beinhalten Bewegungen mit dem Ball (z.B. Prellen, Passen). Die Modellfits ( $p = .80$ ; CFI = 1.00; RMSEA = .00; SRMR = 0.37) sind als sehr gut zu bezeichnen.

(2) Eine Replikation dieser zweifaktoriellen Struktur unter restriktiveren Bedingungen im Rahmen der konfirmatorischen Faktorenanalyse bestätigt zu beiden Messzeitpunkten diese Struktur mit guten Modellfits (t1: CFI = .98; TLI = .97; RMSEA = .030; WRMR = .87; t2: CFI = .98; TLI = .97; RMSEA = .033; WRMR = .87).

(3) Das autoregressive Strukturgleichungsmodell mit faktorieller Invarianz erreicht eine sehr gute Passung auf die Längsschnittdaten (CFI = .98; TLI = .98; RMSEA = .021; WRMR = .97), woraus gefolgert werden kann, dass die Struktur der MOBAQ-Testitems nach einem Jahr in gleicher Form bestand hat. Die Autoregression zwischen den latenten Faktoren „Ganzkörperbewegungen“ beträgt .93, zwischen den latenten Faktoren „Bewegungen mit dem Ball“ .84. Diese hohen Werte weisen darauf hin, dass eine hohe Stabilität interindividueller Unterschiede über die Zeit vorliegt.

Diskussion: Die entwickelte Testbatterie aus MOBAQ-Testaufgaben genügt weitestgehend den psychometrischen Testkriterien und ist damit ein geeignetes Instrument, um Effekte von Schulsport zu bestimmen. Es ist für die Zukunft zu empfehlen, die Testbatterie vor allem im Bereich Ballspiel zu erweitern und den Bereich Ganzkörperbewegungen ausdifferenzieren.

EDK (2008). Evaluation und Schulqualität. Orientierungsrahmen zuhanden von Bildungsbehörden, Aufsichtsorganen und Schulleitungen. Biel: Ediprim.

Balz, E. (2009). Fachdidaktische Konzepte update oder: Woran soll sich der Schulsport orientieren? *sportpädagogik*, 33 (1), 25–32.

Kurz, D., Fritz, T. & Tscherpel, R. (2008). Der MOBAQ-Ansatz als Konzept für Mindeststandards für den Sportunterricht? In V. Oesterheld, J. Hofmann, M. Schimanski, M. Scholz & H. Altenberger (Hrsg.), *Sportpädagogik im Spannungsfeld gesellschaftlicher Erwartungen, wissenschaftlicher Ansprüche und empirischer Befunde* (S. 97–106). Hamburg: Czwalina.

## Vertrauen als Kernaspekt digitaler Kompetenz – Implikationen für Bildung

Chair(s): Stephanie Pieschl, Ricarda Moll (beide: Westfälische Wilhelms-Universität Münster)

Diskutant/in: Joachim Wirth (Ruhr-Universität Bochum)

Die Entwicklung digitaler Technologien und der damit verbundene Zugang zu digitalisierten Welten eröffnet eine Fülle an neuen Möglichkeiten. So ist es Nutzern zum Beispiel jederzeit möglich, sich Informationen zu beschaffen, in Online-Shops einzukaufen, oder in Online Communities und sozialen Netzwerken zu kommunizieren. Der Umgang mit diesen Möglichkeiten stellt Nutzer jedoch vor neue Herausforderungen, da der kompetente Umgang mit digitalen Medien eine Voraussetzung für eine vorteilhafte Nutzung darstellt.

Bislang ist jedoch nur unzureichend geklärt, was der Begriff „digitale Kompetenz“ genau bezeichnet. Auf theoretischer Ebene wird für traditionelle sowie digitale Medien oft eine Einteilung in Medienkritik, -nutzung, -kunde, -und gestaltung vorgenommen (Baacke, 1997). Spezifische Aufgabenbereiche für digitale Medien beinhalten dabei unter anderem Informationsbeschaffung, Kommunikation und Kooperation, sowie ein angemessenes Medialitätsbewusstsein (BMBF, 2010). Auf empirischer Ebene beschäftigen sich – direkt oder indirekt - verschiedene Disziplinen mit dem Thema der digitalen Kompetenz, z.B. die Medienpädagogik, Sozial-, Medien-, kognitive und pädagogische Psychologie, sowie die Kommunikationswissenschaft.

In dem interdisziplinären Graduiertenkolleg „Vertrauen und Kommunikation in einer digitalisierten Welt“ (DFG 1712/1) wird spezifischer der Frage nachgegangen, welche Rolle Vertrauen in medialisierten Umwelten spielt. Die empirischen Beiträge aus diesem Graduiertenkolleg betrachten die Rolle von Vertrauen für digitale Kompetenz aus unterschiedlichsten Perspektiven: Auf inhaltlicher Ebene liegt ein Fokus des Symposiums auf der Beschaffung und kritischen Bewertung (wissenschaftlicher) Informationen und damit zusammenhängend der Beurteilung der Vertrauenswürdigkeit verschiedenster Quellen (Hendriks & Kienhues; vgl. BMBF, 2010). Ein zweiter Fokus dieses Symposiums liegt auf dem Thema Kommunikation und Kooperation (vgl. BMBF, 2010) in Online-Foren, Emails oder in sozialen Netzwerken und in wie weit dortiges Vertrauen abhängt von Variablen wie beispielsweise Öffentlichkeit der Kommunikation (Thon & Jucks) oder dem Bewusstsein über eigene Selbstoffenbarungen (Moll & Bromme). Der vierte Beitrag des Symposiums konzentriert sich dem gegenüber ganz auf die pädagogische Anwendungspraxis in Form eines Trainings digitaler Kompetenz zur Prävention von Cybermobbing (Pieschl). Diese Beiträge nutzen eine Vielzahl methodischer Zugänge wie Interviews (Moll & Bromme), Fragebögen (Hendriks & Kienhues), Experimente (Pieschl; Thon & Jucks) und Feldstudien (Pieschl).

In jedem einzelnen dieser Beiträge werden die Implikationen für die empirische Bildungsforschung hervorgehoben: Die Vermittlung digitaler Kompetenz ist eine Kernaufgabe aller Bildungsinstitutionen (BMBF, 2010) und dabei spielt Vertrauen eine zentrale Rolle (Blöbaum et al., 2010). Aus dieser Perspektive heraus werden die Beiträge abschließend zusammenfassend diskutiert. Dafür konnten wir einen vom Graduiertenkolleg unabhängigen Experten der empirischen Bildungsforschung gewinnen: Prof. Dr. Joachim Wirth.

Baacke, D. (1997). Medienpädagogik. Tübingen: Niemeyer.

Bawden, D. (2008). Origins and concepts of digital literacy. Digital literacies: concepts, policies and practices, 17–32.

Berlo, David, K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for Evaluating the Acceptability of Message Sources. The Public Opinion Quarterly, 33(4), 563–576.

Blöbaum, B. (2010). Vertrauen und Kommunikation in einer digitalisierten Welt: Antrag auf die Einrichtung eines Graduiertenkollegs durch die Deutsche Forschungsgemeinschaft. Westfälische Wilhelms-Universität Münster.

BMBF (Bildungsministerium für Bildung und Forschung; 2010). Kompetenzen in einer digital geprägten Kultur. Medienbildung für die Persönlichkeitsentwicklung, für die gesellschaftliche Teilhabe und für die Entwicklung von Ausbildungs- und Erwerbsfähigkeit. Bielefeld: Bertelsmann Verlag.

Bromme, R., Thomm, E., & Wolf, V. (accepted). From understanding to deference: Laypersons' and medical students' views on conflicts within medicine. International Journal of Science Education, Part B: Communication and Public Engagement.

Eimeren, B. Van, & Frees, B. (2012). 76 Prozent der Deutschen online - neue Nutzungssituationen durch mobile Endgeräte. Media Perspektiven, 7(8), 362–379.

Gummadi, K. P., Krishnamurthy, B., & Mislove, A. (2010). Addressing the Privacy Management Crisis in Online Social Networks. Proceedings of the IAB Workshop on Internet Privacy, Boston, MA, December 2010.

Jourard, S. M., & Lasakow, P. (1958). Some factors in self-disclosure. Journal of Abnormal and Social Psychology, 56(1), 91-98. doi:10.1037/h0043357

Keil, F. C. (2010). The feasibility of folk science. Cognitive Science, 34(5), 826–862.

Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. Cognitive Science, 32(2), 259–300.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. The Academy of Management Review, 20(3), 709–734.

Moll, R., Pieschl, S., & Bromme, R. (2013). (Meta-) Cognitive Factors Relevant for Self-Disclosure on Social Networking Sites. Unpublished Manuscript.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. The Journal of Consumer Affairs, 41(1), 100-126. doi:10.1111/j.1745-6606.2006.00070.x

Pieschl, S., & Porsch, T. (2012). Schluss mit Cybermobbing! Das Trainings- und Präventionsprogramm „Surf-Fair“. Weinheim: Beltz.

Pieschl, S., & Porsch, T. (2013). Das Präventionsprogramm Surf-Fair gegen Cybermobbing. Eine Einführung. Psychologie in Oesterreich, 1, 14-21.

Pieschl, S., & Porsch, T. (akzeptiert). Cybermobbing – mehr als „Ärgern im Internet“. In T. Porsch & S. Pieschl (Hrsg.), Neue Medien und deren Schatten. Mit neuen Medien kompetent umgehen. Göttingen: Hogrefe.

Pieschl, S., & Urbasik, S. (2013). Does the cyberbullying prevention program Surf-Fair work? - An evaluation study. In R. Hanewald (Ed.). From cyber bullying to cyber safety: issues and approaches in educational contexts (pp. 205-224).

Hauppauge, NY: Nova

Science Publishers.

Valkenburg, P. M., & Peter, J. (2011). Online Communication Among Adolescents: An Integrated Model of Its Attraction, Opportunities, and Risks. Journal of Adolescent Health, 48, 121-127.



## ***Dimensionen von Vertrauen in wissenschaftliche Autoren und ihre Implikationen für weitere Informationssuche***

Viele Fragen des Alltags beinhalten wissenschaftliches Wissen, zum Beispiel, ob man ein Medikament einnehmen sollte, das neu auf dem Markt ist. Wenn wir solche Fragen beantworten möchten, stellen wir oft fest, dass wir laienhaftes Wissen haben. Ein Weg, um sich zu informieren, den 79% der Deutschen nutzen, ist das Internet (Eimeren & Frees, 2012). Digitale Kompetenz ist erforderlich, um mit wissenschaftlichen Informationen aus dem Internet umzugehen. Sie beinhaltet nicht nur grundlegende Fertigkeiten im Umgang mit Text und Hypertext, sondern auch Wissen über Kontext von Informationen und soziale Praktiken, in die diese eingebettet ist (Bawden, 2008). Wir postulieren kognitive Arbeitsteilung (Keil, Stein, Webb, Billings, & Rozenblit, 2008; Keil, 2010) als wichtige Kontextbedingung der Wissensverteilung in unserer Gesellschaft. Kognitive Arbeitsteilung beschreibt die Aufteilung von Wissen in verschiedene Expertisedomänen. Aus diesem Grund bleiben wir in den meisten Domänen Laien und stoßen nicht nur im Internet auf wissenschaftliche Informationen, zu denen uns das nötige Vorwissen fehlt. Daher müssen wir, um mit solchen Informationen umgehen zu können, sowohl zuständige Experten identifizieren können (Bromme, Thomm, & Wolf, accepted), als auch überprüfen, ob diesen Experten vertraut werden kann. Insofern ist Quellenbeurteilung von Texten im Internet ein wichtiger Teil digitaler Kompetenz.

Fragestellung der vorliegenden Studie ist daher, ob sich die von Mayer, Davis und Schoorman (1995) für interpersonales Vertrauen vorgeschlagenen Dimensionen für Vertrauen in Quellen von wissenschaftlichen Informationen gelten. Gleichzeitig ist eine weitere Fragestellung, welche Strategien Personen haben, mit kognitiver Arbeitsteilung umzugehen und wie diese Strategien variieren, in Abhängigkeit von dem zuvor einer Quelle von wissenschaftlichen Informationen entgegenbrachten Vertrauen.

Um diese Fragestellungen zu untersuchen, wurde eine Online-Untersuchung durchgeführt, in dem den Probanden das Vertrauen in einen Experten - ein Forscher, der seine Ergebnisse zum (fiktiven) Medikament Serofinil in einem Science-Blog darstellt - einschätzten. Das Vertrauen wurde auf einem 18-Item Inventar eingeschätzt, das wie ein 7-stufiges semantisches Differential aus Gegensatzpaaren aufgebaut war (cf. Berlo, David, Lemert, & Mertz, 1969). Eine exploratorische Faktorenanalyse zeigt, dass sich diese Adjektive den drei Dimensionen Expertise (7 Items,  $\alpha = .91$ ), Integrität (5 Items,  $\alpha = .82$ ) und Wohlwollen (4 Items,  $\alpha = .90$ ) zuordnen lassen.

Anschließend wurden die Probanden gebeten einzuschätzen welche Strategien sie für eine weitere Informationssuche verwenden würden. Hierzu beantworteten die Probanden 18 Items, die Strategien zum Umgang mit kognitiver Arbeitsteilung beschrieben. Es zeigte sich in einer exploratorischen Faktorenanalyse, dass sich fünf Faktoren extrahieren ließen: „Informationen über den Autor suchen“ (2 Items,  $\alpha = .90$ ), „Informationen zum Thema suchen“ (6 Items,  $\alpha = .74$ ), „Auf sich selbst verlassen“ (3 Items,  $\alpha = .70$ ), „Zuständige Experten identifizieren“ (4 Items,  $\alpha = .63$ ) und „Reliabilität der Forschungsmethode prüfen“ (2 Items,  $\alpha = .60$ ).

Des Weiteren testeten wir korrelative Zusammenhänge zwischen den Skalen der Instrumente zu Vertrauen und Strategien. Je höher Probanden die Expertise des Autors einschätzten, desto mehr gaben sie an, nach mehr Informationen zum Thema suchen zu wollen,  $r = -.14$ ,  $p < .05$ . Wenn Probanden dem wissenschaftlichen Autor mehr Wohlwollen zuwies, gaben sie an, mehr nach weiteren Experten suchen zu wollen,  $r = -.17$ ,  $p < .01$ . Je weniger Probanden insgesamt vertrauten, desto eher gaben sie an, dass sie Methodik und Reliabilität der Forschungsmethode selbst beurteilen würden,  $r = -.22$ ,  $p < .01$ . Dieses letzte Ergebnis wird vor allem durch die Vertrauensskalen Expertise,  $r = -.27$ ,  $p < .01$  und Integrität,  $r = -.14$ ,  $p < .05$  bestimmt.

Wir finden also, dass Laien in der Lage sind, spontane Quellenurteile zu machen und bezüglich dreier Vertrauensdimensionen zu unterscheiden. Außerdem spielt das Vertrauen in einen Autor wissenschaftlicher Informationen eine Rolle für die weitere Informationssuche. Dies begreifen wir als wichtiges Ergebnis hinsichtlich digitaler Kompetenz bezüglich wissenschaftlicher Informationen, wie sie im Internet zu finden sind.



## ***Allein oder unter Beobachtung Anderer? Umgang mit Privatheit und Öffentlichkeit in der digitalen Welt***

Die junge Generation wird oft als „digital natives“ bezeichnet, weil sie schon sehr früh mit der digitalen Welt in Kontakt gekommen ist und mit den ständigen technologischen Veränderungen aufgewachsen ist. Doch reicht es, im digitalen Zeitalter aufgewachsen zu sein, um sich kompetent in der digitalen Medienwelt bewegen zu können? Eine Kompetenz, die es aufzubauen gilt betrifft den Umgang mit Privatheit und Öffentlichkeit im Netz. Inwiefern nehmen junge Erwachsene die „spezifischen und sich verändernden Verhältnisse zwischen dem privaten Bereich und der (...) Öffentlichkeit“ in den digitalen Medien wahr und passen ihr Kommunikationsverhalten daran an (vgl. BMBF, S.9)? Ausgangspunkt der Fragestellung bildet der Erfahrungs- und Wissensaustausch zwischen Unbekannten, wie sie häufig in Online-Foren zu finden ist. Während der aktive Austausch dort jedoch meist nur zwischen wenigen Nutzern stattfindet, rezipiert die Mehrheit der Nutzer die Konversationen passiv ohne direkt in Erscheinung zu treten (Stegbauer & Rausch, 2001): Die Grenze von privater und öffentlicher Kommunikation verwischt. Inwiefern junge Menschen dies als Spannungsfeld zwischen privatem Informationsaustausch und öffentlicher Kommunikation wahrnehmen und dies entsprechend in ihrem tatsächlichen Kommunikationsverhalten berücksichtigen, haben wir in einem Online-Experiment untersucht. Im Fokus der Analyse steht dabei der Umgang mit gefühlsbezogener Selbstoffenbarung. Selbstoffenbarung als „the process of making the self known to others“ (Jourard & Lasakow, 1958), ist ein Weg um das Ausmaß an Privatsphäre in der Kommunikation zu regulieren.

Wir haben 163 Studierende gebeten an dem Online-Projekt „Studierende helfen Studierende“ teilzunehmen und die Anfrage einer ratsuchenden Person zum Thema Prokrastination zu beantworten. Das Projekt haben wir für unseren Forschungszweck entsprechend konstruiert. In einem randomisierten Between-Subjects-Design wurde dabei variiert, ob in der Anfrage gefühlsbezogene Selbstoffenbarung (z.B. Äußerungen wie „Mir geht es damit sehr schlecht“) enthalten waren oder nicht (zweigestufteter Faktor) und ob die Kommunikationssituation als öffentlich (in einem Online-Forum) oder privat (in einer persönlichen Email) gestaltet wurde. Des Weiteren wurde die Länge der Anfrage zweigestuft variiert um Effekte der Ausführlichkeit der Anfrage auf das Antwortverhalten zu kontrollieren. Die abhängigen Variablen gliedern sich in drei Bereiche: Die Wahrnehmung der ratsuchenden Person, den Grad der wahrgenommenen Privatsphäre und das tatsächliche Kommunikationsverhalten.

Die Ergebnisse zeigen, dass die Studierenden den Grad der Privatsphäre in den beiden Medien unterschiedlich einschätzen. Dabei wird die Kommunikation im E-Mailkontext als privater empfunden als wenn diese in einem Forum erfolgt. Auch die ratsuchende Person wird entsprechend der Manipulation als selbstoffenbarer wahrgenommen, wenn die Anfrage gefühlsbezogene Selbstoffenbarung enthielt. In dem resultierenden Kommunikationsverhalten der Studierenden zeigten sich jedoch keine systematischen Unterschiede. In der Literatur wird diese Diskrepanz oft als „Privatsphäre-Paradox“ (Norberg, Horne, & Horne, 2007) beschrieben, es gibt aber bislang nur wenige experimentelle Studien, die das tatsächliche Kommunikationsverhalten untersucht haben. Die Studie zeigt, dass die Kompetenz den digitalen Raum einzuschätzen nicht zwingend zu einem sozial verantwortlichen Agieren führt. Die Ergebnisse liefern somit erste Ansatzpunkte für bildungspolitisches Handeln. Die Gründe für diese Diskrepanz sollten Gegenstand zukünftiger Forschung sein.

## ***Sind sich Schüler darüber bewusst, was sie online preisgeben? Vertrauen in das eigene Wissen als Teilaspekt digitaler Kompetenz***

### Theoretischer Hintergrund

Die Nutzung sozialer Netzwerkseiten erfreut sich vor allem unter Jugendlichen und jungen Erwachsenen großer Beliebtheit (Valkenburg & Peter, 2011). Dabei ist von zentraler Bedeutung, dass Nutzer mit dem Spannungsverhältnis zwischen den positiven Aspekten der online-Selbstoffenbarung und den damit verbundenen Privatheitsrisiken kompetent umgehen (vgl. Moll, Pieschl, & Bromme, 2013). Eine notwendige Voraussetzung für den kompetenten Umgang mit den eigenen persönlichen Daten ist Wissen darüber, was man über sich selbst offenbart. Selbstoffenbarung bezieht sich dabei auf zwei Aspekte: a) welche Information wurde online gestellt, und b) für welches Publikum ist die Information potentiell zugänglich. Da dieses Faktenwissen jedoch aufgrund komplizierter Privatheitsoptionen innerhalb sozialer Netzwerke oft nur schwer zu erwerben ist (Gummadi, Krishnamurthy, & Mislove, 2010, p.1), scheint es umso wichtiger, zusätzlich die Verlässlichkeit des eigenen Wissens akkurat einschätzen zu können (Metakognitive Akkuratheit). Bisher fehlen jedoch Studien, die das Ausmaß an Faktenwissen und metakognitiver Akkuratheit von Nutzern sozialer Netzwerkseiten untersuchen. Die vorliegende Studie zielt darauf, diese Forschungslücke zu füllen.

### Methode

In standardisierten Interviews wurden Schüler (N = 45; 25 männlich, 20 weiblich; M = 16.53 Jahre; SD = 1.35) zu ihrer Facebook-Nutzung befragt. Sie machten für 16 inhaltliche Profilkategorien Angaben dazu, ob sie Information über sich selbst in den Kategorien preisgegeben haben, und für wen der entsprechende Inhalt sichtbar ist. Für jede Angabe stufen die Schüler zusätzlich ihre subjektive Sicherheit ein, dass die gemachte Angabe korrekt ist. Im Anschluss loggten die Schüler sich in ihren Facebook-Account ein, sodass die Interviewer dokumentieren konnten, in welchen Kategorien tatsächlich Information preisgegeben wurde und für wen sie tatsächlich sichtbar ist.

### Ergebnisse

Die Schüler machen in der überwiegenden Anzahl der Fälle korrekte Angaben bezüglich ihrer Profilinehalte (Faktenwissen). Die Angaben zu der entsprechenden Publikumseinstellung bereiteten hingegen größere Schwierigkeiten.

Bezüglich der metakognitiven Akkuratheit, d.h. dem Zusammenhang zwischen der subjektiven Sicherheit der Schüler und der tatsächlichen Korrektheit ihrer Angabe, zeigt sich strukturell gesehen ein ähnliches Bild: Während Schüler geringe, jedoch signifikante Akkuratheit bei ihren Angaben zu offenbarten Profilinehalten zeigen (Mittelwert Korrelation:  $M = .27$ ;  $SD = .38$ ,  $T(35) = 4.29$ ;  $p < .001$ ), fehlt dieser Zusammenhang für das potentielle Publikum vollständig (Mittelwert Korrelation:  $M = -.08$ ;  $SD = .56$ ,  $T(30) = -.76$ ;  $p = .451$ ). Insgesamt fällt der Zusammenhang zwischen der subjektiven Sicherheit und der Korrektheit der Angaben sehr gering aus.

### Implikationen

Die vorliegende Studie beleuchtet zwei Punkte, die bisher in Diskussionen zu Kernaspekten digitaler Kompetenz vernachlässigt wurden: Nutzer sozialer Netzwerkseiten wissen zwar noch vergleichsweise gut, was sie dort über sich bekanntgegeben haben, aber sie wissen nicht mehr wem. Da dies auch der Tatsache geschuldet ist, dass Publikumseinstellungen und Privatheitsmanagement auf Facebook sehr kompliziert sind, kann dieses Ergebnis nicht grundsätzlich als Defizit bewertet werden. Problematischer ist jedoch, dass außerdem das metakognitive Wissen über diese fehlende Kenntnis der Zuordnung von Inhalt und Publikum nur sehr gering ausgeprägt ist. Bildungsansätze zur Förderung digitaler Kompetenz bei Kindern und Jugendlichen sollten also darauf abzielen, Nutzer in ihrer Fähigkeit zu stärken, sich selbst, ihr Wissen und ihre Kompetenz realistisch einschätzen zu können. Aus der Methodik unserer Studie entwickeln wir abschließend Vorschläge für instruktionale Maßnahmen, die genau auf dieses metakognitive Wissen bei Schülerinnen und Schülern abzielen.



## ***Förderung digitaler kommunikativer Kompetenz als Schutz gegen Cybermobbing?***

### Theoretischer Hintergrund und Fragestellung

Ein substantieller Anteil deutscher Jugendlicher leidet unter Cybermobbing und dies hat u.a. negative Wirkung auf ihr emotionales Wohlbefinden und auf ihre Bildungsbeteiligung (Pieschl & Porsch, akzeptiert). Bisherige überwiegend korrelative Studien zeigen eine Vielzahl von möglichen Schutz- und Risikofaktoren auf, nicht nur genereller sondern auch spezifisch digitaler Natur (Pieschl & Porsch, akzeptiert): Der größte Risikofaktor ist Cybermobbing selbst, da es eine große Schnittmenge zwischen Tätern und Opfern gibt. Weitere Faktoren präzisieren daher sowohl Täter- als auch Opferschaft, beispielsweise ausgeprägte Computerkenntnisse, viel im Internet verbrachte Zeit und Risikoverhalten im Internet. Andere Faktoren scheinen dagegen spezifisch für Opfer von Cybermobbing, beispielsweise ein aktives Profil in einem sozialen Netzwerk und hohes Vertrauen gegenüber Online-Kommunikationspartnern. Diese Faktoren liefern erste Hinweise darauf, welche digitalen Kernkompetenzen für erfolgreiche digitale Kommunikation ohne Cybermobbing notwendig sind. Auch wenn dies sinnvoll scheinen mag, gehört Medienabstinenz nicht dazu! Vielmehr sollte ein angemessener Umgang mit Medien erlernt werden. Beispielsweise könnte erhöhte Medienkritik (Baacke, 1997) dazu führen, dass Jugendliche kritisch hinterfragen, wie viel Selbstoffenbarung in ihren sozialen Profilen notwendig ist und wie vertrauenswürdig (anonyme) Internetbekanntschaften tatsächlich sind. Diese digitalen Kompetenzen sollen mit Surf-Fair (Pieschl & Porsch, 2012, 2013), einem Präventionsprogramm gegen Cybermobbing, vermittelt werden. In den vorgestellten Studien soll folgender zentraler Fragestellung nachgegangen werden: Ist die Vermittlung digitaler kommunikativer Kompetenzen erfolgreich bezüglich der subjektive Bewertung, des Wissens und des Verhaltens der Teilnehmer/innen?

### Methode

Zur Beantwortung dieser Fragestellung wurden zwei explorative Studien mit jeweils unterschiedlichem Fokus durchgeführt. An der ersten experimentellen Evaluationsstudie (N = 87; Pieschl & Urbasik, 2013) nahmen drei sechste Gymnasialklassen teil (je n = 29), eine Klasse als Kontrollklasse (KK), eine Klasse mit einer Doppelstunde Surf-Fair (EX90) und eine Klasse mit zwei Doppelstunden Surf-Fair (EX180). Vorher (prä), nachher (post) und zwei Monate später (follow-up) wurde das Verhalten (Cybermobbing in den letzten zwei Monaten), das Wissen (über effektive Copingstrategien) und die subjektive Bewertung der Schüler/innen als abhängige Variablen erhoben. In der folgenden Feldstudie (N = 289) wurde Surf-Fair in drei aufeinander folgenden Jahren in den fünften Klassen eines Gymnasiums durchgeführt (je zwei Doppelstunden), im ersten Jahr von geschulten Trainer/innen der Universität Münster und in den folgenden Jahren von Lehrer/innen. Es wurden die gleichen abhängigen Variablen prä und follow-up erhoben wie in der ersten Studie.

### Ergebnisse

Die experimentelle Evaluationsstudie zeigt, dass die Teilnahme an einer ausführlichen Version von Surf-Fair (EX180) im Vergleich zur Kontrollgruppe (KG) zu signifikant weniger Opfer- und Täterschaft von Cybermobbing führt, in der EX90 Klasse änderten sich diese Prävalenzen nicht. Darüber hinaus lernten die Schüler/innen abhängig von der Dosis des Treatments (längeres Treatment = stärkere Effekte) signifikant etwas über angemessene technische Copingstrategien und bewerteten ihren Lernerfolg signifikant positiv. Die Feldstudie II konnte die positiven Ergebnisse bezüglich Wissen und subjektiven Bewertungen replizieren und zeigt darüber hinaus, dass Lehrer/innen Surf-Fair genauso erfolgreich durchführen können wie geschulte Trainer/innen. Zusammengefasst kann Surf-Fair positive Effekte in der subjektiven Bewertung, dem Wissen und Verhalten der Teilnehmer/innen haben und es kann von Lehrer/innen erfolgreich an Schulen implementiert werden.

Diese und ähnliche Studien zeigen, dass digitale kommunikative Kompetenzen erfolgreich trainiert werden können und dass Schulen ein möglicher Ort für solche Trainings sind. Bisher bleibt allerdings unklar, welche Mechanismen genau zu diesen Effekten führen. Eine Möglichkeit ist, dass neben einem Wissenszuwachs vor allem eine kritische Auseinandersetzung mit der eigenen Mediennutzung (Baacke, 1997) zu entsprechenden Verhaltensänderungen geführt hat.

## ***Fertigkeiten für eine reflektierte Nutzung des Social Web***

Das Internet und insbesondere das Social Web, d.h. z.B. Blogs, Foren, Wikis, und soziale Netzwerke, sind zu einem wichtigen Bestandteil der Lebenswelt Jugendlicher geworden und bieten auch für Lehr-Lernkontexte neues Potential. Gleichzeitig sind für eine reflektierte Nutzung des Social Web für Wissenserwerb und Wissensaustausch verschiedene kognitive Fertigkeiten notwendig, die wir in einem theoretischen Rahmenmodell identifizierten. Grundlage des Rahmenmodells waren dabei allgemeine ICT (information and communication technology)-Literacy-Modelle (z.B. ICILS, Eickelmann & Bos, 2011; iSkills, International ICT Literacy Panel, 2002) sowie pädagogisch-psychologische Modelle zum individuellen Lernen und Problemlösen im Internet (IPS-I Modell, Brand-Gruwel, Wopereis, & Walraven, 2009; MD-TRACE Modell, Britt & Rouet, 2012) sowie zu kollaborativen Problemlösefertigkeiten (Hesse, Buder, & Sassenberg, 2011). Das Rahmenmodell umfasst die vier Kompetenz-Facetten Auswählen und Rezipieren von Informationen, Verwalten von Informationen, Kommentieren und Kommunizieren von Informationen sowie Erstellen von Informationen. Zu jeder Kompetenzfacette wurden basierend auf psychologischen Theorien sowie empirischen Befunden konkrete Fertigkeiten operationalisiert. Anschließend wurden zur Überprüfung dieser Fertigkeiten computerbasierte Aufgaben entwickelt (vgl. PISA Online Reading, OECD, 2011), in denen es beispielsweise darum geht Informationen zu bewerten (Facette Auswählen und Rezipieren) oder für sich und andere zu ordnen (Facette Verwalten), eigene Beiträge (gemeinsam) zu gestalten (Facette Erstellen) oder mit anderen zu teilen (Facette Kommentieren und Kommunizieren).

Derzeit werden die Aufgaben mit ca. 100 Schülerinnen und Schülern aus fünf Berufsschulklassen getestet. Die Aufgaben werden den Studienteilnehmern am Laptop präsentiert und sollen innerhalb von 2 Schulstunden individuell bearbeitet werden. Dabei werden konkrete Szenarien mit Webseiten (z.B. Blogs, Foren, soziale Netzwerk-Seiten) präsentiert, in denen die Schülerinnen und Schüler beispielsweise bestimmte Webseiten explorieren und dabei die Angemessenheit der Beiträge bewerten sollen (z.B. ob sie eine bestimmte Webseite für ein Referat verwenden würden oder ob sie bestimmte Beiträge posten würden). Darüber hinaus werden anhand von offenen Fragen Strategien von Schülerinnen und Schülern abgefragt (z.B. wie sie vermeiden, im Internet von zu vielen Informationen überflutet zu werden).

Im Rahmen des geplanten Posterbeitrags soll das Rahmenmodell, sowie die Aufgaben und die Ergebnisse der empirischen Studie vorgestellt werden.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, 53, 1207–1217.

Britt, M., & Rouet, J. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276-314). New York, NY US: Cambridge University Press.

Eickelmann, B., & Bos, W. (2011). Messung computer- und informationsbezogener Kompetenzen von Schülerinnen und Schülern als Schlüsselkompetenz im 21. Jahrhundert. *Zeitschrift Medienimpulse*, 3, 1–11.

Hesse, F.W., Buder, J., & Sassenberg, K. (2011, September). A framework for teachable collaborative problem solving skills. *European Conference on Educational Research (ECER)*. Berlin.

International ICT Literacy Panel (2002). *Digital transformation: A framework for ICT literacy (A report of the International ICT Literacy Panel)*. Princeton, NJ: Educational Testing Service.

[http://www.ets.org/Media/Tests/Information\\_and\\_Communication\\_Technology\\_Literacy/ictreport.pdf](http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf)

OECD (2011). *PISA 2009 Results: Students on Line: Digital Technologies and Performance (Volume VI)*. <http://dx.doi.org/10.1787/9789264112995-en>.

Luisa Friedrich (Universität Duisburg-Essen), Annett Schmeck (Universität Duisburg-Essen), Maria Opfermann (Universität Duisburg-Essen), Detlev Leutner (Universität Duisburg-Essen)

## **Schwierigkeiten und Herausforderungen beim Lernen mit selbstgenerierten Bildern – auf Papier per Hand versus am Computer per Drag-and-Drop**

Schülerinnen und Schüler (SuS) haben häufig Probleme beim Lernen aus naturwissenschaftlichen Texten, da diese oft sehr komplex und somit kognitiv anspruchsvoll sind. Eine Lernstrategie, die die SuS beim Lernen aus naturwissenschaftlichen Texten unterstützt, ist es, die SuS zu instruieren, während des Lesens Bilder zu zeichnen, die die wichtigsten Elemente des Textinhaltes widerspiegeln (sog. Lernen mit selbstgenerierten Bildern; vgl. Leutner & Schmeck, in press; van Meter & Garner, 2005). Das Selbstgenerieren von Bildern zu einem Text hat sich sowohl auf Papier (Schmeck, 2010; Schwamborn et al., 2010) als auch am Computer (Friedrich et al., 2013) als lernförderlich erwiesen. Einerseits kann das Übersetzen von Text in Bilder bei der Bildgenerierung kognitive Prozesse der Informationsverarbeitung fördern, indem die wichtigsten Informationen selegiert, in kohärente Strukturen organisiert und schließlich mit dem vorhandenen Vorwissen integriert werden (Mayer, 2009). Andererseits garantiert das Generieren von Bildern nicht automatisch tiefes Verstehen des Lernmaterials. Wie das Lernmaterial präsentiert wird sowie die Art des Generierens selbst kann kognitive Belastungen erzeugen (Sweller, 2005). So berichten SuS, die Bilder zu einem Text am Computer generieren, von einer höheren Schwierigkeit im Sinne einer höheren kognitiven Belastung als SuS, die keine Bilder generieren (Friedrich et al., 2013). Außerdem ist der Effekt des Selbstgenerierens von Bildern stärker in papier-basierten (Schmeck, 2010) als in computer-basierte Lernumgebungen (Friedrich et al., 2013). Gegenstand der vorliegenden Studie ist es daher, mögliche Erklärungen für Unterschiede in der Effektivität des Lernens mit selbstgenerierten Bildern aufzuzeigen, und zwar abhängig davon, ob SuS Bildern zu einem Text auf Papier oder am Computer generieren. Im Speziellen bezieht sich die Fragestellung darauf, Ursachen der von den SuS empfundenen Schwierigkeit beim Generieren auf Papier gegenüber dem Generieren am Computer zu untersuchen und Empfehlungen abzuleiten, wie das Lernen mit selbstgenerierten Bildern erfolgreich implementiert werden kann. 54 SuS gymnasialer 8. Klassen wurden zufällig auf zwei Experimentalgruppen verteilt. Beide Gruppen wurden instruiert, einen in 6 Abschnitte unterteilten Chemietext zu lesen. Dabei wurden die Textabschnitte, beiden Gruppen, abwechselnd auf Papier (mit der Instruktion, mit einem Stift ein Bild auf Papier zu zeichnen) und am Computer (mit der Instruktion, eine Zeichnung mit Hilfe von drag-and-drop zu erstellen) präsentiert, wobei die eine Experimentalgruppe auf Papier, die andere am Computer begann. Das computerbasierte Lernmaterial präsentierte jeweils neben einem Textabschnitt einen drawing prompt, bestehend (1) aus einer Werkzeuggeste, die alle relevanten Zeichenelemente enthielt, welche die SuS per drag-and-drop auf (2) einem zum Teil vorgezeichneten Zeichenhintergrund platzieren und kombinieren konnten. Das papierbasierte Material enthielt zu jedem Textabschnitt (1) eine gedruckte Werkzeuggeste inklusive der Zeichenelemente, welche die SuS per Hand auf den (2) Zeichenhintergrund zeichnerisch platzieren und kombinieren konnten. Nach der Lernaufgabe beantworteten die SuS einen Fragbogen bestehend aus 13-Items, die sich auf Freude und Motivation beim Umgang mit dem Lernmaterial, Schwierigkeiten beim Genieren der Bilder sowie auf den Umgang mit dem Lerninhalt selbst beziehen. Die Ergebnisse zeigen, dass insgesamt signifikant mehr SuS angaben, dass es für sie einfacher war, am Computer zu lernen als auf Papier. Allerdings entschieden sich bezüglich der Ausrichtung der Zeichenelemente auf dem Zeichenhintergrund mehr SuS für das Papiermaterial. Des Weiteren gaben signifikant mehr SuS an, dass sie den Inhalt des Textes leichter verstehen konnten, wenn sie auf Papier statt am Computer arbeiteten. Zusätzlich gaben signifikant mehr SuS an, in der Papierlernumgebung mehr über die zu zeichnenden Bilder nachzudenken als in der Computerlernumgebung. Die Ergebnisse lassen sich im Hinblick darauf diskutieren, inwieweit die Vorteile beider Lernumgebungen kombiniert werden können, um eine mögliche lernabträgliche kognitive Belastung der Lernenden zu minimieren und mehr kognitive Ressourcen für Informationsverarbeitung vorhalten zu können, die tiefes Lernen fördert (Mayer, 2009; Sweller, 1999, 2005).

Friedrich, L. A., Schmeck, A., Opfermann, M., & Leutner, D. (2013, April). Computer-based visualizations as comprehension aids for science text learning - extending the generative drawing principle. Presented at AERA, San Francisco, USA.

Leutner, D., & Schmeck, A. (in press). The generative drawing principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed). New York: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed). New York: Cambridge University Press.

Schmeck, A. (2010). *Visualisieren naturwissenschaftlicher Sachverhalte: Der Einsatz von vorgegebenen und selbst generierten Visualisierungen als Textverstehenshilfen beim Lernen aus naturwissenschaftlichen Sachtexten* (Dissertation). Essen: Universität Duisburg-Essen.

Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology*, 102, 872–879.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of multimedia learning* (S. 19-30). New York: Cambridge University Press.

van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawings: Literature review and synthesis. *Educational Psychology Review*, 12, 261–312.

## Erfassung des Fachwissens von Lehramtsstudierenden der Physik

Der Forschungsverbund „Professionswissen in der Lehramtsausbildung Physik“ (ProfiLeP) ist ein gemeinsames Projekt der Universitäten Paderborn, Bremen, Potsdam und Duisburg-Essen. Beim Professionswissen werden üblicherweise die folgenden Dimensionen unterschieden (Baumert & Kunter, 2006): Fachwissen, fachdidaktisches Wissen, pädagogisches Wissen. Das Projekt ProfiLeP konzentriert sich auf die Dimensionen Fachwissen und fachdidaktisches Wissen von Studierenden der Lehramter im Fach Physik. In unserem Teilprojekt „Fachwissen im Studium zum Lehramt der Physik“ (FaWis) wird das Fachwissen von Lehramtsstudierenden und Übungsleiterinnen und -leitern untersucht. Ziele sind u.a.:

- Entwicklung eines Modells zur Erfassung des Fachwissens in Physik unter Berücksichtigung des Teilbereichs „vertieften Schulwissens“.
- Entwicklung eines Fachwissenstests für Lehramtsstudierende der Physik.
- Diskriminante Validierung des Modells zum Fachwissen durch Vergleich der Testergebnisse von Lehramtsstudierenden mit den Testergebnissen von Fachstudierenden der Physik.
- Abstimmung des Modells zum Fachwissen mit den beiden Modellen zum fachdidaktischen Wissen zur Aufklärung von vermuteten Zusammenhängen zwischen Fachwissen und fachdidaktischem Wissen.

Für die Itemkonstruktion zum Fachwissenstest wird ein Kompetenzmodell für die Sekundarstufe II (Schoppmeier, Borowski & Fischer, 2012) adaptiert, das auf dem ESNaS-Modell für die Sekundarstufe I (Kauertz, Fischer, Mayer, Sumfleth, & Walpuski, 2010) aufbaut. Dieses Modell besitzt drei Dimensionen: Inhalt, Komplexitätsstufe und Wissensart. Beim Inhalt wurde für größtmögliche Inhaltsvalidität im Gesamtprojekt eine Fokussierung auf die Mechanik vorgenommen. Zur Variation der Schwierigkeit existieren 3 Komplexitätsstufen: 1 Fakt/2 Fakten, 1 Zusammenhang/2 Zusammenhänge und Übergeordnetes Konzept. Die Wissensarten sind Schulwissen, vertieftes Schulwissen und universitäres Wissen. Beim Schulwissen werden Inhalte berücksichtigt, die in den Richtlinien und Lehrplänen für die Sekundarstufen I und II vorgesehen sind. Der Oberstufenstoff ist der obligatorische Inhalt für die Grundkurse. Der Bereich des vertieftes Schulwissens wird in COACTIV (Kunter, Baumert, Blum, Klusmann, Krauss, & Neubrand, 2011) als tieferes Verständnis der curricularen Fachinhalte der Mathematik aufgefasst und in MT21 (Blömeke, Kaiser, & Lehmann, 2008) als Schulmathematik vom höheren Standpunkt. Beide Studien beziehen sich dabei auf Werke von Kirsch (1987) und Klein (1933), ohne jedoch eine weitergehende Akzentuierung vorzunehmen. Im Teilprojekt FaWis wurden Items unter Berücksichtigung der folgenden 6 Fähigkeiten entwickelt:

- Unterschiedliche Wege zur Lösung einer Aufgabe identifizieren und anwenden können.
- Unterschiedliche theoretische Einordnungen zur Lösung einer Aufgabe nutzen können.
- Ausgehend von einer Schulaufgabe die physikalischen Randbedingungen erkennen können, unter denen sie gelöst werden kann.
- Ausgehend von einer universitären Aufgabe physikalisch reduzierte Aufgaben für die Schule auf einem jeweils angemessenen Niveau entwickeln können.
- Ausgehend von einer universitären Aufgabe die einschränkenden physikalischen Bedingungen einer Schulaufgabe erkennen können.
- Zusammenhänge, Gemeinsamkeiten und Unterschiede von physikalischen Phänomenen erkennen können.

Beim universitären Wissen werden Inhalte berücksichtigt die erst innerhalb der physikalischen Ausbildung an Universitäten angeboten werden (Riese, 2009).

Im Sommersemester 2013 fand die Pilotstudie mit 436 Studierenden der Physik an sechs Universitäten statt. Es wurden sechs Testhefte im Multi-Matrix-Format eingesetzt, welche insgesamt 145 Items im Multiple-Choice, Single Select Format erfassten. Die Rasch-Analyse der Daten erfolgte mit der Open-Source-Software R. Aus den 122 Items mit guten Fitwerten ( $0.8 < MNSQ < 1.2$ ) wurden 40 Items für die Hauptstudie ausgewählt. Der Test konnte in der Pilotierung die Fähigkeiten der Studierenden gut auflösen. In der Hauptstudie wird von Wintersemester 2013/14 bis Wintersemester 2014/15 zu vier ausgewählten Messzeitpunkten das Fachwissen der Studierenden getestet. Jedes Testheft enthält einen Test zu kognitiven Fähigkeiten, zur Erfassung der Personenmerkmale und den Fachwissenstest. Physikstudierende des 1. und 3. Semesters werden getestet. Lehramtsstudierende bilden die Zielgruppe und Fachstudierende die Vergleichsgruppe. Im Vortrag werden Aufgaben aus der Pilotierung präsentiert, um die Bereiche der Wissensarten im Kompetenzmodell zu charakterisieren. Zudem werden Ergebnisse der ersten Erhebung der Hauptstudie präsentiert.

- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft (9). 469-520.
- Blömeke, S., Kaiser, G., Lehmann, R. (Hrsg.) (2008). Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Münster: Waxmann.
- Kauertz, A., Fischer, H. E., Mayer, J., Sumfleth, E., Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. Zeitschrift für Didaktik der Naturwissenschaften 16. 135-153.
- Kirsch, A. (1987). Mathematik wirklich verstehen: Eine Einführung in ihre Grundbegriffe und Denkweisen. Köln: Aulis.
- Klein, F. (1933). Elementarmathematik vom höheren Standpunkte aus. Berlin: Springer.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und Lehrer im Rahmen der COACTIV-Studie. Journal für Mathematik-Didaktik 29 3/4, 223-258.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., Neubrand, M. (Hrsg.) (2011). Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann.
- Prenzel, M., (Hrsg.) (2006). Untersuchungen zur Bildungsqualität von Schule. Münster: Waxmann.
- Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. Berlin: Logos.
- Riese, J. & Reinhold, J. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. Zeitschrift für Didaktik der Naturwissenschaften 16. 167-187.
- Schoppmeier, F., Borowski, A., Fischer, H. E. (2012). Mathematische Bereiche in Leistungskursklausuren. Phydid A – Physik und Didaktik in Schule und Hochschule Nr. 11. Band 1.



## **Lernen aus Fehlern im Schulpraktikum**

Lernen im Praktikum ist ein Themenbereich, der in den letzten Jahren zunehmend in den Mittelpunkt der Aufmerksamkeit gerückt ist. Die bisherigen empirischen Ergebnisse raten zur Bescheidenheit, da sie verdeutlichen, dass die Wirksamkeit von Praktika für die professionelle Entwicklung angehender Lehrpersonen tendenziell überschätzt wird (im Überblick siehe z. B. Hascher, 2012). Es ist daher von besonderer Bedeutung, Rahmensituationen und Faktoren zu identifizieren, in denen nachhaltiges Lernen stattfinden kann. Im Projekt soll deshalb das Konzept des Lernens aus Fehlern auf den Kontext des Lernens von Lehramtsstudierenden im Praktikum übertragen werden. Im Sinne der Theorie „negativen Wissens“ (Oser & Spsychiger, 2005) wird davon ausgegangen, dass in Fehlersituationen erworbene Erkenntnisse über das Nicht-Funktionieren von Prozessen nachhaltig im Gedächtnis verankert sind. Ausgangsthese ist, dass das Wissen über fehlerhafte Unterrichtshandlungen dabei helfen kann, adaptive Alternativen für ähnliche unterrichtliche Situationen in der Zukunft zu entwickeln und dadurch die Qualität des Unterrichts zu erhöhen. Vor dem Hintergrund dieser theoretischen Überlegungen stellen sich die Fragen, welche Fehler Praktikant/inn/en erkennen und inwiefern sie diese dazu nutzen, ihre professionellen Handlungskompetenzen zu erweitern – ob also in schulischen Praktika aus Fehlern gelernt wird. Darüber hinaus ist von Interesse, in welchen Dimensionen von Unterrichtsqualität (Helmke, 2010) das Lernen aus Fehlern gegebenenfalls besonders fruchtbar ist: Können Fehlerschwerpunkte im Rahmen schulischer Praktika ermittelt und deren Potential für den Lernprozess der Studierenden identifiziert werden? In der Untersuchung werden subjektiv relevante Misserfolgserlebnisse und damit verbundene Lernprozesse Studierender während schulischer Praktika genauer betrachtet. Dazu wurden bestehende Lerntagebücher von Lehramtsstudierenden (vgl. Hascher & Wepf, 2007) wie folgt reanalysiert: (1) In einem ersten Schritt wurden alle von den Praktikant/inn/en berichteten Lernsituationen daraufhin untersucht, ob es sich um eine Fehlersituation handelt. Eine solche liegt dann vor, wenn der Praktikantin / dem Praktikanten im Zuge der Unterrichtsversuche durch eine der beteiligten Personen (z.B. Praktikant/in selbst, Schüler/innen, Praktikumsleiter/in) deutlich wird, dass ein Fehler aufgetreten ist. Anhand dieser Definition konnten insgesamt 343 Fehlersituationen (55% Anteil der berichteten Lernsituationen) identifiziert werden, was die Bedeutung des Lernens aus Fehlern im Kontext schulischer Praktika verdeutlicht. (2) In einem zweiten Schritt wurden die gefundenen Fehlersituationen nach inhaltlichen Gesichtspunkten von zwei unabhängigen Raterinnen kategorisiert. Dem Kategorienschema wurden die Merkmale guten Unterrichts nach Helmke (2010) zugrunde gelegt, was über alle Analyseschritte hinweg eine Interrater-Reliabilität von über 90% ergab. Die Kategorisierung erfolgte mit dem Ziel, eine Taxonomie für Fehler zu erstellen, die im Zuge von Unterricht im Praktikum auftreten und für Lernprozesse nutzbar gemacht werden können. Mit 26.5% aller Nennungen erweist sich die mangelnde „Schüler/innenorientierung“ als größte Fehlerkategorie. Sie umfasst vor allem eine fehlerhafte Einschätzung des kognitiven Leistungsstands der Lernenden sowie ein daraus resultierendes über- oder unterforderndes Anforderungsniveau des Unterrichts. 21.9% aller genannten Fehlersituationen können der Kategorie „Klarheit und Strukturiertheit der Aufgabenstellung“ zugeordnet werden, in der beispielsweise unklare Instruktionen der angehenden Lehrer/innen zum Ausdruck kommen. Als weitere fehleranfällige Unterrichtsaspekte sind Probleme mit der Klassenführung zu nennen, vor allem ein suboptimaler „Umgang mit Störungen“ (9.0%). Die Merkmale guten Unterrichts besitzen also auch hinsichtlich wahrgenommener Herausforderungen in Praktikum praktische Relevanz mit Bezug auf die Entwicklung professioneller Handlungskompetenz von angehenden Lehrpersonen. Die Befunde der Untersuchung geben Aufschluss darüber, in welchen Bereichen eigene Erfahrungen und damit zusammenhängende Misserfolgserlebnisse im Sinne der Theorie des negativen Wissens zu fruchtbaren Lernprozessen führen können. Diese gilt es im Zuge der Lehramtsausbildung näher zu beleuchten und eine diesbezügliche Kompetenzentwicklung entsprechend zu unterstützen und zu fördern. So ist es wichtig, Studierende frühzeitig an Möglichkeiten der Diagnose des kognitiven Wissensstandes der Schüler/innen heranzuführen, um Unter- und Überforderung der Lernenden zu vermeiden und deren Lernprozesse zu optimieren.

- Hascher, T. (2012). Lernfeld Praktikum – Evidenzbasierte Entwicklungen in der Lehrer/innenbildung. Zeitschrift für Bildungsforschung, 2(1), 109-129.  
Hascher, T. & Wepf, L. (2007). Lerntagebücher im Praktikum von Lehramtsstudierenden. Empirische Pädagogik, 21, 101-118.  
Helmke, A. (2010). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Stuttgart: Klett.  
Oser, F. & Spsychiger, M. (2005). Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur. Weinheim: Beltz.



## **Erwerb und Entwicklung allgemeindidaktischer Planungskompetenz in unterschiedlichen Praktikumsformen**

Das Poster präsentiert eine Analyse der Lernentwicklung von Lehramtsstudierenden an der Universität Hildesheim in unterschiedlichen Praktikumsformen.

In seiner Expertise für die KMK hat Terhart (2002) Zielvorstellungen für Schulpraktische Studien formuliert, die mit den Aufstellungen in einschlägigen Publikationen zur Praktikumsforschung anschlussfähig sind. In der Literatur werden in Regel drei Bereiche besonders hervorgehoben: (1) Eignungsüberprüfung und Berufsfelderkundung, (2) Erwerb und Erweiterung (professionsspezifischer) Kompetenzen sowie (3) die Möglichkeit der Verknüpfung theoretischer und praktischer Elemente der Lehrerbildung (u.a. Reinthoffer & Dörr, 2008). Fraglich ist dabei, ob diese Ziele faktisch erreicht werden und ob sämtliche Ziele für alle Praktikumsmodelle und -formen von gleicher Bedeutung sind.

An der Universität Hildesheim werden seit 2009 Lehramtsstudierende in den schulpraktischen Ausbildungsphasen durch das Projekt ESIS (Entwicklung Studierender in Schulpraktika) untersucht. Schwerpunkt der Forschung bildete bislang der Erwerb bzw. die Entwicklung allgemeindidaktischer Planungskompetenz in den Praxisphasen. Hierzu liegen inzwischen mehrere Publikationen aus dem Projekt vor (u.a. Bach, 2013; Bach, Besa & Arnold, 2013, 2014; Bach, Besa, Brodhäcker & Arnold, 2012). Die Messung des Konstrukts erfolgt mittels einer im Projekt entwickelten Selbsteinschätzungsskala, die, im Rückgriff auf Planungsmodelle der Allgemeinen Didaktik, Aspekte einer allgemeindidaktisch orientierten Unterrichtsplanung abbildet. Die Skala konnte für alle Messzeitpunkte des Projekts validiert werden. In einem aufwändigen Studiendesign wird die Entwicklung der selbsteingeschätzten Kompetenz über zwei Schulpraktika mit insgesamt sechs Messzeitpunkten erfasst. Im Rahmen des Projekts konnte ein Anstieg der selbsteingeschätzten Kompetenz bereits für ein Blockpraktikum, das Allgemeine Schulpraktikum (ASP), nachgewiesen werden (Bach, 2013). In einem weiteren Schritt soll nun überprüft werden, ob diese Entwicklung in verschiedenen Praktikumsformen unterschiedlich verläuft. Dazu sollen neben der Betrachtung des ASP, die vorgelagerten Schulpraktischen Studien (SPS), bei denen es sich um Tagespraktika handelt, in die Analysen mit einbezogen werden. Die Frage nach differentiellen Lernentwicklungen in den unterschiedlichen Praktikumsformen wird unter anderem durch die organisatorischen, aber auch inhaltlichen Rahmenbedingungen der Praktika aufgeworfen: Ist in semesterbegleitenden Tagespraktika, wie beispielsweise in den Schulpraktischen Studien des hier untersuchten Hildesheimer Modells der Lehrerbildung, eine parallele Lehrveranstaltung an der Universität vorgesehen, können die Praxiserfahrungen direkt theoriebezogen reflektiert und Gelingensbedingungen analysiert werden – und umgekehrt, das zuvor erworbene Theoriewissen in der Praxis angewendet werden. Diese Möglichkeit ist in Blockpraktika oftmals nur begrenzt gegeben, da, wenn vorhanden, vor- und nachbereitende hochschulische Lehrveranstaltungen oftmals zu erheblichen Teilen für organisatorische Zwecke genutzt werden. Andererseits bietet ein Blockpraktikum mehr Möglichkeiten der individuellen Planung von Unterrichtssequenzen und -stunden.

Die Analyse der Entwicklung Allgemeindidaktischer Planungskompetenz an einer Kohorte Hildesheimer Lehramtsstudierender soll für die unterschiedlichen Praktikumsformen mittels einer Latent-Change-Analyse erfolgen. In die Modellierung werden jeweils Messzeitpunkte vor und nach den beiden Schulpraktika einbezogen.

Bach, A. (2013). Kompetenzentwicklung im Schulpraktikum. Ausmaß und zeitliche Stabilität von Lerneffekten hochschulischer Praxisphasen. Münster: Waxmann.

Bach, A., Besa, K.-S. & Arnold, K.-H. (2014). Bedingungen von Lernprozessen in Schulpraktika: Ergebnisse aus dem Projekt ESIS (Entwicklung Studierender in Schulpraktika). In K.-H. Arnold, A. Gröschner & T. Hascher (Hrsg.) Schulpraktika in der Lehrerbildung: Theoretische Grundlagen, Konzeptionen, Prozesse und Effekte. Münster: Waxmann.

Reinthoffer, B. & Dörr, G. (2008). Zur Wirksamkeit Schulpraktischer Studien. In Rotermund, M., Dörr, G. & Bodensohn, R. (Hrsg.): Bologna verändert die Lehrerbildung. Auswirkungen der Hochschulreform. Leipzig: Univ.-Verlag.

Terhart, E. (2002). Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz. Westfälische Wilhelms-Universität Münster (ZKL-Texte Nr. 23). Münster.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823

Schulpraktika in der ersten Phase der Lehrerinnen- und Lehrerbildung sind aktuell obligatorisch für alle Lehramtsstudiengänge an deutschen Hochschulen (vgl. Glumpler & Wildt, 2000). Die Vermittlungs-, Erprobungs- und Orientierungsfunktion von Praktika in der universitären Lehrerinnen- und Lehrerbildung ist wenig umstritten, sehr wohl aber die Ausgestaltung ihrer Formen (vgl. Topsch, 2004). Die Verbindung von Theorie und Praxis als zentraler Bestandteil der Ausbildung, aber auch als zentrales Problem, wird immer wieder diskutiert und bleibt dabei nicht auf die Lehramtsstudiengänge beschränkt, sondern wird auch in anderen universitären Berufsausbildungen artikuliert (vgl. Bohnsack, 2000). Der vorliegende Beitrag untersucht den Einfluss verschiedener Variablen auf die Entwicklung der Kompetenzselbsteinschätzung der Studierenden im Rahmen von Schulpraktika und nutzt dabei die Definition von Klieme und Hartig, bei der Kompetenzen „als erlernbare kontextspezifische Leistungsdispositionen verstanden [werden], die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme & Hartig, 2007, S. 17).

Die Lehrerselektionserwartung als Überzeugung der Lehrperson, den Unterricht erfolgreich zu meistern und auch mit schwierigen Schülerinnen und Schülern umgehen zu können, ist eine der Lehrervariablen, deren Korrelation mit Schülerleistungen schon vor über 25 Jahren identifiziert werden konnte (vgl. Armor et al., 1976). Der Begriff der epistemologischen Überzeugungen beschreibt individuelle Ansichten, Auffassungen und Theorien über die Entstehung, Entwicklung, Bedeutung und Gültigkeit von Wissen in den Wissenschaften (vgl. Priemer, 2006). Sie werden u.a. als Erklärungsvariablen für die Lernwirksamkeit von Unterricht verwendet (vgl. u.a. Baumert et al., 2000). Eine weitere Variable innerhalb der Schulpraktika stellen die betreuenden Lehrkräfte dar. Ihre Rolle kann im Rahmen des Modelllernens nach Bandura (1976) interpretiert werden, so dass die Studierenden sowohl Handlungs- und Planungsmuster von diesen Modellen annehmen können, als auch verschiedene Einstellungen und Werte. Die empirische Untersuchung analysiert, wie sich der Einfluss der Mentorinnen und Mentoren in zwei unterschiedlichen Praxisphasen des Lehramtsstudiums auf die Entwicklung der Studierenden bezüglich ihrer selbsteingeschätzten Kompetenz zur Unterrichtsplanung auswirkt.

In einer Längsschnittstudie wurden dazu alle Lehramtsstudierenden der Universität Hildesheim befragt, die zum Wintersemester 2009/2010 im ersten Semester immatrikuliert waren. Die Studierenden wurden jeweils vor und nach den Praktika zu verschiedenen Merkmalen und ihren selbsteingeschätzten Fähigkeiten in der Unterrichtsplanung befragt. Zusätzlich wurden die beteiligten Mentorinnen und Mentoren befragt. Die Allgemeindidaktische Planungskompetenz wurde mit dem Instrument APK (vgl. Bach, Besa, Brodhäcker & Arnold, 2012) erhoben. Als Einflussfaktoren werden neben soziodemographischen Merkmalen die Lehrerselektionserwartung (nach Schmitz & Schwarzer, 2002), epistemologische Überzeugungen mit dem Erhebungsinstrument von Schraw, Bendixen und Dunkle (2004) sowie der Einfluss der Mentorinnen und Mentoren im Sinne des Modelllernens (vgl. Bandura, 1976) genutzt. Um die aufgeführten Fragestellungen zu bearbeiten, wird eine Modellierung der verschiedenen Variablen in Mplus in einem Strukturgleichungsmodell vorgenommen. Die Ergebnisse zeigen, dass nur wenige Merkmale der Studierenden für die Kompetenzselbsteinschätzung statistisch bedeutsam sind. Die Lehrerselektionserwartung sowie die epistemologischen Überzeugungen der Studierenden wirken sich auf die Kompetenzzuwächse in beiden Praktika aus, jedoch in unterschiedliche, zum Teil nicht theoriekonforme Richtungen und zu verschiedenen Zeitpunkten. Die Merkmale der Mentorinnen und Mentoren (Geschlecht sowie epistemologische Überzeugungen) besitzen keinen oder nur einen geringen Einfluss auf die selbsteingeschätzte Kompetenzentwicklung der Studierenden. Gleiches gilt für den Einfluss der Beziehungsqualität zwischen Studierenden und Mentorinnen und Mentoren. Es zeigen sich zwar theoriekonforme positive Einflüsse, diese sind jedoch statistisch nicht bedeutsam. Die selbsteingeschätzte Planungskompetenz der Mentorinnen und Mentoren weist in keinem der Praktika einen signifikanten Einfluss auf die Kompetenzentwicklung der Studierenden auf. Insgesamt zeigen sich also nur wenige Einflussfaktoren, die in der analysierten Stichprobe statistisch bedeutsamen Einfluss auf die selbsteingeschätzte Kompetenzentwicklung der Studierenden haben. Es werden jedoch einige Ansatzpunkte für weitere Analysen innerhalb der Lehrerinnen- und Lehrerbildungsforschung und auch Entwicklungspotentiale für die Untersuchung der Praktika an der Universität Hildesheim deutlich.

- Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A.P.E. & Zellman, G. (1976). Analysis of the school preferred reading program in selected Los Angeles minority schools (R-2007-LAUSD). Santa Monica: Rand.
- Bach, A., Besa, K.-S., Brodhäcker, S. & Arnold, K.-H. (2012). Kompetenzentwicklung in Schulpraktika. Erfassung allgemeindidaktischer Kompetenz zur Planung, Durchführung und Analyse von Unterricht. In T. Hascher & G.H. Neuweg (Hrsg.), *Forschung zur (Wirksamkeit) der LehrerInnenbildung* (S. 105-122). Münster: LIT.
- Bandura, A. (1976). *Lernen am Modell*. Stuttgart: Klett.
- Baumert, J., Bos, W., Brockmann, J., Gruehn, S., Klieme, E., Köller, O., Lehmann, R., Lehrke, M., Neubrand, J., Schnabel, K.U., Schwippert, K., Watermann, R. (2000). TIMSS/III-Deutschland. Der Abschlussbericht. Zusammenfassung ausgewählter Ergebnisse der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur mathematischen und naturwissenschaftlichen Bildung am Ende der Schullaufbahn. Berlin: Max-Planck-Institut für Bildungsforschung.
- Bohnsack, F. (2000). Probleme und Kritik der universitären Lehrerbildung. In M. Bayer, F. Bohnsack, B. Koch-Priewe & J. Wildt (Hrsg.), *Lehrerin und Lehrer werden ohne Kompetenz? - Professionalisierung durch eine andere Lehrerbildung* (S. 521-23). Bad Heilbrunn: Klinkhardt.
- Glumpler, E. & Wildt, J. (2000). Das Berufspraktische Halbjahr im Lehramtsstudium. In M. Bayer, F. Bohnsack, B. Koch-Priewe & J. Wildt (Hrsg.), *Lehrerin und Lehrer werden ohne Kompetenz?* (S. 207-225). Bad Heilbrunn: Klinkhardt.
- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11-29). Wiesbaden: VS Verlag.
- Priemer, B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 159-175.
- Schmitz, G.S. & Schwarzer, R. (2002). Individuelle und kollektive Selbstwirksamkeitserwartung von Lehrern. In D. Hopf & M. Jerusalem (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (Zeitschrift für Pädagogik, Beiheft 44) (S. 192-214). Weinheim: Beltz.
- Schraw, G., Bendixen, L.D. & Dunkle, M.E. (2004). Development and Validation of the Epistemic Belief Inventory (EBI). In B.K. Hofer (Hrsg.), *Personal epistemology: paradigmatic approaches to understanding students' beliefs about knowledge and knowing*, 39 (S. 261-275). Mahwah: Lawrence Erlbaum Associates.
- Topsch, W. (2004). Schulpraxis in der Lehrerbildung. In S. Blömeke, P. Reinhold, G. Tulodziecki & J. Wildt (Hrsg.), *Handbuch Lehrerbildung* (S. 476-486). Kempten: Klinkhardt.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823



Claudia Voigt (Christian-Albrechts-Universität zu Kiel), Ronja Just (Christian-Albrechts-Universität zu Kiel),  
Uta Klusmann (Christian-Albrechts-Universität zu Kiel), Michaela Köller (Christian-Albrechts-Universität zu Kiel)

## **Welche Effekte hat das Selbsterkundungsverfahren Career Counseling for Teachers (CCT)? Eine quasi-experimentelle Studie mit Schülerinnen und Schülern**

Immer mehr Hochschulen empfehlen Studieninteressierten die Bearbeitung zumeist webbasierter Selbsterkundungsverfahren, so genannter Self-Assessments vor Studienbeginn (Hell, 2009). Selbsterkundungsverfahren bestehen aus Fragebögen und Tests zu berufsbezogenen Interessen, der Persönlichkeit und gegebenenfalls studienrelevanten Fähigkeiten. Direkte Rückmeldungen über die Passung der Interessen und Fähigkeiten zu den Anforderungen eines bestimmten Studiengangs bzw. Berufes sollen die Studierenden bei der Orientierung hinsichtlich der Studien- und Berufswahl unterstützen. Im Rahmen fachspezifischer Self-Assessments soll weiterhin der Prozess der Selbstselektion angeregt werden, in welchem geeignete Studieninteressierte in ihrer Wahl bestärkt, die ungeeigneten hingegen nach Rückmeldung von einer Bewerbung um den entsprechenden Studienplatz absehen (Mayr, 2010). Eine gelungene (Aus-)Wahl wirkt sich nach Schuler und Hell (2008) positiv auf das Leistungsverhalten – und damit das Wohlbefinden - der Studierenden aus und schlägt sich in besseren Noten, kürzeren Studienzeiten sowie geringerer Studienabbruchquoten nieder.

Ein prominentes, fachspezifisches Selbsterkundungsverfahren ist das Programm Career Counselling for Teachers (CCT). Das CCT soll angehende Lehramtskandidaten bei ihrer Entscheidung für oder gegen ein Lehramtsstudium unterstützen. Das Programm bietet die Möglichkeit in Internet Tests zur Persönlichkeit und berufsbezogenen Interessen zu bearbeiten sowie erste Praxiserfahrungen zu reflektieren. Durch die Rückmeldung über Passung der Merkmale zu beruflichen Anforderungen sollen angehende Studierende zur Reflexion ihrer Eignung für den Lehrerberuf und somit der Selbstselektion angeregt werden (Köller, Klusmann, Retelsdorf und Möller, 2012; Nieskens und Müller, 2009).

Bisherige Evaluationsergebnisse des CCT verdeutlichen, dass Studieninteressierte mit günstigen Voraussetzungen die Rückmeldung von CCT als Ermutigung erleben, Personen mit ungünstigen Voraussetzungen berichten gegenteilige Impulse (Mayr, 2012). Durch dieses Verfahren lassen sich demnach gut geeignete Personen auf den Lehrerberuf aufmerksam machen und für diesen gewinnen. Fraglich ist jedoch, inwieweit sich Personen mit einer negativen Rückmeldung davon abhalten lassen, Lehramt zu studieren und ihre Studienwahlentscheidung entsprechend der Rückmeldung verändern. Es steht aus, zu untersuchen, welchen Effekt die Information und Beratung auf die tatsächliche Studienwahlentscheidung hat. (Klusmann, Köller, Kunter, 2011).

Ziel der Arbeit ist es, erstens zu untersuchen, welche Studieninteressierten sich nach Bearbeitung des CCT in ihrer Studienwahl umentscheiden. Weiter soll zweitens die Frage untersucht werden, ob sich Studieninteressierte, die Interesse an einem Lehramtsstudium zeigen, besser durch das CCT informiert fühlen als durch allgemeine Studieninteressentests. Drittens wird der Einfluss der Vorstellung unterrichten oder erziehen zu können bzw. Lehrer/in zu werden auf diese Informationsfunktion des CCT und die Bewertung dieser im Vergleich zu anderen Verfahren untersucht.

An der quasi-experimentellen Untersuchung nahmen insgesamt 529 Schüler/innen der Sekundarstufe II von 10 Gymnasien in Schleswig-Holstein teil. Im Prä-Post-Vergleich wurde dann randomisiert eines von vier Verfahren (CCT, „was-studiere-ich.de“; Hell, Päßler & Schuler, 2009, EXPLORIX®; Jörn, Stoll, Bergmann & Eder, 2012; oder eine allgemeine Studienberatung) zugeteilt. Alle Self-Assessments und Fragebögen wurden computerbasiert durchgeführt, die allgemeine Studienberatung wurde von Psychologiestudentinnen begleitet und fand in einem separaten Klassenraum statt. Insgesamt bearbeiteten durch diese Zuteilung 52 Schüler/innen das CCT.

Erste Ergebnisse zeigen, dass 21% der 52 Studieninteressierten nach Bearbeitung sowie Rückmeldung ihren Studienwunsch veränderten. Um auf die Untersuchung der Funktion der Selbstselektion des CCT hinzuleiten, wurde bei Untersuchung der zweiten Fragestellung mittels Varianzanalysen (ANOVA) deutlich, dass Studieninteressierte über die verschiedenen Self-Assessments (CCT, Explorix®; Jörn, Stoll, Bergmann & Eder, 2012; „was-studiere-de.de“; Hell, Päßler & Schuler, 2009) einen signifikanten Unterschied hinsichtlich der Wahrnehmung allgemeiner Informationen, persönlicher Informationen sowie der Bewertung dieser aufzeigen. Die Tendenz wird sichtbar, dass das CCT zukünftigen Studierenden weniger Informationen darbietet, welche im Vergleich zu Verfahren allgemeiner Studienwahl weniger positiv bewertet werden. Über eine ANCOVA wurde im Hinblick auf die dritte Fragestellung deutlich, dass die Vorstellung zu unterrichten oder erziehen bzw. Lehrer/in zu werden einen signifikanten Einfluss auf die Wahrnehmung persönlicher Informationen durch die Verfahren darstellt.

Career Counseling for Teachers, CCT (o.J.). Zum Hintergrund der Selbsterkundungs-Verfahren in CCT URL: <http://www.cct-germany.de/de/0/pages/Index/48> [letzter Zugriff: 23.10.2013]

Hell, B. (2009). Selbsttests zur Studienorientierung: nützliche Vielfalt oder unnützer Wildwuchs? In G. Rudinger & K. Hörsch (Hrsg.), Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung (S. 9–20). Göttingen: V&R.

Hell, B., Päßler, K. & Schuler, H. (2009). "was-studiere-ich.de" - Konzept, Nutzen und Anwendungsmöglichkeiten. Zeitschrift für Beratung und Studium, 4 (1), 9–14.

Jörn, S., Stoll, F., Bergmann, C. & Eder, F. (2012). EXPLORIX - Das Werkzeug zur Berufswahl und Laufbahnplanung. Bern: Huber.

Klusmann, U., Köller, M. & Kunter, M. (2011). Anmerkungen zur Validität eignungsdiagnostischer Verfahren bei angehenden Lehrkräften. Zeitschrift für Pädagogik, 57, 711-721.

Köller, M., Klusmann, U., Retelsdorf, J. & Möller, J. (2012). Geeignet für den Lehrerberuf? Self-Assessments auf dem Prüfstand. Zeitschrift für Unterrichtswissenschaft, 40, 121-139.

Mayr, J. (2010). Ein Lehrerstudium beginnen. In T. Janik & P. Knecht (Hrsg.), New pathways in the professional development of teachers (pp. 11-18). Münster: Lit.

Mayr, J. (2012). Ein Lehramtsstudium beginnen? Ein Lehramtsstudium beginnen lassen? Laufbahnberatung und Bewerberauswahl konstruktiv gestalten. In B. Weyand, M. Justus, M. Schratz (Hrsg.) Auf unsere Lehrerinnen und Lehrer kommt es an. Geeignete Lehrer/-innen gewinnen, (aus-)bilden und fördern (S. 38-57). Essen: Stifterverband für die Deutsche Wissenschaft

Nieskens, B. & Müller, F. (2009). Soll ich LehrerIn werden? Web-basierte Selbsterkundung persönlicher Voraussetzungen und Interessen. Erziehung und Unterricht, 159, 41-49.

Schuler, H., Hell, B. (2008) Studierendenauswahl und Studienentscheidung aus eignungsdiagnostischer Sicht. In H. Schuler (Hrsg.), Studierendenauswahl und Studienentscheidung (S. 11-17). Göttingen: Hogrefe



## Welche Lerneffekte zeigen sich im Zusammenhang mit dem Planspiel „Schulalltag“?

Im Klassenzimmer zu unterrichten ist eine komplexe Aufgabe (Brophy & Good, 1997). Vom Umgang mit Komplexität ist bekannt, dass stets Emotionen mit im Spiel sind, die das Nachdenken, Problemlösen und Handeln beeinflussen (Abele, 1995; Hänze, 1998; Spering, Wagner, & Funke, 2005). Menschen, die erfahren und gelernt haben, ihre Emotionen zu erkennen und zu regulieren, gelangen auch zu besseren Problemlösungen in einer komplexen Problemsituation (Starker, 2012). Das Planspiel macht komplexe Anforderungen und emotionale Reaktionen für Lernprozesse in komplexen Situationen verfügbar. Auf dem Gebiet des Managementtrainings, sowie des Notfall- und Krisenstabtrainings wird sie bereits erfolgreich eingesetzt. (Badke-Schaub, Lauche, & Hofinger, 2008; Pawlowsky & Mistele, 2008). Bei der Planspielmethode besteht die Möglichkeit reflektiertes Handlungswissen zu erwerben, was normalerweise erfahrenen Praktikern vorbehalten ist (Schön, 1983). Ein Planspiel dieses Typs wurde im Rahmen dieser erstmals für die Situation von Lehrern und Lehrerinnen im Unterricht entwickelt.

Methode: Planspielkonzept: Im komplexen Planspiel Schulalltag besteht die Möglichkeit, eigene Erfahrungen in verschiedenen Rollen des Schulgeschehens zu machen und systematisch aufzuarbeiten. Ziel ist es, Dimensionen der eigenen und fremden Emotionen wahrzunehmen, theoretisches Wissen für das Handlungsrepertoire verfügbar zu machen und Möglichkeiten der Emotionsregulation kennenzulernen und auf ihre Wirkung auszuloten. Im Planspiel werden die Teilnehmenden gebeten, jeweils für eine fiktive Zeit von zwei Schultagen (Realzeit vier Stunden) Rollen als Lehrende oder Lernende zu übernehmen. Anschließend werden die Dynamiken, die sich beispielsweise zwischen den Schülern, oder den Schülern und den Lehrenden ergeben reflektiert und vor einem theoretischen Hintergrund reflektiert.

Erfassung von Verhalten und Ergebnis: Es fragt sich nun, ob diese Lerneffekte hinsichtlich Emotionsregulation, Fähigkeit zur Analyse und Reflexion schwieriger Situationen. Dies soll zunächst im Rahmen dieser Pilotstudie (n=21) überprüft werden. Da dieses Planspiel und seine Zielsetzungen bisher so noch nicht durchgeführt wurden, wurden sowohl qualitative Daten aus offenen Interviews gesammelt, um mögliche weitere Indikatoren für den Einfluss dieses Verhaltens auf den Umgang mit Komplexität zu entdecken als auch Beobachtungen auf der Basis von a-priori definierten Kategorien erhoben. Diese bestanden zum einen aus einem Rating bezüglich der Gesamtstimmung, das alle 10 Minuten erhoben wurde, daneben wurden kritische Ereignisse festgehalten. Zum Vorher-Nachher-Vergleich des Reflexionsgrades der Beteiligten wurden Fallvignetten verwendet und verglichen, welche Eigenschaften die Vorschläge zur Problemlösung hatten und hinsichtlich feststehender Kategorien verglichen. Die Reliabilität der a-priori-Kategorien wurde mit dem Intraklassenkoeffizient nach Wirtz und Caspar (2002) geprüft und war zufriedenstellend. Auf qualitativer Datenebene (Bortz & Döring, 2009) wurde die offene Beschreibung einzelner Verläufe durch Beobachter, die auch in die Reflexion des fiktiven Unterrichts herangezogen und so kommunikativ validiert. Die generelle Fragestellung bestand darin, festzustellen, welche Lerneffekte sich im Zusammenhang mit dem Planspiel zeigen. Diese lassen sich in den folgenden beiden Hypothesen konkretisieren: Hypothese 1: Es ändern sich Handlungstendenzen in emotionaler Hinsicht. Hypothese 2: In den Fallanalysen wird theoriebezogen reflektiert. Ergebnisse: Global wurde das Training in der Nachbesprechung als nützlich bewertet, dies spiegelt sich z.B. in folgenden Aussagen wieder: Hoher Spaßfaktor aber auch hoher Lernfaktor, gute Vorbereitung, die Tipps und Anleitungen sehr passend und die Vorbereitung mit Theorie teil beim ersten Termin war sehr gelungen und informativ. Wertet man die Verlaufsdaten des Stimmungsratings aus, erweisen diese sich als sehr volatil. In der detaillierten Analyse kann den Teilnehmern rückgemeldet werden, wie durch Wechselwirkungen zwischen Lehrerverhalten und Reaktionen schnell sich aufschaukelnde Stimmungsschwankungen entstehen. Die deutlichsten Effekte zeigen sich beim Vorher-Nachher-Vergleich der Fallvignetten. Hier zeigte sich als Ergebnis mit Chi-Quadrat eine stärkere Bevorzugung reflektierter Handlungsweisen bei der zweiten Durchführung nach der Veranstaltung.

Diskussion und Ausblick: Die erste Durchführung des Planspiels Schultag wurde positiv bewertet und zeigte die gewünschten Effekte. Dies ermutigt dazu, dieses Format weiter zu entwickeln und einen verallgemeinerten Nachweis der beschriebenen Effekte anzustreben, da diese bei einer einmaligen Durchführung zum Teil natürlich von den Rahmenbedingungen und den Lehrenden abhängig sein können.

- Abele, A. (1995). Stimmung und Leistung. Göttingen: Hogrefe.  
Badke-Schaub, P., Lauche, K. & Hofinger, G. (2007). Human Factors. Berlin: Springer.  
Bortz, J. & Döring, N. (2009). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin: Springer.  
Good, T.L., & Brophy, J.E. (1997). Looking in classrooms. New York: Longman.  
Hänze, M. (1998). Denken und Gefühl. Weinheim: Beltz.  
Pawlowsky, P. & Mistele, P. (2008). Hochleistungsmanagement. Leistungspotenziale in Organisationen gezielt fördern. Wiesbaden: Gabler.  
Schön, D. (1983). Reflective practitioner. New York: Basic Books.  
Spering, M., Wagner, D. & Funke, J. (2005). The role of emotions in complex problem-solving. Cognition and Emotion, 19, 1252-1261.  
Starker, U. (2012). Emotionale Adaptivität. Lengerich: Pabst.  
Sutton, R. E. (2004): Emotional regulation goals and strategies of teachers. Social Psychology and Education, 7. 379-398.  
Von der Weth, Strohschneider (1993). Ja, mach nur einen Plan. Bern: Huber.  
Wirtz, M. & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe.



## Aspekte gelingender Qualifizierung – Eine Studie aus verschiedenen Blickwinkeln

Die Professionalisierung pädagogischer Fachkräfte wird in der Diskussion um die Entwicklung pädagogischer Qualität in frühpädagogischen Institutionen zunehmend in den Mittelpunkt gerückt. Im Bereich der sprachlichen Bildung wurde versucht, über spezifische Sprachförderansätze und -konzepte dem Bildungsauftrag gerecht zu werden, allerdings zeigen diese Ansätze nur geringe oder keine Wirkungen, wobei evidenzbasiert nicht beantwortet werden kann, warum keine Wirkungen erzielt werden (Schöler & Roos, 2010; Wolf et al., 2011; Sachse et al., 2012; Jungmann et al., 2013). Theoretisch werden eine mangelhafte Fort- und Weiterbildung der Fachkräfte, die der theoretischen und praktischen Komplexität der Handlungsanforderungen nicht gerecht wird, sowie fehlende Reflexions- und Supervisionsangebote als Ursachen diskutiert (Schöler & Roos, 2010). An dieser Diskussion setzt das Projekt „Qualifizierungsoffensive: Sprachliche Bildung und Förderung für Kinder unter Drei“ an, in dessen Rahmen zu dem alltagsintegrierten Sprachförderansatz „Die Sprache der Jüngsten entdecken und begleiten“ (Jampert et al., 2011; vgl. auch Best et al., 2011) ein Qualifizierungskonzept für die Fort- und Weiterbildung entwickelt wird. Das Qualifizierungskonzept sieht vor, dass über einen längeren Zeitraum ganze Kita-Teams prozessbegleitend geschult und gecoacht werden, wobei Theorie-, Praxis- und Reflexionsphasen miteinander verschränkt werden. Das Projekt ist in die Offensive „Frühe Chancen: Schwerpunkt-Kitas Sprache & Integration“ eingebunden, in der zwischen 2011 und 2014 insgesamt 4000 Einrichtungen hinsichtlich alltagsintegrierter Sprachbildung gefördert werden. Zwischen Ende 2011 und Frühjahr 2013 wurden etwa 120 dieser Kindertageseinrichtungen von 60 speziell geschulten MultiplikatorInnen nach dem DJI-Ansatz qualifiziert und gecoacht. Die wissenschaftliche Begleitung der Qualifizierung nach dem DJI-Ansatz verfolgt zwei Ziele: Erstens soll die Gestaltung der Qualifizierung in den Einrichtungen beleuchtet werden und zweitens die Veränderungen im pädagogischen Alltag durch die Qualifizierung analysiert werden. Für diese Zielstellungen wird multiperspektivisch und methodentriangulierend vorgegangen. Neben den Perspektiven der MultiplikatorInnen werden die Perspektiven der Teams durch die SprachexpertInnen der Einrichtungen einbezogen und zueinander in Beziehung gesetzt. Eine längsschnittlich angelegte, standardisierte Online-Befragung mit drei Erhebungszeitpunkten (zu Beginn, in der Mitte, nach Ende der Qualifizierung) die Verlaufsanalysen (N=107) ermöglicht, stellt einen methodischen Zugang dar. In dieser Befragung wurden geschlossene wie offene Fragen integriert. Methodentriangulierend werden diese Erhebungen durch Gruppendiskussionen mit einer Teilgruppe der SprachexpertInnen ergänzt, um vertiefende Informationen über die Umsetzung der Qualifizierung und angestoßene Veränderungen aus Sicht der Praxis zu erhalten. Durch das Design können sowohl Auskünfte zu den einzelnen Erhebungszeitpunkten als auch Längsschnittanalysen durchgeführt werden. Die statistische Datenanalyse erfolgt sowohl uni-, bi- als auch multivariat (z.B. Varianzanalysen, Zusammenhangsanalysen). Für die Auswertung der offenen Fragen und der Gruppendiskussionen wird inhaltsanalytisch vorgegangen. Die Posterpräsentation soll dazu genutzt werden, neben dem Forschungskontext und dem Forschungsdesign wesentliche Erkenntnisse der wissenschaftlichen Begleitung zu präsentieren. So geben die ersten Ergebnisse Hinweise, welche Aspekte des Qualifizierungsprozesses als gelungen angesehen, sowie welche Veränderungen in der pädagogischen Arbeit angestoßen werden können. In den Ergebnissen zur Gestaltung der Qualifizierung zeigt sich beispielsweise, dass für die Umsetzung der Inhalte die Verschränkung von Theorie-, Praxis- und Reflexionsphasen bedeutsam ist. Verlaufsanalysen zeigen auf, inwiefern die Inhalte umgesetzt werden, welche didaktischen Methoden für die Vermittlung hilfreich sind und welche Wissensentwicklung sich in den Teams vollzieht. Durch die Qualifizierung werden in den Einrichtungen auf verschiedene Weise Veränderungen angestoßen, wobei besonders Wissenserweiterung, Haltungsveränderung, Kontinuität in Beobachtung und Dokumentation, verstärkter Teamaustausch und vermehrte Reflexionen genannt werden können. Die Gruppendiskussionen untermauern diese Veränderungen durch vielfältige Umsetzungsstrategien. Aus den Erhebungen wird ersichtlich, dass es nicht eine Umsetzung des DJI-Konzepts gibt, sondern vielfältige Wege. Analysiert werden soll, ob sich zu Qualifizierungsende Muster erkennen lassen, in welchen Kontexten (z.B. große Teams, unterschiedliche Vorerfahrungen) und mit welchen Qualifizierungsmethoden sich verschiedene Qualifizierungsverläufe sowie Praxisveränderungen erkennen lassen.

- Best, P./Laier, M./Jampert, K. & Leuckefeld, K. (2011). Dialoge mit Kindern führen. Die Sprache der Kinder im dritten Lebensjahr beobachten, entdecken und anregen. Her-gegeben von der Stiftung Baden-Württemberg. Weimar/Berlin.
- Jampert, K./Thanner, V./Schattel, D./Sens, A./Zehnauer, A./Best, P. & Laier, M. (2011). Die Sprache der Jüngsten entdecken und begleiten. Sprachliche Bildung und Förderung für Kinder unter Drei. Deutsches Jugendinstitut. Weimar, Berlin: verlag das netz.
- Jungmann, T., Koch, K. & Etzien, M. (2013). Effektivität alltagsintegrierter Sprachförderung bei ein- und zwei- bzw. mehrsprachig aufwachsenden Vorschulkindern. In: Frühe Bildung, 2 (3), 110-121.
- Sachse, S./Budde, N./Rinker, T. & Groth, K. (2012): Evaluation einer Sprachfördermaß-nahme für Vorschul Kinder. In: Frühe Bildung, 1 (4), 194-201.
- Schöler, H. & Roos, J. (2010): Ergebnisse einer Evaluation von Sprachfördermaßnahmen in Mannheimer und Heidelberger Kitas. In: Fröhlich-Gildhoff, K., Nentwig-Gesemann, I. & Strehmel, P. (Hrsg.): Forschung in der Frühpädagogik III. Schwerpunkt: Sprachentwicklung & Sprachförderung. Materialien zur Frühpädagogik, Band 5. Freiburg, 35-74.
- Wolf, K. M./Felbrich, A./Stanat, P. & Wendt, W. (2011): Evaluation der kompensatorischen Sprachförderung in Brandenburger Kindertagesstätten. In: Empirische Pädagogik, 25 (4), 423-438.





## **Der Einfluss der Domäne auf die allgemeinen epistemologischen Überzeugungen am Beispiel der Biologie und der Wirtschaftswissenschaften**

Theoretischer Hintergrund: Epistemologische Überzeugungen sind individuelle Überzeugungen zu Wissen und Wissenserwerb (vgl. Hofer & Pintrich, 1997). Ihnen werden sowohl handlungsleitende als auch handlungssteuernde Funktionen zugeschrieben (vgl. Müller, 2007; Wahl, 1979). Empirische Studien belegen Zusammenhänge zwischen epistemologischen Überzeugungen von Lernenden und Faktoren wie z. B. Lernstrategien und Motivation (vgl. Mokwinski, 2011; Zinn, 2013). Darüber hinaus zeigen sich Zusammenhänge zwischen den epistemologischen Überzeugungen von Lehrenden und ihrem unterrichtlichen Handeln (vgl. z. B. Aypay, 2011; Chai Teo & Lee, 2010; Chan & Elliot, 2004). Folglich kommt den epistemologischen Überzeugungen von Lehrenden und Lernenden eine wesentliche Bedeutung für Lehr-Lern-Prozesse zu. Während frühere Publikationen davon ausgehen, dass sich epistemologische Überzeugungen in einer Dimension abbilden (vgl. z. B. King & Kitchener, 2002; Perry, 1970), postulieren spätere Veröffentlichungen ein mehrdimensionales Modell (vgl. z. B. Buehl & Alexander, 2006; Hofer & Pintrich, 1997). Vor allem durch den Ansatz von Schommer (1990) vollzog sich dieser konzeptionelle Wechsel hin zu einem multidimensionalen Ansatz voneinander unabhängiger Dimensionen, bei denen es auch zu rekursiven Entwicklungen kommen kann. Schommers Modell beinhaltet neben Vorstellungen zu Wissen auch Überzeugungen zum Lernen und ist folglich für Untersuchungen schulischen Lehrens und Lernens besonders relevant (vgl. Schommer-Aikins, 2002). Ein weiterer Konsens besteht dahingehend, dass sowohl allgemeine als auch domänenbezogene epistemologische Überzeugungen existieren (vgl. z. B. Buehl & Alexander, 2006; Hofer, 2006; Zinn, 2013). Domäne wird hier oft als akademische Disziplin verstanden. So lässt sich beispielsweise die Domäne „Biologie“ nach dem Klassifikationsschema von Biglan (1973) eher als „hart“ und „theoretisch“ einschätzen, die Wirtschaftswissenschaften hingegen eher als „weich“ und „anwendungsorientiert“. Bislang vergleicht eine Vielzahl an Studien verschiedene Domänen entlang vorher festgelegter Dimensionen (vgl. z. B. Schommer & Walker, 1995; Paulsen & Wells, 1998). Es gibt allerdings nur wenige Studien, welche das Vorliegen verschiedener Dimensionen in unterschiedlichen Domänen untersuchen (vgl. z. B. Müller, 2007). So nimmt insbesondere das Modell von Buehl und Alexander (2006) an, dass bestimmte Dimensionen nur in bestimmten Domänen existieren und allgemeine sowie bereichsspezifische Überzeugungen in einer wechselseitigen Abhängigkeit stehen. Vor diesem Hintergrund stellt sich die Frage: „Unterscheiden sich die allgemeinen epistemologischen Überzeugungen in den Wirtschaftswissenschaften von denen in der Biologie?“

Methode: Zur Beantwortung der Forschungsfrage wurden in einer ersten Annäherung 150 angehende Lehrende der Biologie, 276 angehende Handelslehrende und 653 angehende Bankkaufleute zu ihren allgemeinen epistemologischen Überzeugungen befragt. Als Erhebungsinstrument diente der von Paechter, Rebmann, Schlömer, Mokwinski, Hanekamp und Arendasy (2013) entwickelte Oldenburg Epistemic Beliefs Questionnaire (OLEQ), der auf dem Modell von Schommer (1990) basiert.

Ergebnisse: Im Rahmen einer konfirmatorischen Faktorenanalyse konnte die in der Entwicklungsstudie von Paechter et al. (2013) ermittelte Faktorstruktur des OLEQs in keiner der drei Stichproben repliziert werden. Eine sich anschließende exploratorische Faktorenanalyse deutet hingegen an, dass bei den Bankkaufleuten und Handelslehrenden sechs Faktoren relevant sind, von denen sich fünf als die Dimensionen aus dem Modell von Schommer (1990) interpretieren lassen. Bei den Handelslehrenden zielt der sechste Faktor auf die Relativität von Wissen ab, bei den Bankauszubildenden entsteht ein Faktor, der die Entwicklung von Verständnis beschreibt. Bei den Biologielehrenden lassen sich nur zwei Faktoren in Form der „Fähigkeit zu Lernen“ und der „Quelle des Wissens“ mit dem Modell von Schommer (1990) vereinbaren. Darüber hinaus besteht ein Faktor, der als „Überzeugungen über Lernen“ gedeutet werden kann. Schließlich lässt sich ein weiterer Faktor identifizieren, der sich als Überzeugungen zur „Eindeutigkeit von Wissen“ interpretieren lässt, wie er ähnlich auch in der TIMSS/III-Studie berücksichtigt wird (vgl. Köller, Baumert & Neubrand, 2000). Die Ergebnisse zeigen, dass die Domäne bereits Einflüsse auf die Struktur allgemeiner epistemologischer Überzeugungen nimmt. Die Resultate bestätigen und erweitern das Modell von Buehl und Alexander (2006) und machen deutlich, dass die Domäne bei der Erhebung allgemeiner Vorstellungen zu Wissen und Wissenserwerb berücksichtigt werden muss.

- Aypay, A. (2011). The adaption of the Teaching-Learning Conceptions Questionnaire and its relationships with epistemological beliefs. *Educational Sciences: Theory and Practice*, 11(1), 21–29.
- Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57(3), 195–203.
- Buehl, M. M. & Alexander, P. A. (2006). Examining the dual nature of epistemological beliefs. *Educational Research*, 45, 28–42.
- Chai, C. S.; Teo, T. & Lee, C. B. (2010). Modelling the relationships among beliefs about learning, knowledge, and teaching of pre-service teachers in Singapore. *The Asia-Pacific Education Researcher*, 19(1), 25–42.
- Chan, K.-W. & Elliott, R. G. (2004). Relational analysis of personal epistemology and conceptions about teaching and learning. *Teaching and Teacher Education*, 20(8), 817 – 831.
- Conley, A. M.; Pintrich, P. R.; Vekiri, I. & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29, 186–204.
- Hofer, B. K. (2006). Beliefs about knowledge and knowing: Integration domain specificity and domain generality: A response to Muis, Benedixen, and Haerle (2006). *Educational Psychology Review*, 18(1), 67–76.
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.
- King, P. M. & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer and P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*, (pp. 37–61). Mahway, NJ: Lawrence Erlbaum, Publisher.
- Köller, O.; Baumert, J. & Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (S. 229 - 269). Opladen: Leske + Budrich.
- Mokwinski, B. (2011). Entwicklungen von epistemologischen Überzeugungen in ausgewählten Berufsfeldern der Berufsausbildung. München: Hampp.
- Müller, S. (2007). Epistemologische Überzeugungen zu Wissen und Wissenserwerb im europäischen Vergleich. In B. Graue, B. A. Mester, G. Siehlmann & M. Westhaus (Hrsg.), *International - Europäisch - Regional* (S. 247 - 265). Oldenburg: BIS.
- Paechter, M. R.; Rebmann, R.; Schlömer, T.; Mokwinski, B.; Hanekamp, Y. & Arendasy, M. (2013). Development of the Oldenburg Epistemic Belief Questionnaire (OLEQ), a German questionnaire based on the Epistemic Belief Inventory (EBI). *Current Issues in Education*, 16(1), 1–17.
- Paulsen, M. B. & Wells, C. T. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education*, 39(4), 365–384.
- Perry, W. G. (1970). Forms of intellectual and ethical development in college years. New York: Holt, Rinehart, and Winston.
- Schommer, M. (1990). Effect of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498–504.
- Schommer, M. & Walker, K. (1995). Are epistemological beliefs similar across domains? *Journal of Educational Psychology*, 87(3), 424–432.
- Schommer-Aikins, M. (2002). An evolving theoretical framework for an epistemological belief system. In B. K. Hofer & B. R. Pintrich (Eds.), *Personal epistemology. The psychology of beliefs about knowledge and knowing* (pp. 103–118). Mahway, NJ: Lawrence Erlbaum, Publisher.
- Wahl, D. (1979). Methodische Probleme bei der Erfassung handlungsleitender und handlungsrechtfertigender subjektiver psychologischer Theorien von Lehrern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 11 (3), 208–217.
- Zinn, B. (2013). *Überzeugungen zu Wissen und Wissenserwerb von Auszubildenden*. Münster: Waxmann.



## Diagnostische Kompetenz von Sekundarstufenlehrkräften im Bereich der Leseleistungen

Theoretischer Hintergrund: Diagnostische Kompetenzen von Lehrkräften stellen im Zusammenspiel mit hoher Unterrichtsqualität eine Voraussetzung für hohe Lernzuwächse dar (Behrmann & Souvignier, 2013). Entsprechend wurden zuletzt Forschungsaktivitäten zur differenzierten Beschreibung diagnostischer Kompetenzen intensiviert. Die meisten dieser Untersuchungen konzentrieren sich jedoch auf Grundschullehrkräfte. Aufgrund des jahrelang bestehenden Klassenlehrerprinzips wird vermutet, dass Grundschullehrer ihre Schüler besser einschätzen können als Fach-Lehrkräfte der Sekundarstufe. Ziel unserer Studie war es daher zu prüfen, ob sich vorliegende Befunde zur diagnostischen Kompetenz bei Lehrkräften weiterführender Schulen replizieren lassen.

Beispielsweise wurde metaanalytisch eine mittlere Korrelation von  $r=.53$  zwischen Lehrerurteilen und Testleistungen gefunden (Südkamp, Kaiser & Möller, 2012). Dabei gelingen Lehrkräften umso treffendere Einschätzungen, je heterogener Schülerleistungen ausfallen (Karing, 2009). Die Urteilsgenauigkeit ist offenbar nicht von der Klassengröße abhängig (Hurwitz, Elliot & Braden, 2007), jedoch werden leistungsstärkere Schüler oft präziser beurteilt (Feinberg & Shapiro, 2009). Die Urteilspräzision hängt aber auch entscheidend vom einzuschätzenden Konstrukt ab (Spinath, 2005). Eine themen- und fachübergreifende Urteilsfähigkeit lässt sich folgerichtig nicht abbilden – inhaltlich verwandte Konstrukte können jedoch offenbar ähnlich gut eingeschätzt werden (Lorenz & Artelt, 2009).

Fragestellungen: Für folgende vier Aspekte wurde bei Sekundarstufen-Lehrkräften geprüft, ob sich diese Befunde replizieren lassen:

1. Wie hoch ist das absolute Niveau der diagnostischen Sensitivität der Lehrkräfte?
2. Wird die Urteilsgenauigkeit der Lehrkräfte von der Leistungsheterogenität ihrer Schüler und/oder der Klassengröße beeinflusst?
3. Geben Lehrkräfte akkuratere Leistungsbeurteilungen für leistungsstarke als für leistungsschwache Schüler ab?
4. Sind Lehrkräfte, die ihre Schüler in einer bestimmten Fähigkeit akkurat beurteilen können, ähnlich genau in ihren Urteilen bei einer inhaltlich verwandten Fähigkeit?

Methode: Die Untersuchung wurde mit 66 Sekundarstufenlehrkräften (1489 Schüler, Klassen 5-7) durchgeführt. Das Durchschnittsalter der Lehrkräfte betrug 44.9 Jahre, das der Schulkinder 11.3 Jahre. Die Leistungen der Schüler wurden im Hinblick auf das Leseverständnis und die Leseflüssigkeit als zwei inhaltlich eng verwandte Konstrukte erhoben. Das Leseverständnis wurde mit dem „Frankfurter Leseverständnistest“ (FLVT), die Leseflüssigkeit mit dem „Salzburger Lesescreening“ (SLS) erfasst. Die Lehrerurteile über das Leseverständnis und die Leseflüssigkeit der Schüler wurden jeweils mit einem fünfstufigen Likert-skalierten Item erhoben. Zur Bestimmung der diagnostischen Kompetenzen der Lehrkräfte wurden die Einschätzungen der Lehrkräfte mit dem entsprechenden objektiven Maß der Leseleistung (FLVT-, bzw. SLS-Ergebnisse) korreliert.

Ergebnisse: Fragestellung 1. Der durchschnittliche „diagnostische Kompetenz“-Koeffizient der Lehrkräfte bezüglich des Leseverständnisses betrug  $r=.35$  und bezüglich der Leseflüssigkeit  $r=.44$  und lag damit etwas niedriger als in vorliegenden Metaanalysen. Fragestellung 2: Zur Überprüfung der zweiten Fragestellungen wurden beide „Kompetenzkoeffizienten“ jeweils durch die Standardabweichung der entsprechenden Leseleistungsvariablen (Leseverständnis, Leseflüssigkeit) sowie durch die Klassengröße vorhergesagt. Die Urteilsgenauigkeit in Bezug auf das Leseverständnis wurde signifikant vom Prädiktor „Streuung der Schülerleistungen“ vorhergesagt ( $\beta=.26$ ). Die Klassengröße wies jedoch keinen prädiktiven Wert auf ( $\beta=-.01$ ). Die Urteilsgenauigkeit bezüglich der Leseflüssigkeit konnte durch keinen der beiden Prädiktoren bedeutsam vorhergesagt werden („Streuung“,  $\beta=-.05$ ; „Klassengröße“,  $\beta=-.04$ ). Fragestellung 3. Um zu überprüfen, ob Lehrkräfte leistungsstarke besser als leistungsschwache Schüler identifizieren können, wurden die Schulkinder für beide Leseleistungsvariablen in je drei gleichgroße Gruppen (schwache, mittelgute, starke Leser) aufgeteilt, für die jeweils ein separater „Kompetenzkoeffizient“ bestimmt wurde. Anschließend wurden jeweils die Leseverständnis- bzw. die Leseflüchtigkeits-Kompetenzkoeffizienten für die leseschwachen Gruppen mit denen der lesestarken Gruppen miteinander verglichen. Es konnten keine signifikanten Unterschiede zwischen den Kompetenzkoeffizienten gefunden werden.

Fragestellung 4. Zur Überprüfung der vierten Fragestellung wurden die Lehrkräfte nach der Güte ihrer Beurteilungen mittels Median-Splits in zwei Gruppen aufgeteilt (bessere vs. schwächere Diagnostiker). Dies geschah sowohl bezüglich des Leseverständnis-Kompetenzkoeffizienten als auch bezüglich des Leseflüchtigkeits-Koeffizienten. Mittels  $\chi^2$ -Tests wurde überprüft, ob Lehrkräfte in Bezug auf die beiden inhaltlich verwandten Bereiche vergleichbar gute Diagnostiker sind. Es ließ sich kein solcher Zusammenhang statistisch nachweisen ( $p = 1.0$ ). Insgesamt zeigt sich über alle Einzelergebnisse hinweg, dass die diagnostischen Kompetenzen von Sekundarstufen-Lehrkräften etwas niedriger ausfallen als in der allgemeinen Literatur berichtet.

Behrmann, L., & Souvignier, E. (2013). The relation between teachers' diagnostic sensitivity, their instructional activities, and their students' achievement gains in reading. *Zeitschrift für Pädagogische Psychologie*, 27, 283-293. doi: 10.1024/1010-0652/a000112

Hurwitz, J. T., Elliott, S. N., & Braden, J. P. (2007). The influence of test familiarity and student disability status upon teachers' judgments of students' test performance. *School Psychology Quarterly*, 22, 115-144. doi: 10.1037/1045-3830.22.2.115

Karing, C. (2009). Diagnostische Kompetenz von Grundschul- und Gymnasiallehrkräften im Leistungsbereich und im Bereich Interessen. *Zeitschrift für Pädagogische Psychologie*, 23, 197-209. doi: 10.1024/1010-0652.23.34.197

Lorenz, C. & Artelt, C. (2009). Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. *Zeitschrift für Pädagogische Psychologie*, 23, 211-222. doi: 10.1024/1010-0652.23.34.211

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19, 85-95. doi: 10.1024/1010-0652.19.12.85

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104, 743-762. doi: 10.1037/a0027627

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
Raum 823



## Der Einfluss individueller Merkmale und schulischer Kontextbedingungen auf die inklusiven Einstellungen von Grundschullehrkräften

Professionelles Lehrerhandeln in inklusiven Kontexten steht derzeit im Fokus vielfältiger Forschungsinitiativen. In diesem Zusammenhang werden insbesondere die Einstellungen der Lehrerinnen und Lehrer als Bestandteil der Handlungskompetenz (Baumert & Kunter, 2006) untersucht und ihre Einflüsse auf Schülervariablen betrachtet (Hartinger, Kleickmann & Hawelka, 2006; Küsting, Billich & Lipowsky, 2009; Voss, Kleickmann, Kunter & Hachfeld, 2011). Darüber hinaus wird erforscht, welche Variablen wiederum die inklusiven Lehrereinstellungen determinieren. Sowohl demografische Merkmale, z.B. das Alter, als auch berufsbezogene Merkmale der Lehrkräfte, wie z.B. die bisherigen Unterrichtserfahrungen, die Erfahrungen in inklusiven Settings, Selbstwirksamkeitserwartungen und die Bereitschaft an Fortbildungsmaßnahmen teilzunehmen, spielen dabei eine bedeutsame Rolle (de Boer, 2011; Forlin & Chambers, 2011; Kim, 2010; Savolainen, Engelbrecht, Nel & Malinen, 2012). Hingegen wurde bislang kaum untersucht, inwiefern schulische Kontextbedingungen die Einstellungen der Lehrkräfte zur Inklusion beeinflussen. So hat z.B. das Kontextmerkmal Größe der Schule, also die Anzahl der Lernenden und Lehrenden, einen Einfluss auf Schülermerkmale. Newman, Garrett, Elbourne, Bradley, Noden, Taylor und West (2006) fanden in ihrer Meta-Analyse heraus, dass die untersuchten Studien in Bezug auf den Zusammenhang von Schulgröße und den Leistungen der Schülerschaft ein inkonsistentes Bild erzeugen. Im Hinblick auf die das Schulklima jedoch gilt, dass Schülerinnen und Schüler und auch die Lehrkräfte an kleineren Schulen das Klima an ihren Schulen positiver einschätzen als Schülerinnen und Schüler sowie Lehrkräfte an größeren Schulen (Newman et al., 2006). Bislang bleibt jedoch offen, ob ein bedeutsamer Zusammenhang zwischen diesem Kontextmerkmal und den Lehrereinstellungen besteht. Ein weiteres bedeutsames Schulkontextmerkmal ist die Organisation der Schuleingangsphase. Derzeit ist kaum erforscht, wie sich die flexible Schuleingangsphase (im Sinne eines jahrgangsübergreifenden Lernens mit höherer Heterogenität auf Schülerebene) auf die Lehrkräfte auswirkt. Lang, Grittner, Rehle und Hartinger (2010) konnten bislang zeigen, dass Lehrkräfte, die aktuell jahrgangsgemischt unterrichten, die Heterogenität als Bereicherung wahrnehmen und seltener eine Homogenisierung der Schülerleistungen anstreben. Weitere Befunde stehen allerdings noch aus. Ein drittes Kontextmerkmal bezieht sich auf eine Organisationsform von Schule, die Ganztagschule. Aktuelle Untersuchungen befassen sich damit, welche Bedingungsgefüge zwischen Schüler- und Lehrermerkmalen innerhalb von Ganztagschulen bestehen (Willems, Jarsinski, Holtappels & Rollett, 2013). Offen bleibt der Vergleich von Schulen mit und ohne Ganztagskonzept und der jeweiligen Auswirkungen auf die Lehrermerkmale. Die genannten Desiderata stehen im Mittelpunkt der vorliegenden Studie, mit der untersucht werden soll, inwiefern Zusammenhänge zwischen den Einstellungen von Lehrkräften an inklusiven Grundschulen und (a) individuellen und (b) schulstrukturellen Merkmalen bestehen. Konkret rücken folgende Hypothesen in den Blick:

- (1) Jüngere Lehrerinnen und Lehrerinnen haben höhere Einstellungswerte als ältere Kollegen.
  - (2) Sowohl die persönlichen Erfahrungen mit Menschen mit Behinderungen als auch das Ausmaß bisheriger inklusiver bzw. integrativer Unterrichtserfahrungen haben positive Auswirkungen auf die Einstellungen zur Inklusion.
  - (3) Lehrkräfte, die an Schulen mit flexibler Schuleingangsphase im Sinne des jahrgangsübergreifenden Lernens beschäftigt sind, haben höhere Einstellungswerte.
  - (4) Arbeiten Lehrkräfte an Schulen mit einem Ganztagskonzept, so sind ihre Einstellungen zur Inklusion positiver.
- Schließlich soll explorativ untersucht werden, welchen Einfluss die Größe einer Schule auf die inklusiven Einstellungen von Lehrerinnen und Lehrer hat. Die Untersuchung ist eingebettet in die wissenschaftliche Begleitforschung des Brandenburger Pilotprojektes „Inklusive Grundschule“ durch die Universität Potsdam. Es liegen strukturelle Daten zur Schülerzahl, zum Schulkonzept und weiteren Rahmenbedingungen von N = 75 inklusiv unterrichtenden Schulen vor. Zudem wurden N = 630 Lehrerinnen und Lehrer, die in 2. und 3. Klassen unterrichten, querschnittlich per Fragebogen zu ihren Einstellungen zum inklusiven Unterricht (Einstellungen zum gemeinsamen Unterricht,  $\alpha = .82$ ; Einstellungen zur Wirksamkeit inklusiven Unterrichts,  $\alpha = .83$ ; Einstellungen zum Verhalten von Kindern im inklusiven Unterricht,  $\alpha = .76$ ) sowie zu demografischen Merkmalen befragt. Die auf Individual- und Schulebene gewonnenen Daten werden mehrerebenen- und regressionsanalytisch im Hinblick auf die genannten Einstellungsfacetten ausgewertet. Erste Analysen bestätigen Unterschiede in den Einstellungen der Lehrkräfte hinsichtlich des Alters.

- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469–520.
- Boer, A. de, Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: a review of the literature. International Journal of Inclusive Education, 15(3), 331–353.
- Forlin, C., & Chambers, D. (2011). Teacher preparation for inclusive education: increasing knowledge but raising concerns. Asia-Pacific Journal of Teacher Education, 39(1), 17–32.
- Hartinger, A., Kleickmann, T., & Hawelka, B. (2006). Der Einfluss von Lehrervorstellungen zum Lernen und Lehren auf die Gestaltung des Unterrichts und auf motivationale Schülervariablen. Zeitschrift für Erziehungswissenschaft, 9(1), 110–126.
- Kim, J. (2011). Influence of teacher preparation programmes on preservice teachers' attitudes toward inclusion. International Journal of Inclusive Education, 15(3), 355–377.
- Küsting, J., Billich, M., & Lipowsky, F. (2009). Der Einfluss von Lehrerkompetenzen und Lehrerhandeln auf den Schülerfolg von Lernenden. In O. Zlatkin-Troitschanskaia (Ed.), Beltz-Bibliothek. Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung (pp. 655–667). Weinheim, Basel: Beltz.
- Lang, E., Grittner, F., Rehle, C., & Hartinger, A. (2010). Das Heterogenitätsverständnis von Lehrkräften im jahrgangsgemischten Unterricht der Grundschule. In J. Hagedorn (Ed.), Ethnizität, Geschlecht, Familie und Schule. Heterogenität als erziehungswissenschaftliche Herausforderung ; [Festschrift für Leonie Herwartz-Emden] (1st ed., pp. 315–331). Wiesbaden: VS, Verl. für Sozialwiss.
- Newman, M., Garrett, D., Elbourne, D., Bradley, S., Noden, P., Taylor, J., & West, A. (2006). Does secondary school size make a difference? Educational Research Review, 1(1), 41–60.
- Savolainen, H., Engelbrecht, P., Nel, M., & Malinen, O.-P. (2012). Understanding teachers' attitudes and self-efficacy in inclusive education: implications for pre-service and in-service teacher education. European Journal of Special Needs Education, 27(1), 51–68.
- Sermier Dessemontet, R., Benoit, V., & Bless, G. (2011). Schulische Integration von Kindern mit einer geistigen Behinderung: Untersuchung der Entwicklung der Schulleistungen und der adaptiven Fähigkeiten, der Wirkung auf die Lernentwicklung der Mitschüler sowie der Lehrereinstellungen zur Integration. Empirische Sonderpädagogik, (4), 291–307.
- Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2011). Überzeugungen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV (pp. 235–257). Münster, New York, NY, München, Berlin: Waxmann.
- Willems, A. S., Jarsinski, S., Holtappels, H. G., & Rollett, W. (2013). Schulische Qualitätsmerkmale von Ganztagsgrundschulen aus Sicht der Lehrkräfte: Zur Bedeutung der sozialen Komposition der Schülerschaft. Zeitschrift für Grundschulforschung, 6(2), 69–84.



## **Professionelle Kompetenz von Grundschullehrkräften: Bedeutung für die motivational-emotionale Entwicklung der Lernenden und Zusammenhänge zum Schulkontext**

Theoretischer Hintergrund: Die Lehrkraft gilt als zentraler Akteur im Bildungssystem, weshalb die systematische Untersuchung der Lehrerkompetenz zunehmend in den Fokus der Bildungsforschung rückt. Professionelle Lehrerkompetenz wird dabei als Zusammenspiel von kognitiven (professionelles Wissen und Überzeugungen) und motivationalen Aspekten aufgefasst (Kunter, Baumert, Blum, & Neubrand, 2011). Zusammenhänge zwischen Aspekten der professionellen Lehrerkompetenz und Unterrichtsqualität sowie Lernerfolg konnten bereits gezeigt werden (z.B. Dubberke, Kunter, McElvany, Brunner, & Baumert 2008; Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009; Kunter et al., 2013; Voss, Seiz, Hoehne, Kunter, & Baumert, in press). Außerdem weisen Ergebnisse darauf hin, dass motivationale Aspekte der Lehrerkompetenz mit dem Schulkontext (z.B. Kooperation im Kollegium, Führungsstil des Schulleiters/der Schulleiterin) assoziiert sind (z.B. Klusmann, Kunter, Trautwein, Lüdtke, & Baumert, 2008). Ein Großteil der bisherigen Forschung bezieht sich dabei auf den Sekundarschulbereich.

Zielsetzung: Ziel dieser Arbeit ist es, die professionelle Kompetenz von Grundschullehrkräften zu untersuchen. Dabei wird Lehrerkompetenz nicht nur als unabhängige Variable in ihrer Wirkung auf die motivational-emotionale Entwicklung der SchülerInnen untersucht, sondern auch als abhängige Variable betrachtet, wobei Zusammenhänge mit dem Schulkontext (berichtet aus Sicht der SchulleiterInnen) analysiert werden. Die Grundschule stellt hierbei einen besonders interessanten Bereich dar, da der/die KlassenlehrerIn eine besonders wichtige Rolle für die Entwicklung der SchülerInnen spielen sollte.

Methode: Den Analysen liegt ein Längsschnittdesign mit zwei Messzeitpunkten und drei unterschiedlichen Datenquellen (SchülerInnen, LehrerInnen und SchulleiterInnen) zugrunde. Die Befragungen fanden im Rahmen einer Studie der wissenschaftlichen Begleitung der Hector-Kinderakademien statt, einem außerunterrichtlichen Förderprogramm für besonders begabte und hoch begabte Grundschulkinder. Die Variablen auf Schülerseite wurden zweimal erfasst (zu Beginn des dritten Schuljahres und ein halbes Jahr später), die Variablen auf Lehrkraft- und SchulleiterInnen-Seite einmal. Die Stichprobe bestand aus 111 Grundschulen und deren SchulleiterInnen mit 189 dritten Klassen und deren über 2000 SchülerInnen.

Eingesetzte Skalen: Schülerskalen ( $\alpha$  jeweils zu t1/t2): - Schulentzug ( $\alpha = .89/.91$ ; Entwicklung im Rahmen der Studie); - Akademische Interessen ( $\alpha = .82/.82$ ; in Anlehnung an bestehende Interessensskalen); - Emotionales Wohlbefinden ( $\alpha = .63/.61$ ; KINDL; Ravens-Sieberer & Bullinger, 1998)

Lehrerskalen (übernommen aus COACTIV, vgl. Kunter et al., 2011): - Enthusiasmus für das Unterrichten ( $\alpha = .87$ ); - Emotionale Erschöpfung ( $\alpha = .71$ ); - Rollenverständnis als Erzieher ( $\alpha = .80$ )

Schulleiterskalen: - Führungsstil (5 Unterskalen;  $.49 < \alpha < .67$ ; Bolman & Deal, 1997); - Kooperation im Kollegium ( $\alpha = .87$ ; Kunter et al., 2002)  
- Beeinträchtigung des Schulklimas von Lehrer-/Schülerseite ( $\alpha = .81/.79$ ; OECD, 2005)

Die Daten werden mittels Mehrebenenstrukturgleichungsmodellen in Mplus analysiert.

Erste Ergebnisse: Da die Daten sich noch in einem frühen Stadium der Aufbereitung befinden, konnten bisher nur vorläufige Analysen mittels hierarchischer Regression auf manifester Ebene berechnet werden. Die motivationalen und emotionalen Variablen auf Schülerseite weisen eine mittlere Stabilität über das halbe Jahr auf (erklärte Varianz zu t2 durch die jeweilige Variable zu t1 bei 34/25/20% für Engagement/Interesse/emotionales Wohlbefinden). Die Voralysen deuten darauf hin, dass sowohl Enthusiasmus für das Unterrichten als auch emotionale Erschöpfung auf Lehrerseite einen Zusammenhang mit den emotionalen und motivationalen Variablen auf Schülerseite zu t2 zeigen (kontrolliert für die jeweiligen Ausprägungen zu t1), hingegen das Rollenverständnis der Lehrkräfte keinen Einfluss auf die Schülervariablen hat. Erste korrelative Voralysen zur Untersuchung der Zusammenhänge zwischen Lehrervariablen und Schulkontext deuten darauf hin, dass vor allem die emotionale Erschöpfung mit den durch die SchulleiterInnen berichteten Variablen zum Schulkontext assoziiert ist (positiv mit Beeinträchtigung des Schulklimas von Lehrer-/Schülerseite und negativ mit Kooperation im Kollegium). Auch einzelne Unterskalen des Führungsstils hängen mit der Lehrerkompetenz zusammen. Weitere differenzierte Analysen werden vorgestellt. Diese Studie soll einen wichtigen Beitrag zum Verständnis des Zusammenspiels von Lehrerkompetenz, Schulkontext und sozial-emotionaler Entwicklung der SchülerInnen leisten.

Bolman, L. G., & Deal, T. E. (1997). Reframing organizations: Artistry, choice, and leadership (2nd ed.). The Jossey-Bass higher and adult education series. San Francisco: Jossey-Bass.

Dubberke, T., Kunter, M., McElvany, N., Brunner, M., & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. Zeitschrift für Pädagogische Psychologie, 22(3), 193–206.

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. Journal of Educational Psychology, 101(3), 705–716.

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Engagement and emotional exhaustion in teachers: Does the school context make a difference? Applied Psychology: An International Review, 57(1), 127–151.

Kunter, M., Baumert, J., Blum, W., & Neubrand, M. (Eds.). (2011). Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. Journal of Educational Psychology, 105(3), 805–820.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., ...Weiß, M. (2002). PISA 2000: Dokumentation der Erhebungsinstrumente. Berlin: Max-Planck-Institut für Bildungsforschung.

OECD. (2005). PISA 2003 Technical report. Paris: OECD.

Ravens-Sieberer, U., & Bullinger, M. (1998). Assessing health related quality of life in chronically ill children with the German KINDL: first psychometric and content-analytical results. Quality of Life Research, 4(7), 399-407.

Voss, T., Seiz, J., Hoehne, V., Kunter, M., & Baumert, J. (in press). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. Zeitschrift für Pädagogik.





## **Optimierung eines deutschsprachigen Fragebogens zur Erfassung epistemischer Überzeugungen von Lehramtsstudierenden**

Epistemische Überzeugungen bezeichnen individuelle Überzeugungen über die Struktur des Wissens, den Wissenserwerb und die Wissenserzeugung (Schommer-Aikins, 2004; Urhahne & Hopf, 2004). Nach dem Modell von Hofer (2001) haben epistemische Überzeugungen von Lehrpersonen Einfluss auf Motivation, Lernstrategien, implizite Theorien über Lernen und Unterricht und epistemische Überzeugungen von Schülerinnen und Schülern. Schommer entwickelte eine Theorie, die epistemische Überzeugungen in fünf mehr oder weniger unabhängige Dimensionen fasst (Schommer, 1990; Schommer-Aikins, 2002). In Weiterentwicklung dieser Grundlage gehen Conley, Pintrich, Vekiri und Harrison (2004) von vier Dimensionen aus (source (Quelle), certainty (Sicherheit), development (Entwicklung) und justification (Rechtfertigung/Begründung). Die einzelnen Dimensionen zeichnen sich durch zwei dichotome, gegenüberstehende Ausprägungen von weniger und weiter entwickelten epistemischen Überzeugungen aus. Die Dimension Sicherheit des Wissens repräsentiert ein Kontinuum zwischen der weniger entwickelten Überzeugung, dass es nur eine richtige und wahre Antwort auf wissenschaftliche Fragen gibt und der Überzeugung, dass es mehrere Antworten auf komplexe Fragen geben kann. Die Dimension Entwicklung des Wissens erfasst Überzeugungen über die Veränderlichkeit/Festigkeit von Wissen. Während die Dimension Quelle des Wissens thematisiert, ob Wissen durch allwissende Autoritäten weitergegeben oder subjektiv aufgenommen wird, wird durch die Dimension Rechtfertigung des Wissens die Rolle von Studien und Experimenten und den daraus folgenden Argumenten für die Begründung von Wissen erfasst. Conley et al. (2004) entwickelten einen Fragebogen zur Erfassung epistemischer Überzeugungen von Schülerinnen und Schülern in den Naturwissenschaften. Dieses Instrument wurde von Urhahne und Hopf (2004) ins Deutsche übersetzt und in einer Studie mit 167 Schülerinnen und Schülern getestet. Die Ergebnisse dieser Untersuchung zeigten niedrigere Validitäts- und Reliabilitätswerte als die englischsprachige Originalstudie. Urhahne und Hopf (2004) vermuten, dass „auch Übersetzungsschwierigkeiten zu einer geringer ausfallenden Reliabilität beigetragen haben“ (Urhahne & Hopf, 2004, S. 81) könnten. Ziel der vorliegenden Untersuchung war es, die englischen Items des Fragebogens von Conley et al. (2004) sprachlich zu überarbeiten, sodass sie (a) dem deutschen Sprachgebrauch entsprechen aber dennoch (b) den originalen Inhalt wiedergeben. 84 Studierende (52 Frauen, 32 Männer; M= 23.14 Jahre; SD = 3.67; 5.-7. Semester;) bearbeiteten den sprachlich-modifizierten Fragebogen und 91 Personen (28 Männer, 62 Frauen; M = 22.29 Jahre; SD = 1,89; 5. Semester) den Fragebogen in der Originalübersetzung von Urhahne und Hopf (2004). Die Varianzaufklärung der einzelnen Skalen konnte mit dem sprachlich-modifizierten Fragebogen erhöht werden; insgesamt zeigte sich eine Steigerung der Gesamtvarianzaufklärung von 39.54% auf 62.51%. Zudem erzielte der sprachlich-modifizierte Fragebogen höhere Reliabilitätswerte (Quelle des Wissens  $\alpha = .87$ ; Sicherheit des Wissens  $\alpha = .77$ ; Entwicklung des Wissens  $\alpha = .85$ ; Rechtfertigung des Wissens  $\alpha = .76$ ) als der Fragebogen in der Übersetzung von Urhahne und Hopf (2004) (Quelle des Wissens  $\alpha = .67$ ; Sicherheit des Wissens  $\alpha = .41$ ; Entwicklung des Wissens  $\alpha = .66$ ; Rechtfertigung des Wissens  $\alpha = .68$ ). In einer zweiten Studie (N= 677) wurde die Reliabilität und Validität des sprachlich-modifizierten Fragebogens erneut getestet und zudem untersucht, inwiefern dieses Instrument zur Domänen übergreifenden Erfassung von epistemischen Überzeugungen geeignet ist oder ob sich Unterschiede in Abhängigkeit des Studienfachbereiches der Studierenden zeigen. Die Reliabilitätswerte für die Ergebnisse der großen Stichprobe des sprachlich-modifizierten Fragebogens waren niedriger als in der ersten Studie (Quelle des Wissens  $\alpha = .76$ ; Sicherheit des Wissens  $\alpha = .69$ ; Entwicklung des Wissens  $\alpha = .77$ ; Rechtfertigung des Wissens  $\alpha = .60$ ). Bezogen auf die Gesamtvarianzaufklärung ergab sich ein Wert von 46.12%. Die Varianzaufklärung der einzelnen Skalen war für alle Dimensionen – mit Ausnahme der Entwicklung des Wissens – niedriger als in der ersten Studie, jedoch immer noch höher als in der Studie von Urhahne und Hopf (2004). Der Vergleich Studierender naturwissenschaftlicher, geisteswissenschaftlicher oder gemischter Fächerkombinationen zeigte geringe Unterschiede in der Varianzaufklärung (48.09% für Studierende zweier geisteswissenschaftlicher Fächer; 46.04% für Studierende naturwissenschaftlicher Fächer; 42.61% gemischter Fächerkombinationen).

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186–204.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review* (Vol. 13, pp. 353–383).

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498–504.

Schommer-Aikins, M. (2002). An Evolving Theoretical Framework for an Epistemological Belief System. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing* (pp. 103–118). New York, London: Taylor and Francis.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist*, (July 2013), 19–29.

Urhahne, D., & Hopf, M. (2004). Epistemologische Überzeugungen in den Naturwissenschaften und ihre Zusammenhänge mit Motivation, Selbstkonzept und Lernstrategien. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 71–87.





## **Das Wissen von (angehenden) Lehrkräften zur Diagnose von Lernstrategien: Misskonzepte und „Knowledge in pieces“**

Theoretischer Hintergrund. Die Standards für die Lehrerbildung (KMK, 2004, S. 9, 11) führen das Diagnostizieren und Fördern von Lernstrategien als zentrale Kompetenzen von Lehrkräften an. Jedoch betrachtet die Forschung zu diagnostischen Kompetenzen bislang kaum das Wissen und Können von Lehrkräften bezüglich Lernstrategiediagnose (Brunner, Anders, Hachfeld, & Krauss, 2011; vgl. Helmke, 2003). Clift, Ghatala, Naus und Poole (1990) fanden, vor mehr als 20 Jahren, dass viele Lehrkräfte erstens Lernstrategien nicht im wissenschaftlichen Sinne definierten (55%). Die wissenschaftliche Definition sollte hier mindestens umfassen, dass eine Lernstrategie eine vom Lerner initiierte, mentale Aktivität ist, um Inhalte zu lernen und zu verstehen. Die Probanden von Clift et al. erwähnten lehrgesteuerte Lernaktivitäten und trennten nicht klar zwischen Lernstrategien und basalen Informationsverarbeitungsprozessen. Zweitens priorisierten die Lehrkräfte Lernstrategien für bestimmte Lernaufgaben nicht erwartungskonform: Sie wählten selten das Organisieren und Elaborieren, obwohl diese beiden Strategien für verständnisorientiertes Lernen besonders wichtig sind. Deshalb könnte man hier von einer schlechten Priorisierung von Wissen sprechen. Es wurden drittens Verwechslungen von einzelnen Substrategien gefunden. Dies lässt vermuten, dass das Vorwissen der Lehrkräfte bezüglich Lernstrategien aus einer Mischung von (1) misskonzeptartigen Wissensbausteinen (ähnlich den Fehl kategorisierungen nach Chi, 1992) und (2a) zwar richtigen, aber nicht kohärent organisierten und (2b) fehlpriorisierten Wissensbausteinen (ähnlich „knowledge in pieces“, DiSessa, 1988; Kali, Goodyear, & Markauskaite, 2011) besteht.

Fragestellung. Auf diesem Hintergrund wurde in der vorliegenden Studie das Vorwissen von angehenden Lehrkräften analysiert. Folgende Forschungsfragen wurden bearbeitet: (1) Wie häufig zeigen (angehende) Lehrkräfte Misskonzepte (Fehl kategorisierungen, vgl. Chi, 1992) in Bezug auf verständnisorientierte Lernstrategien und deren Diagnose? (2) Welche Hinweise auf nicht priorisierte und nicht kohärent organisierte und Wissensbausteine zeigen sie in Bezug auf Lernstrategien und deren Diagnose? (3) Lassen sich Untergruppen mit spezifischen Defiziten bei verschiedenen (angehenden) Lehrkräften finden?

Methode. Offene Antworten in Vorwissenstests aus Studien mit Lehramtsstudierenden wurden reanalysiert (N = 129; M [Alter] = 22.7, SD = 3.32; 92 weiblich und 37 männlich; Glogger et al., 2010, 2013; Ohst et al., 2012). Die Tests fragten zu Beginn jeweils nach wichtigen, besonders lernförderlichen Lernstrategien allgemein. Eine kontextualisierte Frage bezog sich darauf, welche Lernstrategien beim Schreiben von Lerntagebüchern von Schülern verwendet werden können. Die offenen Antworten wurden mit einem Kodiersystem ausgewertet, das Misskonzepte (z.B. Lehrstrategien), primäre verständnisorientierte (z.B. Elaboration) und nicht-primäre oder Stützstrategien (z.B. in einer Lerngruppe lernen; angenehme Lernatmosphäre schaffen) erfasste. Nicht kohärent organisiertes Wissen wurde als Kohärenzrating der Antworten auf die verschiedenen Fragen erfasst (Skala von 1 bis 6, 1 = sehr kohärentes Wissen; vgl. Braasch, Goldman, & Wiley, 2012). Wenn vorwiegend nicht-primäre Lernstrategien genannt wurden, wurde von fehlpriorisiertem Wissen ausgegangen. Zur Beantwortung der Frage 3 wurden Clusteranalysen durchgeführt.

Ergebnisse. 46% der (zukünftigen) Lehrkräfte nannten mindestens ein „Misskonzept“ (z.B. Lehrstrategie, genannt von 26.4% der Lehrkräfte). 45% nannten Wiederholung. Im Gegensatz zu Elaboration und Organisation ist Wiederholung nicht als wichtige verständnisorientierte Lernstrategie anzusehen. 73 % der Lehrkräfte nannten mehr nicht-primäre als primäre Lernstrategien. Die letzten beiden Ergebnisse weisen auf fehlpriorisiertes Wissen hin. Die Kohärenz des Wissens war eher schwach (M = 4.09, SD = 1.58). Bei der Häufigkeitsverteilung des Kohärenz-Ratings fiel auf, dass Stufe 6 (sehr inkohärent) am häufigsten besetzt war. Die geringe globale Kohärenz bedeutet, dass auf jede Frage hin anderes Lernstrategiewissen abgerufen und kaum Verknüpfungen hergestellt wurden. Die Clusteranalyse ergab vier Cluster mit den folgenden Vorwissensarten: (1) insgesamt gut, (2) primäre Lernstrategien bekannt, aber eher fragmentiertes Wissen, (3) vorwiegend fehlpriorisiert, (4) misskonzeptartig und inkohärent. Bezüglich der Angehörigkeit einer Hochschule fällt auf, dass sich im Cluster mit dem besten Vorwissen keine Probanden der pädagogischen Hochschule befinden (Likelihood Quotient  $p = .011$ ). Bei explorativen Analysen zeigten sich Hinweise, dass Lehrkräfte des Clusters „misskonzepthaft“ anders auf die experimentellen Interventionen der herangezogenen Studien reagierten als andre Cluster.

Braasch, J. L., Goldman, S. R., & Wiley, J. (2012). The role of text and reader characteristics in science learning from refutational texts. Paper presented at the Annual Conference of the American Educational Research Association (AERA). Vancouver, Canada.

Brunner, M., Anders, Y., Hachfeld, A. & Krauss, S. (2011). Diagnostische Fähigkeiten von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften (S. 215-234). Münster: Waxmann.

Chi, M. T. H. (1992). Conceptual change in and across ontological categories: Examples from learning and discovery in science. In R. Giere (Hrsg.), Cognitive models of science (S. 129-160). Minneapolis, MN: University of Minnesota Press.

Clift, R. T., Ghatala, E. S., Naus, M. M. & Poole, J. (1990). Exploring teachers' knowledge of strategic study activity. Journal of Experimental Education, 58, 253-263.

DiSessa, A. (1988). Knowledge in pieces. In G. Forman & P. Fullall (Hrsg.), Constructivism in the computer age (S. 49-70). Hillsdale, NJ: Erlbaum.

Glogger, I., Kappich, J., Schwonke, R., Holzaepfel, L., Nueckles, M., & Renkl, A. (2010). Inventing prepares learning not only cognitively, but also motivationally. In M. Hopp, & F. Wagner (Eds.), Instructional design for motivated and competent learning in a digital world - Proceedings of the EARLI SIG 6 and 7 Conference (S. 41-43). Ulm, Germany: University of Ulm.

Glogger, I., Kappich, J., Schwonke, R., Holzaepfel, L., Nueckles, M., & Renkl, A. (2013b). Inventing prepares learning motivationally, but a worked-out solution enhances learning outcomes. In: Knauff M., Pauen, M., Sebanz N., & Wachsmuth I. (Eds.), Proceedings of the 35th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society, 2013; 2416-2421

Helmke, A. (2003). Unterrichtsqualität erfassen, bewerten, verbessern. Seelze-Velber: Kallmeyer.

Kali, Y., Goodyear, P., & Markauskaite, L. (2011). Researching design practices and design cognition: Contexts, experiences and pedagogical knowledge-in-pieces. Learning, Media and Technology, 36(2), 129-149.

doi:10.1080/17439884.2011.553621

Ohst, A., Glogger, I., Nueckles, M., & Renkl, A. (2013). An approach to overcome the barrier of teachers' prior pedagogical knowledge in pieces. Paper presented at the 15th Biennial Conference for Research on Learning and Instruction (EARLI), Germany, Munich.



## Quereinsteiger in berufsbegleitenden Pflegestudiengängen

Im Kontext des demografischen Wandels gewinnen im Pflegebereich die Fachkräftesicherung und das Konzept des lebenslangen Lernens immer mehr an Bedeutung. Zur Vermeidung von gravierenden Personalengpässen gilt es daher, neue Zielgruppen, wie z.B. Quereinsteiger, zu gewinnen und zu qualifizieren. Als Quereinsteiger in der Pflege gelten Personen, die zunächst in einem fachfremden Bereich tätig waren oder fachfremd studiert und anschließend eine Ausbildung als Altenpfleger/in oder Gesundheits- und (Kinder-)Krankenpfleger/in absolviert haben. Im Zuge der Akademisierung der Pflege steigt in den letzten Jahren auch bei Quereinsteigern unter Pflegefachkräften das Interesse, ein (berufsbegleitendes) Pflegestudium aufzunehmen.

Die ersten Pflegestudiengänge an deutschen Fachhochschulen entstanden in den 1990er Jahren, in den vergangenen Jahren ist ihre Anzahl angestiegen. Studieninteressierte Pflegefachkräfte, die zwar über Berufspraxis, aber nicht über eine schulische Hochschulzugangsberechtigung verfügen, profitieren dabei von einem Beschluss der Kultusministerkonferenz aus dem Jahr 2009, der beruflich Qualifizierten unter bestimmten Voraussetzungen die Aufnahme eines Studiums ermöglicht.

Dennoch erweist sich die nur begrenzte Durchlässigkeit zwischen beruflicher und hochschulischer Bildung zunehmend als Hindernis, das die Fachkräftesicherung zum Beispiel in Pflegeberufen erschwert. Deshalb initiierte das Bundesministerium für Bildung und Forschung (BMBF, 2011) unter dem Titel „Aufstieg durch Bildung: offene Hochschulen“ einen bundesweiten Wettbewerb, um innovative, nachfrageorientierte und nachhaltige Hochschulkonzepte zu fördern, die den Übergang zwischen beruflicher und akademischer Bildung verbessern. Zu den 26 Förderprojekten zählt auch das Verbundvorhaben „BEST WSG – Berufsintegrierte Studiengänge zur Weiterqualifizierung im Sozial- und Gesundheitswesen“ der Fachhochschule der Diakonie (FHdD), Bielefeld, und der Hochschule der Bundesagentur für Arbeit (HdBA), Mannheim. Während die FHdD innovative Konzepte für berufsbegleitende Studiengänge im Sozial- und Gesundheitswesen entwickelt, übernimmt die HdBA die wissenschaftliche Begleitforschung zur Analyse der personen- und strukturbezogenen Rahmenbedingungen.

Im Rahmen des Forschungsprojektes fand im Sommer 2013 erstmals eine bundesweite Online-Befragung von Studierenden und Absolventen berufsbegleitender Pflegestudiengänge statt. Die Teilnehmerakquise erfolgte über die jeweiligen Leiter bzw. Koordinatoren der Studiengänge an insgesamt 25 Hochschulen. In der quantitativen Umfrage wurden berufliche Vorerfahrungen, Beweggründe für einen etwaigen Quereinstieg auf Fachkräfteebene, Erfahrungen während des Studiums und – bei Absolventen – der berufliche Werdegang nach dem Hochschulabschluss erhoben. Insgesamt haben 445 Personen teilgenommen, von ihnen haben 380 den Fragebogen vollständig ausgefüllt. Unter diesen befinden sich 107 Quereinsteigern – darunter 18 ohne schulische Hochschulzugangsberechtigung – und 273 Pflegefachkräfte ohne andere Berufs- und Studierenerfahrungen (sogenannter „linearer Typ“). Die Einbeziehung von Vertretern des „linearen Typs“ (LT) ermöglicht eine Datenauswertung in vergleichender Perspektive. So treten spezifische Merkmale von Quereinsteigern (QE) deutlicher hervor. Anhand dieser Ergebnisse konnten sowohl Förderfaktoren als auch strukturelle Hürden für die Aufnahme und das Absolvieren eines berufsbegleitenden Studiums identifiziert werden. Als förderlich werden vor allem die Freistellung durch den Arbeitgeber (QE: 89,7%, LT: 85,7), eine finanzielle Absicherung – auch bei Arbeitszeitreduzierung – (QE: 75,7%, LT: 76,9) und die Unterstützung durch den/die Partner/in (QE: 71%, LT: 61,2%) angesehen. Die größten Barrieren stellen demgegenüber die verhältnismäßig geringe Bezahlung (QE: 79,4%, LT: 82,4%) und die fehlenden adäquaten Stellen für akademisch qualifizierte Fachkräfte (QE: 74,8%, LT: 69,6%) sowie die fehlende Unterstützung durch den Arbeitgeber (QE: 71%, LT: 67,8%) dar. Bei 48,6 % der Quereinsteigern und bei 50,5% der Vertreter des „linearen Typs“ fördert/e der Arbeitgeber das Studium weder ideell noch finanziell. Hinzu kommt, dass bei 65,4% der Quereinsteiger und bei 66,3% der Vertreter des „linearen Typs“ keine Anrechnung der beruflichen Kompetenzen auf die im Curriculum geforderten Studienleistungen erfolgte. Die Befragung lieferte auch Orientierungspunkte für die Konzeption der an der FHdD entstehenden berufsintegrierten Studiengänge im Sozial- und Pflegebereich. So werden insbesondere Vorkurse (QE: 38,3%, LT: 42,5%) zum wissenschaftlichen Arbeiten, eine bessere Betreuung durch Mitarbeiter an der Hochschule (QE: 30,8%, LT 21,2%) und zusätzliche Lehrveranstaltungen für Quereinsteiger (QE: 19,6, LT: 18,3%) gewünscht. Die spezifische Perspektive der Quereinsteiger verbindet Erkenntnisse zu strukturellen Aspekten der berufsbegleitenden akademischen Weiterbildung mit individuellen Einblicken in vielfältige Berufsbiografien.

Ayan, Türkan (Hrsg.) (2013): Einsteigen, Umsteigen, Aufsteigen. Personenbezogene und strukturelle Rahmenbedingungen für Berufe und Bildungschancen im Sozial- und Gesundheitssektor. Köln: Kölner Wissenschaftsverlag.

Bettig, Uwe/Frommelt, Mona/Schmidt, Roland (Hrsg.) (2012): Fachkräftemangel in der Pflege. Konzepte, Strategien, Lösungen. Heidelberg: medhochzwei.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

Raum 823



## Handlungsorientierte Erfassung emotionsbezogener Pflegekompetenzen

Theoretischer Hintergrund Die Messung beruflicher Handlungskompetenz steht im Bereich der Gesundheitsdienstleistungen allgemein, und der Altenpflege im Speziellen, vor besonderen Herausforderungen: Neben dem hohen fachlichen Standard sind Pflegeberufe stärker als andere Berufe durch emotionale, selbstreflexive und beziehungsbezogene Anforderungen geprägt. Der Stellenwert emotionsbezogener Anforderungen für die Konstruktion pflegespezifischer Kompetenzmodelle und die Feststellung beruflicher Handlungskompetenzen wird anhaltend kontrovers diskutiert (Dingwall & Allen, 2001; Neumann-Poensch & Höller, 2011). Sowohl berufspädagogische als auch pflegewissenschaftliche Argumentationen greifen dabei auf die Konzepte der Emotionsarbeit (Hochschild, 1990) und Empathiefähigkeit (Rogers, 1961) zurück, erschöpfen sich häufig jedoch in Tätigkeitsbeschreibungen oder Einstellungsmessungen; ihr Bezug zum kompetenten Pflegehandeln in verschiedenen Arbeitsfeldern bleibt somit weitgehend unklar (Böhle & Glaser, 2006; Yu & Kirk, 2009). Obwohl ein mehrdimensionales und sequenzielles Verständnis von Empathie und Emotionsarbeit vorherrscht, wurden die im Pflegekontext beschriebenen Facetten (z.B. empathische Akkuratheit, Perspektivenübernahme, Kommunikation empathischen Verständnisses; Erkennen institutioneller und individueller Empfindungs- und Ausdrucksregeln, Kontrolle von erlebten Gefühlen und des emotionalen Ausdrucksverhaltens in der Kommunikation mit gepflegten und ihren Angehörigen) bislang noch nicht empirisch im Zusammenhang mit dem Pflegeprozess selbst betrachtet (Bischoff-Wanner, 2002).

Fragestellung Diese Arbeit prüft die Angemessenheit konkurrierender Modelle zur Struktur emotionsbezogener Facetten pflegerischer Handlungskompetenz. Dabei kann erstmalig auf Berufschülerleistungen aus einem standardisierten simulationsorientierten Testverfahren zurückgegriffen werden.

Methode Die verwendeten Daten stammen aus der Pilotierung eines computergestützten Tests zur Abbildung eines Teilbereiches des im Projektverbund erarbeiteten Kompetenzmodelles für die Pflege. Die anfallende Stichprobe umfasst insgesamt 100 Absolventen bzw. Schülerinnen und Schüler im dritten Ausbildungsjahr aus Altenpflegeschulen in Bayern und Nordrhein-Westfalen. Der Gesamttest umfasst 66 Aufgaben zur Abbildung der drei theoretisch postulierten Dimensionen diagnostisch-reflexiver, praktisch-technischer und interaktiv-kommunikativer Kompetenz in der unmittelbar bewohnerbezogenen Pflege. Insgesamt 17 Aufgaben wurden mit systematischem Bezug zu emotionsbezogenen Anforderungen in verschiedenen Einsatzfeldern (stationäre Pflege demenzkranker und sterbender alter Menschen, ambulante Pflege chronisch kranker alter Menschen) und Handlungssituationen (Problemfällen) der Pflege entwickelt. Auf der Grundlage der Modellgüte alternativer Item-Response-Modelle wird geprüft, inwieweit sich verschiedene Facetten der zugrunde gelegten Konzepte der Empathie und Emotionsarbeit entweder (1.) wie vorgesehen als Indikatoren einer handlungsorientierten Kompetenzstruktur begreifen lassen, (2.) strukturell die für die Aufgabengenerierung genutzten originären emotionsbezogenen Fähigkeitskonstrukte repräsentieren, oder aber (3.) weder inhaltlich noch handlungsbezogen klar voneinander abgegrenzt werden können.

Ergebnisse Aus einer vorangegangenen Pilotierungsstudie mit insgesamt 16 Schülerinnen und Schülern konnten erste deskriptive Hinweise gewonnen werden, wonach die Items mit Bezug zu Emotionsregulation und Empathiefähigkeit distinkte Konzepte anzeigen, deren handlungsbezogene (d.h. kommunikativen) Aspekte nur in geringem Maße miteinander kovariieren. Die Belastbarkeit dieser Einschätzung wird auf der Grundlage eines zwischenzeitlich optimierten Testverfahrens und eines angemesseneren Stichprobenumfangs überprüft.

Bischoff-Wanner, C. (2002). Empathie in der Pflege. Bern: Verlag Hans Huber

Böhle, F., & Glaser, J. (2006). Arbeit in der Interaktion - Interaktion als Arbeit. Arbeitsorganisation und Interaktionsarbeit in der Dienstleistung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Dingwall, R., & Allen, D. (2001). The implications of healthcare reforms for the profession of nursing. *Nursing inquiry*, 8(2), 64-74.

Hochschild, A. R. (1990). Das gekaufte Herz. Zur Kommerzialisierung der Gefühle. Frankfurt a. M.: Campus.

Neumann-Poensch, S., Höller, A. (2011). Gefühlsarbeit in Pflege und Betreuung: Sichtbarkeit und Bewertung gelungener Beziehungsarbeit. Wien: Springer-Verlag Vienna.

Rogers, C. R. (1961). On becoming a person: A therapist's view of psychotherapy. Oxford, England: Houghton Mifflin.

Yu, J., & Kirk, M. (2009). Evaluation of empathy measurement tools in nursing: systematic review. *Journal of advanced nursing*, 65(9), 1790-1806.



Tobias Kärner (Otto-Friedrich-Universität Bamberg), Klaudia Golyszny (Otto-Friedrich-Universität Bamberg),  
Detlef Sembill (Otto-Friedrich-Universität Bamberg)

### ***Physiologische, psychologische und situative Effekte auf das Stresserleben von Schülern – Eine Mehrebenen-Längsschnittanalyse im Unterrichtsprozess.***

Schulischer Unterricht wird mit unterschiedlichen Belastungsquellen wie beispielsweise zeitlichem Druck oder Erfolgsdruck in Verbindung gebracht, wobei sich der daraus resultierende Stress u. a. auf individuelle Lernprozesse auszuwirken vermag. Individuelle Stressreaktionen sind hierbei als mehrdimensional anzusehen, da sowohl physiologische als auch psychische Dispositionen und Prozesse involviert sind. So ist eine Erhöhung der individuellen Cortisol-Baseline mit chronischer Belastung assoziiert (Takahashi et al., 2005), das sympathische Nervensystem dominiert unter situativer Stresseinwirkung gegenüber dem parasympathischen Nervensystem (Baert et al., 2012) und die Bewertung individueller Bewältigungsressourcen ist hinsichtlich der Stressverarbeitung von großer Bedeutung (Goetz et al., 2008).

Vor diesem Hintergrund liegt unser Forschungsanliegen in der simultanen Analyse von Einflüssen physiologischer, psychologischer und situativer Merkmale bzw. Zustände auf das subjektive Stresserleben von Schülern.

Der Analyse liegen Daten von insgesamt 28 Schülern einer bayerischen Berufsschule zu-grunde (M = 19.2, SD = 3.8 Jahre), welche über einen Zeitraum von 3 Wochen (9 x 45 Min.) begleitet wurden. Die Videostudie wurde als Produkt-Prozess-Design realisiert, wobei zu-nächst individuelle Ausgangsbedingungen erhoben wurden, bevor situationsabhängige States sowie objektive Umgebungsmerkmale erfasst wurden. Es wurde somit auf 775 synchrone State-Messungen und aggregierte Beobachtungsintervalle zurückgegriffen.

Mittels SPSS® MIXED wurde ein 2-Ebenen-Modell mit autoregressiver Kovarianzstruktur geschätzt, welches die in den Individuen geschachtelten wiederholten State-Messungen berücksichtigt. Es zeigen sich signifikante Effekte individueller Bewältigungsressourcen, situativer Zustände sowie der sympathovagalen Balance auf das subjektive Stresserleben (SSE). Ebenen übergreifende Effekte zwischen physiologischen Merkmalen und psychischen Zuständen, welche die individuellen SSE-Wachstumskurven betreffen, werden herausgestellt.

Unsere Befunde unterstreichen einmal mehr die Bedeutung von individuellen Lerntempi, Zeitdruck, emotionalem Erleben von Unterrichtssituationen und Belastung im Kontext zu erbringender Lernleistungen (sensu Sembill, 2008; Sembill & Zilch, 2010; Sembill, 2012).

Baert S., Casier A. & De Raedt R. (2012). The effects of attentional training on physiological stress recovery after induced social threat. *Anxiety, Stress & Coping*, 25, 365-379.

Goetz, Th., Preckel, F., Zeidner, M. & Schleyer, E. (2008). Big fish in big ponds: A multilevel analysis of test anxiety and achievement in special gifted classes. *Anxiety, Stress & Coping*, 21, 185-198.

Sembill, D. (2008). Zeitver(sch)wendung in Bildungsprozessen. In Gläser-Zikuda, M., Seifried, J. (Hrsg.), *Lehrerexpertise – Analyse und Bedeutung unterrichtlichen Handelns*, (S. 19-46) Münster: Waxmann.

Sembill, D. (2012). Was bedeutet die Hirnforschung für die Schul- und Lernkultur? In Warwas, J., Harder, P., Sembill, D. (Hrsg.), *Kultur der Schule – Schule der Kultur(en)?* (S. 85-112).

Baltmannsweiler: Schneider Verlag Hohengehren.

Sembill, D., Zilch, C. (2010). Leistungsbereitschaft und Leistungseffekte unter dem Druck des Wandels. In Warwas, J., Sembill, D. (Hrsg.), *Schule zwischen Effizienzkriterien und Sinnfragen*, (S. 245-268). Baltmannsweiler: Schneider Verlag Hohengehren.

SPSS (2005). *Linear Mixed-Effects Modeling in SPSS: An Introduction to the MIXED Pro-cedure* [PDF file]. Retrieved from [http://www.spss.ch/upload/1126184451\\_Linear%20Mixed%20Effects%20Modeling%20in%20SPSS.pdf](http://www.spss.ch/upload/1126184451_Linear%20Mixed%20Effects%20Modeling%20in%20SPSS.pdf). Armonk, New York: IBM.

Takahashi, T., Ikeda, K., Ishikawa, M., Kitamura, N., Tsukasaki, T., Nakama, D. & Kameda, T. (2005). Anxiety, reactivity, and social stress-induced cortisol elevation in humans. *Neuroendocrinology Letters*, 26, 351-354.



## **Einflüsse der Faktoren Geschlecht und Leistung auf den Stress im Schulalltag**

Eine langfristig andauernde Belastung durch Stress führt häufig zu vielfältigen psychischen oder physischen Beeinträchtigungen, so dass Stress ein Auslöser für verschiedene Krankheiten sein kann (Noll & Kirschbaum, 2006). Die Weltgesundheitsorganisation erklärte chronischen Stress sogar zur größten Gesundheitsgefahr unseres Jahrhunderts.

Häufig wird Stress im beruflichen Rahmen erforscht. Zu Stressfaktoren im Umfeld Schule gibt es bisher nur wenige Untersuchungen. Zudem konzentrierte sich die Stressforschung bisher wenig auf die Untersuchung von naturalistischen Stressoren (z. B. Martinek, 2003).

Daher sollte in der vorliegenden Studie die Frage beantwortet werden, welche Faktoren die physiologische Stressreaktion sowie das subjektive Stressempfinden im Schulalltag beeinflussen. Hierbei wurden sowohl das Geschlecht als auch die Leistung als Einflussfaktoren betrachtet.

Die Probandengruppe setzt sich aus Abiturienten und Abiturientinnen des doppelten Abiturjahrganges 2013 in Nordrhein-Westfalen zusammen. Die Untersuchung fand an drei verschiedenen Gymnasien des Ruhrgebiets statt. Von den insgesamt 25 Probanden wurden neben den individuellen Stresseinschätzungen auch Stresshormone im Speichel erfasst.

Die Probensammlung fand im Zeitraum der Abiturvorprüfungsphase statt. Hierzu wurden an drei Wochentagen zu je sechs Messzeitpunkten die genannten Stressparameter erfasst.

Zur Messung des subjektiven Stressempfindens stuften die Probanden ihren empfundenen Stress zum Zeitpunkt der Probenentnahme anhand einer visuellen Analogskala (engl. Visual Analogue Scale, VAS) ein. Bei der VAS handelte es sich um eine horizontale Linie (10 cm) mit fixierten Endpunkten, wobei der linke Punkt den Zustand "kein Stress" und der rechte "maximaler Stress" darstellt (Folstein & Luria, 1973).

Zudem wurde die Konzentration des Stresshormons Cortisol, welches als Reaktion auf einen Stressor ausgeschüttet wird, im Speichel mit einem enzymgekoppelten Immunadsorptionstest (IBL Kit, Multiscan FC) gemessen.

Die Untersuchungsergebnisse zeigen, dass das subjektive Stressempfinden der Probanden im Tagesverlauf leicht ansteigt, wobei sich Frauen im Schulalltag signifikant gestresster fühlen als Männer ( $p < 0,001$ , Mann-Whitney U-Test). Im Gegensatz dazu sinkt die Cortisolkonzentration im Tagesverlauf. Frauen zeigen über alle Messzeitpunkte gemittelt eine signifikant höhere Cortisolkonzentrationen als Männer ( $p < 0,05$ , Mann-Whitney U-Test). Zudem wird deutlich, dass leistungsstärkere Abiturienten und Abiturientinnen sich im Vergleich zu leistungsschwächeren subjektiv gestresster ( $p < 0,001$ , Mann-Whitney U-Test) fühlen. Diese Beobachtung spiegelt sich in der Messung des Speichelcortisols wider, leistungsstärkere Abiturienten und Abiturientinnen zeigen eine höhere Cortisolkonzentration als leistungsschwächere ( $p < 0,05$ , Mann-Whitney U-Test).

Die Ergebnisse der Studie zeigen somit, dass sich das subjektive Stressempfinden von der physiologischen Stressreaktion unterscheidet.

Folstein, M. F.; Luria, R. (1973): Reliability, validity and clinical application of the visual analogue mood scale. In: Psychological Medicine 3, S. 479-486.

Martinek, L.; Oberascher-Holzinger, K.; Weishuhn, S.; Klimesch, W.; Kerschbaum, H.H. (2003): Anticipated academic examinations induce distinct cortisol responses in adolescent pupils. In: Neuroendocrinology Letters 24 (6), S. 449-453.

Noll, A; Kirschbaum C. (2006): Stresskrankheiten vorbeugen und behandeln mit chinesischer Medizin. München, Urban & Fischer Verlag.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer





## ***Determinanten kognitiver Angstmanifestation während schulischer Tests: Eine Längsschnittstudie***

Theoretischer Hintergrund und Fragestellung: Schulische Tests dienen der Feststellung von Wissen und Können. Wie jede Messung kann auch diese durch Störvariablen verzerrt werden. Während Tests erlebte Angst ist ein solcher potentieller Störeinfluss: Angst kann Testleistungen beeinträchtigen, da sie sich hinderlich auf die Informationsverarbeitung auswirken kann (Störung des Wissensabrufs und des Denkens; Zeidner, 1998). Dieser kognitiven Angstmanifestation (Rost & Schermer, 1997) liegt die angstbedingte automatische Aufmerksamkeitsorientierung auf aufgaben-irrelevante Gedanken während der Testbearbeitung zugrunde (Eysenck, Derakshan, Santos & Calvo, 2007). Es hat sich allerdings gezeigt, dass Angst nicht grundsätzlich mit Testleistungseinbußen einhergeht (Eysenck et al., 2007). Wir gehen davon aus, dass die Kapazität zur Selbstkontrolle bzw. Aufmerksamkeitsregulation (Baumeister, Muraven & Tice, 2000; Schmeichel & Baumeister, 2010) ein Faktor ist, der bedingt, wie stark kognitive Angstmanifestation während Tests und damit zusammenhängende Leistungseinbußen auftreten. Ängstliche Lernende mit höherer Aufmerksamkeitsregulationskapazität sollten ihrer automatischen Aufmerksamkeitsorientierung auf aufgaben-irrelevante Gedanken erfolgreicher entgegensteuern und ihre Aufmerksamkeit weiterhin auf die Aufgabenbearbeitung richten können. Dem entsprechend sollten sie trotz Angst weniger kognitive Angstmanifestation während Tests erleben als ängstliche Lernende mit geringerer Aufmerksamkeitsregulationskapazität. Auch die Selbstwirksamkeitserwartung und das Selbstwertgefühl sind Variablen, die in der Vergangenheit mit der Bewältigung von Angst und Stress in Verbindung gebracht wurden (Bandura, 1977; Baumeister, Campbell, Krueger & Vohs, 2003). Daher wurden diese Variablen als weitere Prädiktoren berücksichtigt. Es wurde die Hypothese getestet, dass die dispositionelle Aufmerksamkeitsregulationskapazität über die dispositionelle Selbstwirksamkeitserwartung und das dispositionelle Selbstwertgefühl hinaus Veränderungen in der kognitiven Angstmanifestation während Mathematiktests in einer Wirtschaftsschülerstichprobe vorhersagt. Es wurde des Weiteren davon ausgegangen, dass eine indirekte Verbindung zwischen der Aufmerksamkeitsregulationskapazität und der Veränderung in den Mathematiknoten, vermittelt über die Veränderung in der kognitiven Angstmanifestation, besteht.

Methode: Einhundertachtundfünfzig Wirtschaftsschüler bearbeiteten im September 2011 (T1) einen Fragebogen, der die folgenden Messungen enthielt:

- Subskala Kognitive Angstmanifestation aus dem Differentiellen Leistungsangstinventar (Rost & Schermer, 1997) bezogen auf Mathematiktests (Sparfeldt, Schilling, Rost, Stelz & Peipert, 2005); Alpha = .90; - Skala zur dispositionellen Aufmerksamkeitsregulationskapazität (Bertrams & Englert, 2013); Alpha = .88; - Skala zur Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1995); Alpha = .83; - Skala zum Selbstwertgefühl (von Collani & Herzberg, 2003); Alpha = .83; - Angabe der letzten Mathematikzeugnisnote  
Im Februar 2012 (T2), also nach 5 Monaten und kurz nach dem Erhalt des Halbjahreszeugnisses, gaben die Schüler erneut ihre kognitive Angstmanifestation während Mathematiktests (Alpha = .93) und ihre letzte Mathematikzeugnisnote an.

Ergebnisse: Die Daten wurden mittels Korrelationsanalyse, multipler Regressionsanalyse und Bootstrapping ausgewertet. Die Aufmerksamkeitsregulationskapazität, die Selbstwirksamkeitserwartung und das Selbstwertgefühl (alle zu T1) waren positiv interkorreliert,  $r = .50/.59/.59$ . Diese Variablen wurden gemeinsam als Prädiktoren in ein Regressionsmodell zur Vorhersage der kognitiven Angstmanifestation zu T2 eingefügt. Gleichzeitig wurde die kognitive Angstmanifestation zu T1 konstant gehalten. Es zeigte sich, dass die Aufmerksamkeitsregulationskapazität erwartungskonform die Veränderungen in der kognitiven Angstmanifestation vorhersagte,  $\text{Beta} = -.21, p = .02$ . Das heißt, dass höhere Aufmerksamkeitsregulationskapazität zu T1 mit verringerter kognitiver Angstmanifestation zu T2 einherging. Die Selbstwirksamkeitserwartung,  $\text{Beta} = .12, p = .14$ , und das Selbstwertgefühl,  $\text{Beta} = .05, p = .54$ , hatten hingegen keinen eigenen Vorhersagewert für die Veränderungen in der kognitiven Angstmanifestation. Des Weiteren ergab eine Mediationsanalyse mittels Bootstrapping (bias-corrected bootstrap 95% confidence interval, 5000 resamples; siehe Hayes & Scharkow, in press), dass die Aufmerksamkeitsregulationskapazität (T1), vermittelt über die Veränderung in der kognitiven Angstmanifestation, indirekt mit der Veränderung in der Mathematikleistung verbunden war (d.h. das Bootstrap-Konfidenzintervall schloss nicht die Null ein; CI [0.01, 0.24]). Für die Selbstwirksamkeitserwartung und das Selbstwertgefühl fand sich keine analoge indirekte Verbindung zur Mathematikleistung.

Fazit: Die Befunde verweisen auf die Bedeutsamkeit der Aufmerksamkeitsregulationskapazität für die Bewältigung kognitiver Angstreaktionen während schulischer Tests.

Losgelöst von der Aufmerksamkeitsregulationskapazität scheinen positive Erwartungen und ein positives Selbstbild keine protektive Wirkung hinsichtlich der leistungsbeeinträchtigenden kognitiven Angstmanifestation während Mathematiktests zu besitzen.

Bandura, A. (1977). Self-efficacy: Towards a unifying theory of human behavior. *Psychological Review*, 84, 191-215.

Baumeister, R. F., Campbell, J. D., Krueger, J. I. & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4, 1-44.

Baumeister, R. F., Muraven, M. & Tice, D. M. (2000). Ego depletion: A resource model of volition, self-regulation, and controlled processing. *Social Cognition*, 18, 130-150.

Bertrams, A. & Englert, C. (2013, July 27). Umsetzung subjektiver Sporthäufigkeitsstandards - Zur Rolle dispositionell verfügbarer Selbstkontrollkraft. *Sportwissenschaft*. Advance online publication.

Collani, G. von & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 3-7.

Eysenck, M. W., Derakshan, N., Santos, R. & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7, 336-353.

Hayes, A. F. & Scharkow, M. (in press). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*.

Rost, D. H. & Schermer, F. J. (1997). *Handbuch Differentielles Leistungsangstinventar*. Frankfurt: Swets Test Services.

Schmeichel, B. J. & Baumeister, R. F. (2010). Effortful attention control. In B. Braya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 29-49). Cambridge, MA: MIT Press.

Schwarzer, R. & Jerusalem, M. (1995). Generalized Self-Efficacy Scale. In J. Weinman, S. Wright & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

Sparfeldt, J. R., Schilling, S. R., Rost, D. H., Stelz, I. & Peipert, D. (2005). Leistungsängstlichkeit: Facetten, Fächer, Fachfacetten? Zur Trennbarkeit nach Angstfacette und Inhaltsbereich. *Zeitschrift für Pädagogische Psychologie*, 9, 225-236.

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York, NY: Plenum Press.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
im Foyer



Wiebke Evers (ZNL – TransferZentrum für Neurowissenschaften und Lernen), Sonja Quante (ZNL), Laura Walk (ZNL),  
Melanie Otto (ZNL), Pamela Amuri (ZNL)

## **Ergebnisse der wissenschaftlichen Begleitung des Projekts „EMIL – Emotionen regulieren lernen“ zur Förderung der sozial-emotionalen Kompetenz im Kindergarten**

Ziel der Studie: Das Forschungsprojekt „EMIL – Emotionen regulieren lernen“ wurde vom ZNL TransferZentrum für Neurowissenschaften und Lernen mit Unterstützung der Baden-Württemberg Stiftung im Jahr 2011 ins Leben gerufen. Ziel des Projekts ist es die sozial-emotionalen Kompetenzen von Kindern im Kindergartenalter durch Stärkung von Selbstregulationsfähigkeiten und Perspektivenübernahme zu fördern. Dies geschieht durch ein alltagsintegriertes Förderkonzept: Durch Fortbildungen und Teambegleitungen sowie zahlreiche Praxisanregungen werden pädagogische Fachkräfte befähigt, die vielfältigen Möglichkeiten für eine alltagsintegrierte Förderung im Kindergarten zu erkennen und zu nutzen.

Hintergrund: Ein zentraler Ansatzpunkt bei der Förderung der sozial-emotionalen Kompetenz stellen die exekutiven Funktionen dar. Exekutive Funktionen beschreiben Kontroll- und Regulierungsprozesse, die eine bewusste Steuerung von Denkprozessen und Aufmerksamkeit, dem Verhalten und Emotionen ermöglichen (z.B. Blair, 2002; Drechsler, 2007; Miller & Cohen, 2001). Gut ausgebildete exekutive Funktionen sind eine wichtige Voraussetzung für den Erfolg in Schule, Beruf und im Sozialleben. Sie stehen langfristig in Zusammenhang mit einer der körperlichen und mentalen Gesundheit (z.B. Moffitt et al., 2011). Zudem verringern gut ausgebildete exekutive Funktionen und selbstregulatorische Fähigkeiten das Risiko zu delinquenten und antisozialen Verhaltensweisen (vgl. Hughes & Ensor, 2009; Trentacosta & Shaw, 2009). Kinder mit gut ausgebildeten exekutiven Funktionen sind in der Lage, die eigenen Wünsche und Bedürfnisse auch mal nach zurückzustellen und andere Meinungen zu akzeptieren. In frustrierenden Situationen unterstützen die exekutiven Funktionen den Umgang mit negativen Gefühlen und die Regulierung des eigenen Verhaltens und nehmen damit erheblichen Einfluss auf die sozial-emotionale Kompetenz (Diamond, Barnett, Thomas, & Munro, 2007; Diamond, Kirkham, & Amso, 2002).

Methode: Für diese Studie wurden zwischen Januar und Oktober 2012 32 pädagogische Fachkräfte aus vier Ulmer Kindergärten durch Referenten des ZNL qualifiziert. Zu drei Messzeitpunkten (November 2011, Juli 2012 und November 2012) wurden in diesen vier Kindergärten (Interventionsgruppe) sowie in vier weiteren Kindergärten (Wartegruppe) bei insgesamt 142 Kindern entwicklungspsychologische Tests zur Messung von exekutiven Funktionen (Arbeitsgedächtnis, Inhibition, kognitive Flexibilität), Perspektivenübernahme, Beliebtheit (Soziogramm) und der sozialer Selbstwirksamkeitserwartung durchgeführt. Diese sollten der Beantwortung der folgenden Forschungsfrage dienen: Zeigen sich durch Qualifizierung der pädagogischen Fachkräfte Veränderungen in Bezug auf die exekutiven Funktionen und die Fähigkeit zur Perspektivenübernahme von Kindern?

Ergebnisse: Es werden die signifikanten Ergebnisse der folgenden Testverfahren vorgestellt: Digit Span-Test vorwärts zur Messung des auditiven Kurzzeitgedächtnisses (Petermann & Petermann, 2008), Block Recall-Test zur Erfassung des visuellen Arbeitsgedächtnisses (Gathercole, Pickering, Ambridge, & Wearing, 2004) und der Tower Building-Test zur Messung der Verhaltenssteuerung und Perspektivenübernahme (Kochanska, Murray, & Harlan, 2000). Varianzanalysen mit Messwiederholung zeigen, dass sich die Kinder der Interventionsgruppe in Bezug auf ihr Arbeitsgedächtnis und ihre Fähigkeit zur Verhaltenssteuerung und Perspektivenübernahme im Vergleich zur Wartegruppe über die drei Messzeitpunkte hinweg signifikant besser entwickelten. Die Ergebnisse des Digit Span-Tests zeigten, dass sich die Kinder der Interventionsgruppe signifikant besser entwickelten als die Kinder der Wartegruppe. Gleiches zeigte auch der Block Recall-Test, bei dem es darum geht, Holzblöcke in der gleichen Reihenfolge anzutippen wie es der Versuchsleiter vormacht. Auch beim Tower Building-Test entwickelten sich die Kinder der Interventionsgruppe signifikant besser als die Kinder der Wartegruppe. Bei diesem Test ist das Kind angehalten, mit dem Versuchsleiter gemeinsam einen Turm zu bauen und sich dabei mit dem Legen von Bausteinen mit dem Versuchsleiter abzuwechseln. Hier fiel es den Kinder der, trotz Wartezeit den Versuchsleiter mitzudenken und mit ihm zusammenspielen.

Schlussfolgerung: Diese Studie konnte zeigen, dass eine alltagsintegrierte Förderung von qualifizierten pädagogischen Fachkräften die exekutiven Funktionen und die Perspektivenübernahme von Kindern stärken kann.

- Blair, C. (2002). Early intervention for low birth weight, preterm infants: The role of negative emotionality in the specification of effects. *Development and Psychopathology*, 14(2), 311-332.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, 318(5855), 1387-1388. doi: 10.1126/science.1151148
- Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38(3), 352-362. doi: 10.1037//0012-1649.38.3.352
- Drechsler, R. (2007). Exekutive Funktionen. *Zeitschrift für Neuropsychologie*, 18(3), 233-248. doi: 10.1024/1016-264x.18.3.233
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, 40(2), 177-190. doi: 10.1037/0012-1649.40.2.177
- Hughes, C. H., & Ensor, R. A. (2009). How Do Families Help or Hinder the Emergence of Early Executive Function? Social Interaction and the Development of Executive Function, 123, 35-50. doi: 10.1002/Cd.234
- Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: continuity and change, antecedents, and implications for social development. *Dev Psychol*, 36(2), 220-232.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. [Article]. *Annual Review of Neuroscience*, 24(1), 167.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc Natl Acad Sci U S A*, 108(7), 2693-2698. doi: 10.1007/6108 [pii] 10.1073/pnas.1010076108
- Petermann, F., & Petermann, U. (2008). HAWIK-IV: Hamburg-Wechsler Intelligenztest für Kinder-IV: H. Huber;[New York].
- Trentacosta, C. J., & Shaw, D. S. (2009). Emotional Self-Regulation, Peer Rejection, and Antisocial Behavior: Developmental Associations from Early Childhood to Early Adolescence. *Journal of Applied Developmental Psychology*, 30(3), 356-365. doi: 10.1016/j.appdev.2008.12.016



## **Zusammenhänge der frühen familiären Lernumwelt und der kindlichen Fähigkeit zur Verhaltensregulation mit schulischen Vorläuferfertigkeiten im Kindergartenalter**

Lesen und Schreiben sind grundlegende Fähigkeiten zum erfolgreichen Meistern schulischer Anforderungen. Bereits beim Eintritt in den Kindergarten existieren interindividuelle Unterschiede in entsprechenden Vorläuferfertigkeiten, wie Wortschatz oder grammatikalisches Sprachverständnis. Inwiefern Eltern die häusliche Lernumwelt („Home learning environment“, HLE) aktiv und passiv gestalten, kann einen wesentlichen Einfluss auf die Entwicklung dieser Vorläuferfertigkeiten haben (Niklas & Schneider, 2010; Davidse et al., 2011). Ebenso werden Merkmale der sozialen Herkunft als maßgeblich bestimmt (Artelt et al., 2001; Niklas & Schneider, 2013). Nur wenige Studien berücksichtigen neben diesen kontextuellen Einflussfaktoren weitere individuelle Fähigkeiten des Kindes, die sich zur gleichen Zeit entwickeln. So ist weitgehend bestätigt, dass Kinder mit einer verzögerten Entwicklung der Fähigkeit und Motivation zur Selbstregulation später vermehrt Schwierigkeiten mit schulischen Anforderungen haben (Blair, 2002; von Suchodoletz et al., 2012). Ziel dieser Studie ist die Identifizierung bedeutsamer Einflussfaktoren auf die Entwicklung schulischer Vorläuferfertigkeiten für Lesen und Schreiben im frühen Kindergartenalter. Dazu werden Zusammenhänge dieser schulischen Vorläuferfertigkeiten mit kontextuellen Faktoren in der Familie (HLE, strukturelle Herkunftsmerkmale) sowie mit individuellen Faktoren (kindliche Fähigkeit zur Verhaltensregulation) im selben Modell untersucht.

Teilnehmer waren N = 218 Kinder (49% Mädchen) am Ende des ersten Kindergartenjahres und ihre Mütter. Die Daten wurden zum ersten Messzeitpunkt einer längsschnittlichen Studie im süddeutschen Raum erhoben. Die Mütter beantworteten einen Fragebogen zur familiären Lernumwelt (Niklas & Schneider, 2010) und zu strukturellen Herkunftsmerkmalen (Bildungsniveau der Eltern, Migrationshintergrund des Kindes, Zugehörigkeit zur sozioökonomischen Schicht). Die Kinder (Alter M = 37.57 Monate, SD = 4.39) wurden in Einzelsitzungen zu ihrem Verständnis und ihrer Imitationsfähigkeit grammatikalischer Strukturen (Heidelberger Sprachentwicklungstest; Grimm & Schöler, 1991), ihrem Wortschatz (Peabody Picture Vocabulary Test-4; Dunn & Dunn, 2007) und ihrer Fähigkeit zur Verhaltensregulation (Head-Toes-Knees-and-Shoulders Task; Cameron Ponitz et al., 2009) getestet. In den Analysen wurde neben dem Alter und Geschlecht der Kinder, auch für ihre Arbeitsgedächtnisleistung (Subtests Zahlennachsprechen und Handbewegungen der Kaufman Assessment Battery for Children; Kaufmann et al., 2006) und den Aufenthalt im Kindergarten in Stunden pro Tag kontrolliert.

Die ersten Ergebnisse hierarchischer Regressionen mit Bayes Schätzung zeigten, dass das Verständnis grammatikalischer Strukturen mit der familiären Lernumwelt ( $\beta = .10$ , Bayesian Confidence Intervall (BCI) .047, .352) als auch der Fähigkeit zur Verhaltensregulation ( $\beta = .28$ , BCI .116, .434), nicht aber mit den strukturellen Herkunftsmerkmalen zusammen hängt. Demgegenüber zeigten die Ergebnisse keine signifikanten Zusammenhänge zwischen der Imitationsfähigkeit grammatikalischer Strukturen und dem Wortschatz weder mit der familiären Lernumwelt noch der Fähigkeit zur Verhaltensregulation. Für diese Indikatoren rezeptiver Sprachentwicklung waren die strukturellen Merkmale Migrationshintergrund ( $\beta = -.24$ , BCI -.397, -.079 für Imitation,  $\beta = -.33$ , BCI -.489, -.170 für Wortschatz), Bildungsniveau des Vaters ( $\beta = .19$ , BCI .012, .368 für Wortschatz) und die Arbeitsgedächtnisleistung ( $\beta = .290$ , BCI .134, .440 für Imitation,  $\beta = .29$ , BCI .136, .432 für Wortschatz) bedeutsamer. Die Studie betont die differenzielle Bedeutung der familiären Lernumwelt und der Fähigkeit zur Verhaltensregulation für die Entwicklung schulischer Vorläuferfertigkeiten für das Lesen und Schreiben bei Kindern zu Beginn des Kindergartenalters. Insbesondere das Verständnis grammatikalischer Strukturen hängt mit der häuslichen Ausgestaltung der Lernumwelt und der Fähigkeit zur Verhaltensregulation positiv zusammen. Gleichzeitig haben strukturelle Herkunftsmerkmale wesentlichen Einfluss auf die Imitationsfähigkeit grammatikalischer Strukturen und den Wortschatz. Implikationen dieser Ergebnisse werden sowohl für die zukünftige Forschung als auch für die praktische Anwendung zur Unterstützung der Kinder bei der Entwicklung schulischer Vorläuferfertigkeiten diskutiert. Darüber hinaus werden für das endgültige Poster zusätzliche längsschnittliche Daten in die Analysen einbezogen, um Aussagen zur Entwicklung über das Kindergartenalter hinweg treffen zu können.

- Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., et al. (Hrsg.) (2001). PISA 2000. Zusammenfassung zentraler Befunde. Berlin: Max-Planck-Institut für Bildungsforschung.
- Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, 57, 111-127.
- Cameron Ponitz, C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, 45, 605-619.
- Davidse, N. J., de Jong, M. T., Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (2011). Cognitive and environmental predictors of early literacy skills. *Reading and Writing*, 24, 395-412.
- Dunn, M., & Dunn, L. M. (2007). Peabody Picture Vocabulary Test-4. Circle Pines, MN: AGS.
- Grimm, H. & Schöler, H. (1991). Der Heidelberger Sprachentwicklungstest H-S-E-T (2., verbesserte Auflage). Göttingen: Hogrefe.
- Kaufman, A. S., Kaufman, N. L., Melchers, P., & Preuß, U. (2006). K-ABC Kaufman Assessment Battery for Children - German Version. Göttingen: Hogrefe.
- Niklas, F., & Schneider, W. (2010). Der Zusammenhang von familiärer Lernumwelt mit schulrelevanten Kompetenzen im Vorschulalter. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 30(2), 148-164.
- Niklas, F., & Schneider, W. (2013). Home literacy environment and the beginning of reading and spelling. *Contemporary Educational Psychology*, 38, 40-50.
- von Suchodoletz, A., Gestsdottir, S., Wanless, S. B., McClelland, M. M., Birgisdottir, F., Gunzenhauser, C., & Ragnarsdottir, H. (2013). Behavioral self-regulation and relations to emergent academic skills among children in Germany and Iceland. *Early Childhood Research Quarterly*, 28(1), 62-73.



## **Das Freizeitbudget von ganztägig betreuten Kindergartenkindern**

Der fortschreitende Ausbau der institutionellen ganztägigen Kindertagesbetreuung, insbesondere in den alten Bundesländern, lässt eine veränderte Zeitstruktur des Kinderalltags vermuten. Eine folgenreiche Zeitverdichtung der kindlichen Freizeitgestaltung außerhalb der Kindertagesstätte ist wahrscheinlich. Es ist außerdem davon auszugehen, dass die Eltern einen bedeutsamen Einfluss auf das Freizeitverhalten ihrer Kinder haben. Ihr erhöhter Fürsorgebedarf macht Kindergartenkinder weitestgehend von ihren Eltern abhängig, sie sind an deren Pflichten und Zeitrhythmus gebunden, ihre (Frei-)Zeitverwendung wird maßgeblich von den Eltern angeregt. Das Freizeitverhalten deutscher Kinder ist äußerst heterogen, wobei sich bedeutsame Ungleichheiten unter anderem in den verschiedenen strukturellen Rahmenbedingungen, wie z.B. den sozioökonomischen Lebenslagen der Familien finden. Vermutlich beeinflusst zusätzlich die familiäre Anregungsqualität, z.B. das Familienklima im Allgemeinen oder spezifische Anregungen das kindliche Freizeitbudget. Der vorliegende Beitrag untersucht, ob eine ganztägige Betreuung der Kindergartenkinder eine geringere und andere Freizeitverwendung im ersten und letzten Kindergartenjahr zur Folge hat. Inwieweit ausgewählte familiäre Hintergrundmerkmale sowie die globale und bereichsspezifische Anregungsqualität früherer Les- und Schreibfertigkeiten einen Einfluss auf die Freizeitverwendung von Kindern haben, stellt einen weiteren Interessenschwerpunkt der Analyse dar. Im Fokus stehen die Freizeitdimensionen: „Mediennutzung“ und „Wort- und Sprachspiele/ Lesen und Vorlesen“, da sie in der öffentlichen Diskussion um zeitgenössische Kinderfreizeit häufig als konkurrierende Aktivitäten vermutet werden. Datengrundlage ist die Längsschnittstudie „Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter (BiKS) 3-10. Die kindlichen Aktivitäten, deren Dauer und Reihenfolge an einem Wochentag konnten mittels eines eigens entwickelten Zeittagebuches von insgesamt 388 Kindergartenkindern erfasst werden, das während der gesamten Kindergartenzeit jährlich zum Einsatz kam. Verglichen wurde die frei verfügbare Zeit der Kinder, die an einem 15-stündigen Wochentag (von 6 Uhr bis 21 Uhr) im Durchschnitt weniger als fünf Stunden in der Kindertagesstätte verbrachten, mit der disponiblen Zeit derer, die mehr als fünf Stunden außerfamiliär betreut wurden. Um fundierte Aussagen über Unterschiede durch die Ganztagsbetreuung geben zu können, wurden multivariate Regressionsanalysen unter Einbezug familiärer Hintergrundmerkmale und Anregungsqualitäten durchgeführt. Insgesamt berichten die Ergebnisse, dass das Zeitvolumen des Medienkonsums sowie der Vorlesenszeit und der Beschäftigung mit Büchern, Zeitschriften, etc. jenseits eines ganztägigen Kindergartenbesuches nur geringfügig einschränkt wird. Im ersten Kindergartenjahr verbringen halb- wie ganztags betreute Kinder ca. 20 Minuten mit Aktivitäten rund um das Lesen. Im Vorschuljahr beschäftigen sich die Halbtagskinder eine Viertelstunde signifikant länger mit Wort- und Sprachspielen, Lesen und Vorlesen. Beim Fernsehen, CD hören und Computerspielen verhält es sich umgekehrt. Die halbtags betreuten Kindergartenkinder verwenden im ersten Jahr signifikant mehr Zeit mit technologischen Medien, wenngleich die Differenz lediglich sieben Minuten beträgt. Ein Grund für die ähnlichen Zeitmuster können ähnliche Tageszeitenpräferenzen für diese Aktivitäten sein, die am Nachmittag bzw. Abend liegen und deshalb kaum von der Betreuungsdauer des Kindergartens tangiert werden. Unter Kontrolle der Betreuungsdauer der Kinder scheinen Ungleichheiten aufgrund familiärer Hintergrundmerkmale nur abgeschwächt zu gelten. Die Annahme, dass Benachteiligungen im Freizeitbudget aufgrund des ethnischen Hintergrundes vorhanden sein könnten, kann für keines der Aktivitäten zu keinem Zeitpunkt bestätigt werden. Gleiches gilt für die durchschnittliche Wochenarbeitszeit der Eltern. Der sozioökonomische Status beeinflusst die Mediennutzungsdauer im ersten Kindergartenjahr negativ, was sich im letzten Kindergartenjahr nur noch tendenziell zeigt. Die Erklärungskraft sozio-struktureller Familienmerkmale fällt allerdings nur schwach aus. Vielmehr kann die familiäre Anregung in Literacy zur Varianzaufklärung beitragen. Steigt die bereichsspezifische Anregungsqualität, erhöht sich das Zeitvolumen, das mit Wort- und Sprachspielen, Lesen und Vorlesen gefüllt wird. Gleichzeitig verringert sich der Medienkonsum. Das trifft für das erste und letzte Kindergartenjahr gleichermaßen zu. Der soziale Hintergrund verliert im Laufe der Kindergartenzeit an Bedeutung, wohingegen die familiäre Stimulation schriftsprachlicher Vorläuferfertigkeiten im ersten und letzten Kindergartenjahr die Dauer des Lesens und ähnlicher Aktivitäten anregen und gleichzeitig den Medienkonsum reduzieren. Die Ergebnisse werden im Zusammenhang mit der Freizeit als Bildungsort diskutiert, dem eine hohe Bedeutung für eine gelingende Schulkarriere zugeschrieben wird.



## **Ethnische Unterschiede bei der Inanspruchnahme und Wirkung von außerfamiliärer U-3-Kinderbetreuung**

Theoretischer Hintergrund und Hypothesen: Die außerfamiliäre Betreuung Unterdreijähriger ist in den letzten Jahren zu einem viel diskutierten Thema im öffentlichen und wissenschaftlichen Kontext und bildungspolitisch zu einem relevanten Ziel geworden. Unter anderem erhofft man sich, dass frühkindliche Bildung die kognitiven und sprachlichen Fähigkeiten fördert und somit langfristig zur Reduktion von sozialer und ethnischer Bildungsungleichheit beiträgt (Cunha et al., 2006; Heckman & Masterov, 2007). Unser Beitrag greift zwei wichtige Aspekte dieses Themas auf und fokussiert zugleich den Aspekt möglicher ethnischer Ungleichheiten. Zunächst untersuchen wir soziale und ethnische Unterschiede der Inanspruchnahme externer U-3-Kinderbetreuung. Nach ökonomischen Ansätzen wird Kinderbetreuung zum einen als Instrument zur Maximierung des Haushaltseinkommens genutzt, zum anderen zur Erhöhung des Humankapitals der Kinder (Geier & Riedel, 2009; Heckman, 1974). Soziologische Ansätze betonen dagegen die Wichtigkeit von Werten und Einstellungen (Meyers & Jordan, 2006). In Einklang mit den genannten theoretischen Argumenten und der bisher nur spärlich vorhandenen empirischen Evidenz (Geier & Riedel, 2009; Schlack et al., 2007) vermuten wir, dass die Inanspruchnahme vor allem von der Erwerbstätigkeit der Mutter abhängt (Hypothese 1). Darüber hinaus erwarten wir eine größere Inanspruchnahme durch höher gebildete Eltern (Hypothese 2). Schließlich gehen wir davon aus, dass Familien mit Migrationshintergrund aufgrund traditionellerer Rollenbilder U-3-Betreuung seltener nutzen (Hypothese 3). In unserer zweiten Analyse untersuchen wir, ob sich eine U-3-Kinderbetreuung positiv auf die Entwicklung von Kindern mit und ohne Migrationshintergrund auswirkt. Hierbei greifen wir auf Theorien von Vygotski (1962) und Bronfenbrenner (1976, 1977) zurück. Frühkindliche Betreuungsinstitutionen können, neben der familiären Umwelt, als bedeutende „Mikrosysteme“ für die Kindesentwicklung verstanden werden. Durch den spielerischen Umgang mit gleichaltrigen bzw. älteren Kindern und Erziehern innerhalb der Betreuungseinrichtungen können Kinder in ihrer „Zone der proximalen Entwicklung“ unterstützt werden. Konkret gehen wir davon aus, dass sich der Besuch einer U-3-Kinderbetreuung positiv auf die allgemeinen kognitiven Fähigkeiten von Kindern mit und ohne Migrationshintergrund auswirkt (Hypothese 4). Zusätzlich erwarten wir unterschiedliche Effekte für die beiden Gruppen bezüglich der Entwicklung der deutschen Sprachfähigkeit. Da sich das familiäre „Mikrosystem“ von Kindern mit Migrationshintergrund vermutlich von dem autochthoner Kinder bezüglich der familiären Umgangssprache unterscheidet und der Kontakt jener zur deutschen Sprache hauptsächlich im außerfamiliären Umfeld stattfindet, erwarten wir einen besonders ausgeprägten positiven Effekt externer frühkindlicher Betreuung auf ihre Deutschsprachfähigkeiten. (Hypothese 5). Spezielle Studien, die sich mit U-3-Kinderbetreuung beschäftigen, sind in Deutschland rar (Tietze et al., 2012), internationale Studien kommen zu gemischten Ergebnissen (vgl. Bornstein et al., 2006; Bowes et al., 2009; Sammons et al., 2002).

Daten und Methoden: In unseren Analysen verwenden wir den Datensatz „Erwerb von sprachlichen und kulturellen Kompetenzen von Migrantenkindern in der Vorschulzeit“. Insgesamt wurden jeweils ca. 600 Familien mit und ohne türkischen Migrationshintergrund mit Kindern zwischen 3 und 4 Jahren befragt. Neben persönlichen Interviews mit der Hauptbetreuungsperson des Zielkindes wurde mit diesem der standardisierte Entwicklungstest K-ABC durchgeführt. Anhand logistischer Regressionen werden familiäre Bedingungen ermittelt, welche die Wahrscheinlichkeit der Nutzung einer externen U-3-Kinderbetreuung beeinflussen. In weiterführenden Analysen untersuchen wir durch lineare Regressionen, ob sich eine solche Betreuung förderlich auf die Testergebnisse zu kognitiven Fähigkeiten und frühen deutschen Sprachfähigkeiten auswirkt. Ergebnisse: Die Analyse zur Nutzung einer U-3-Betreuung zeigt, dass diese von Familien mit türkischem Migrationshintergrund deutlich seltener in Anspruch genommen werden. Hochgebildete Eltern und Familien mit einer erwerbstätigen Hauptbetreuungsperson nehmen sie hingegen häufiger in Anspruch. Somit werden unsere Hypothesen 1 bis 3 bestätigt. Bei den Analysen zur entwicklungsfördernden Wirkung der Betreuung von Kindern unter drei Jahren zeigen sich weder im Gesamtsample noch nach Migrationshintergrund unterschieden positive Effekte für die kognitiven Fähigkeiten. Hypothese 4 kann somit nicht bestätigt werden. Hinsichtlich des expressiven Wortschatzes zeigt sich in Übereinstimmung mit unserer Hypothese 5 erst nach getrennter Betrachtung von Kindern mit und ohne Migrationshintergrund ein positiver Effekt auf die Leistung ersterer.

- Bornstein, M. H., Hahn, C. S., Gist, N. F. & Haynes, O. M. 2006, Long-term cumulative effects of childcare on children's mental development and socioemotional adjustment in a non-risk sample: the moderating effects of gender. *Early Child Development and Care*, 176 (2), 129-156.
- Bowes, J., Harrison, L., Sweller, N., Taylor, A. & Neilsen-Hewett, C. 2009, From child care to school: Influences on children's adjustment and achievement in the year before school and the first year of school. Findings from the child care choices longitudinal extension study. Research Report to the NSW Department of Community Services, Macquarie University; Charles Sturt University.
- Bronfenbrenner, U. 1976, *Experimental Ecology Of Education*. Teachers College Record, 78 (2), 157-204.
- Bronfenbrenner, U. 1977, *Toward An Experimental Ecology Of Human-Development*. *American Psychologist*, 32 (7), 513-531.
- Cunha, F., Heckman, J. J., Lochner, L. J. & Masterov, D. V. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." Pp. 697-812 in *Handbook of the Economics of Education*, edited by E. A. Hanushek & F. Welch. Amsterdam: North-Holland.
- Geier, B. & Riedel, B. 2009. "Ungleichheiten der Inanspruchnahme öffentlicher frühpädagogischer Angebote. Einflussfaktoren und Restriktionen elterlicher Betreuungsentscheidungen." Pp. 11-28 in *Frühpädagogische Förderung in Institutionen*, edited by H.-G. Roßbach & H.-P. Blossfeld: VS Verlag für Sozialwissenschaften.
- Heckman, J. J. 1974, Effects of Child-Care Programs on Womens Work Effort. *Journal of Political Economy*, 82 (2), S136-S163.
- Heckman, J. J. & Masterov, D. V. 2007, The Productivity Argument for Investing in Young Children. *Review of Agricultural Economics*, 29 (3), 446-493.
- Meyers, M. K. & Jordan, L. P. 2006, Choice and Accommodation in Parental Child Care Decisions. *Community Development*, 37 (2), 53-70.
- Schlack, R., Hölling, H. & Kurth, B. M. 2007, Inanspruchnahme außerfamiliärer vorschulischer Kindertagesbetreuung und Einfluss auf Merkmale psychischer Gesundheit bei Kindern. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 50 (10), 1249-1258.
- Tietze, W., Bekker-Stoll, F., Bense, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., Keller, H. & Leyendecker, B. 2012, NUBBEK - Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit. Fragestellung und Ergebnisse im Überblick. Forschungsbericht, verlag das netz.
- Vygotskij, L. S., Hanfmann, E. & Vakar, G. 1962. *Thought and language*, M.I.T. Press [u.a.].





Hendrik Lohse-Bossenz (PädQUIS gGmbH), Daniel Schmerse (PädQUIS gGmbH), Wolfgang Tietze (PädQUIS gGmbH), Miriam Janke (Forscherstation Heidelberg gGmbH), Monika Zimmermann (Forscherstation Heidelberg gGmbH)

## **Entwicklung eines Beobachtungsverfahrens zur Erfassung pädagogischer Qualität früher naturwissenschaftlicher Bildung**

Hintergrund: Ziel des vorliegenden Beitrags ist die Darstellung der Entwicklung eines standardisierten Beobachtungsverfahrens, anhand dessen die pädagogische Qualität in Kindertageseinrichtungen in Bezug auf frühe naturwissenschaftliche Bildung beurteilt werden kann. Die zunehmende Bedeutung der frühen naturwissenschaftlichen Bildung führt zu einer Vielzahl von Fortbildungskonzepten und pädagogischen Ansätzen (Evanschitzky, Lohr & Hille, 2008; French, 2004; Fthenakis, Wendell, Eitel, Daut & Schmitt, 2012; Gelman & Brennemann, 2004; Steffensky, Lankes, Carstensen & Nölke, 2012; Zimmermann, 2011), die – jeweils vor verschiedenen theoretischen Hintergründen – unterschiedliche Zielsetzungen und Schwerpunkte verfolgen.

Der Nachweis der Wirksamkeit entsprechender Fortbildungsangebote im Rahmen entsprechender Evaluationen nimmt selten das tatsächliche Verhalten der pädagogischen Fachkräfte in den Blick (Prozessqualität), sondern fokussiert in erster Linie auf den angenommenen Nutzen der Fortbildungsinhalte bzw. auf pädagogische Einstellungen und Orientierungen (z.B. Zimmermann, 2011). Eine Ausnahme hiervon bilden aufwändige und kostenintensive Videoanalysen, die für eine flächendeckende Evaluation von Fortbildungsprogrammen ungeeignet sind. Zusätzlich ist zu konstatieren, dass die Evaluationen oft an dem der Fortbildung zugrunde liegenden pädagogischem Konzept orientiert sind. Hierzu zählen z.B. der Situationsansatz (Preissing, 2007), der Ko-Konstruktivismus (Fthenakis et al., 2012) oder die Didaktik Wagenscheins (Wagenschein, 1997). Dies erschwert vergleichende Analysen verschiedener Fortbildungsprogramme auf der Ebene der tatsächlichen pädagogischen Prozesse.

Zielstellung: Basierend auf dem Modell pädagogischer Qualität in Kindertageseinrichtungen (Orientierungs-, Struktur- und Prozessqualität, Tietze et al., 1998) wird daher in der vorliegenden Studie ein konzeptunabhängiges standardisiertes Beobachtungsverfahren entwickelt, mit den Zielen a) die Qualität pädagogischer Angebote im Bereich der naturwissenschaftlichen Bildung zu beschreiben, b) Interaktionsunterschiede auf Grund verschiedener pädagogischer Ansätzen zu identifizieren und c) Veränderungen durch Fortbildungen abzubilden.

Methode: Ausgehend von bestehenden Verfahren (z.B. KES-R und KES-R-E) erfolgt eine Fokussierung der Inhalte auf frühe naturwissenschaftliche Bildung unter besonderer Berücksichtigung der Erzieher-Kind-Interaktion und der Unterscheidung verschiedener pädagogischer Ansätze.

Eine inhaltliche Validierung erfolgt anhand von Urteilen ausgewiesener Expert(inn)en (Februar 2014). Nach der Schulung von Beobachter(inn)en wird das so entwickelte Verfahren im Frühjahr 2014 in ca. 80 frühpädagogischen Bildungseinrichtungen in Baden-Württemberg eingesetzt. Nach ca. einem Jahr werden alle Erzieher(innen) erneut beobachtet. Ein Teil dieser Stichprobe besucht in der Zwischenzeit eine mehrmonatige Fortbildung im Bereich früher naturwissenschaftlicher Bildung.

Hypothesen: Im Rahmen eines Prä-Post-Vergleichs erfolgt ein Nachweis der Veränderungssensitivität und der Wirkung der Fortbildung auf verschiedene Qualitätsbereiche. Da die Fortbildung auf der Didaktik Wagenscheins (1997) beruht, sollten sich Verhaltensänderungen bei den Erzieher(inn)en der Fortbildungsgruppe zeigen, die vor allem die Selbsttätigkeit der Kinder unterstützen. Detaillierte Analysen von Videosequenzen, bei denen Interaktionen in frühen naturwissenschaftlichen Bildungsangeboten differenzierter betrachtet werden können, sollen zusätzliche Hinweise liefern, die zeigen, dass auch mittels einer Vor-Ort-Beobachtung zentrale Interaktionsmerkmale valide erfasst werden können.

Diskussion: Eine erfolgreiche Entwicklung eines solchen konzeptübergreifenden Beobachtungsverfahrens ermöglicht verhaltensnahe komparative Analysen verschiedener Ansätze aus allgemeiner Pädagogik, Fachdidaktik und pädagogischer Psychologie im Bereich früher naturwissenschaftlicher Bildung sowie eine umfassende Evaluation bestehender und neu entwickelter Fortbildungsangebote.

Evanschitzky, P., Lohr, Ch. & Hille, K. (2008). Mathematische und naturwissenschaftlich-technische Bildung im Kindergarten – Untersuchung der Wirksamkeit einer beruflichen Weiterbildung von Erzieherinnen. *Diskurs Kindheits- und Jugendforschung*, 4, 469-481.

French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, 19, 138-149.

Fthenakis, W. E., Wendell, A., Eitel, A., Daut, M. & Schmitt, A. (2012). *Natur-Wissen schaffen – Band 3: Frühe naturwissenschaftliche Bildung*. Essen: LOGO-Lernspiel Verlag.

Gelman, R & Brennemann, K. (2004). Science learning pathways for young children. *Early Childhood Research Quarterly*, 19, 150-158.

Preissing, Ch. (2007). *Qualität im Situationsansatz. Qualitätskriterien und Materialien für die Qualitätsentwicklung in Kindertageseinrichtungen*. Berlin: Cornelsen Scriptor.

Steffensky, M., Lankes, E-M., Carstensen, C. & Nölke, Ch. (2012). Alltagssituationen und Experimente: Was sind geeignete naturwissenschaftliche Lerngelegenheiten für Kindergartenkinder? *Ergebnisse aus dem SNaKE-Projekt. Zeitschrift für Erziehungswissenschaft*, 15, 37-54.

Tietze, W., Meischner, T., Gänsfuß, R., Grenner, K., Schuster, K.-M., Völkel, P. & Roßbach, H.-G. (1998). *Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten*. Berlin: Luchterhand.

Wagenschein, M. (1997). *Kinder auf dem Wege zur Physik*. (2. Aufl.). Weinheim: Beltz.

Zimmermann, M. (2011). *Naturwissenschaftliche Bildung im Kindergarten: Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen. Studien zum Physik- und Chemielernen (Bd. 128)*. Berlin: Logos.



## ***(Zweit-)Sprachliche Interaktionen im Kindergarten: Eine qualitative Fallstudie***

Die Erforschung der sprachlichen Entwicklungsprozesse bei Kindern im Elementarbereich erfreut sich in Deutschland in der letzten Zeit eines großen Interesses. Ein besonderes Augenmerk gilt dabei den Kindern, die die deutsche Sprache sukzessiv (meistens mit dem Kindergarteneintritt) erwerben (vgl. Ahrenholz, 2012). Dennoch weiß man heute jenseits des Syntaxerwerbs noch relativ wenig über den frühen DaZ-Erwerb und es gibt bis dato nur wenige Studien, die den kindlichen DaZ-Erwerb in natürlichen Kommunikations- und Interaktionssituationen erforschen. Dabei kommt der natürlichen Umgebung im Zweitspracherwerb eine wichtige Rolle zu und vor allem Interaktionen mit pädagogischen Fachkräften und mit deutsch- und mehrsprachigen kindlichen Peers sind von entscheidender Bedeutung für den Erwerb einer zweitsprachlichen Kompetenz (vgl. Albers, 2009).

Der Beitrag fokussiert (zweit-)sprachliche Interaktionen im Kita-Alltag und geht der Frage nach, wie drei- bis vierjährige Kinder mit (noch) geringen Deutschkenntnissen ihre Kommunikations- und Interaktionsprozesse im Kindergarten sprachlich gestalten. Besonders interessant sind dabei auch die Fragen, welche Interaktions- und Kommunikationsstrategien sich bei Kindern mit DaZ in ihren Interaktionen mit Peers und mit ErzieherInnen rekonstruieren lassen und welche Reaktionen auf diese Strategien die jeweiligen Gesprächspartner zeigen.

Für die als qualitative Fallstudie konzipierte Untersuchung wurden fünf dreijährige Kinder mit Deutsch als Zweitsprache ausgewählt. Im Rahmen einer dreimonatigen ethnographischen Feldforschung im Kindergarten wurden alltägliche Interaktionen, an den die ausgewählten Kinder beteiligt waren, teilnehmend beobachtet und audiographisch aufgenommen. Diese Spontansprachaufnahmen bilden zusammen mit den im Feld erstellten Feldnotizen und Beobachtungsprotokollen, eingesetzten Kurzfragebögen sowie durchgeführten Feldgesprächen und Leitfadeninterviews mit Eltern und ErzieherInnen die Datenbasis der Studie. Ein solches methodisches Vorgehen gewährleistet einen möglichst umfassenden Blick auf die kindliche Zweitsprachentwicklung.

Eine gesprächsanalytische Auswertung der spontansprachlichen Daten sowie die Analyse der ethnographischen Daten bilden den methodischen Kern der Studie. Die Diskussion der Ergebnisse geschieht von dem Hintergrund der soziokulturellen Theorien des Zweitspracherwerbs, entwicklungspsychologischer Erkenntnisse zur Entwicklung sozial-kommunikativer Kompetenzen sowie der Erkenntnisse der Gesprochenen-Sprache-Forschung.

Ahrenholz, Bernt. (2012): Methodische Verfahren der Zweitspracherwerbsforschung - zur Einführung. In: Ahrenholz, B. (Hrsg.): Einblicke in die Zweitspracherwerbsforschung und ihre methodischen Verfahren. Berlin [u.a.]: de Gruyter, 1–26.

Albers, Timm (2009): Sprache und Interaktion im Kindergarten. Eine quantitativ-qualitative Analyse der sprachlichen und kommunikativen Kompetenzen von drei- bis sechsjährigen Kindern. Bad Heilbrunn: Klinkhardt.

## **Bildung und Region: Zur Bedeutung von Unterstützungsformaten am Studienanfang in Regionen mit heterogen Schullandschaften**

In diesem Beitrag wird der Frage nachgegangen, inwiefern Unterstützungsformate zu Studienbeginn das spätere Studium erfolgreich beeinflussen. Am Beispiel der Studienanfänger der Westfälischen Hochschule wird anhand einer quantitativen empirischen Analyse untersucht, welchen Effekt Unterstützungsangebote haben und welche Studierendengruppen hiervon besonders profitieren. Es besteht die Vermutung, dass Förderungsmaßnahmen vor Studienbeginn den Studienverlauf positiv beeinflussen und insbesondere Studienanfänger profitieren, die ihre Studienberechtigung nicht an Gymnasien erworben haben. Studienanfänger von Fachhochschulen kommen heute mit äußerst heterogenen Studienvoraussetzungen an die Hochschule. Hierbei gibt es jedoch starke regionale Unterschiede, je nach Struktur der regionalen Schullandschaft. So liegt der Anteil der Hochschulzugangsberechtigten, die ihren Schulabschluss nicht am Gymnasium erworben haben, im Ruhrgebiet höher als in den übrigen Regionen Nordrhein-Westfalens. Im Ruhrgebiet sind die Hochschulen eng mit der Region verknüpft (Kriegesmann & Kottmann, 2010) und der Großteil der Studierenden hat ihre Studienberechtigung in der Region erworben (IT NRW, 2012). Daher ist das regionale Umfeld für die Hochschulen entscheidend. Es gibt deutliche Hinweise darauf, dass sich die Wege der Studienanfänger an die Hochschulen im Ruhrgebiet erheblich von denjenigen in bildungsbürgerlich geprägten Regionen unterscheiden (Kriegesmann & Kottmann, 2010). Dieses gilt insbesondere für das nördliche Ruhrgebiet. Da mit unterschiedlichen schulischen Bildungswegen andere Studienvoraussetzungen einhergehen, stellt die Heterogenität der Studienanfänger zugleich eine Herausforderung für Hochschulen dar (Köller et al., 2004; Trautwein et al., 2007). Mitten im nördlichen Ruhrgebiet gelegen, hat die Westfälische Hochschule einen extrem hohen Anteil Studierender aus dem direkten Umkreis (Kottmann & Kriegesmann, 2011). Das Fächerspektrum der Westfälischen Hochschule weist ein technisch-ökonomisches Profil auf. Gerade in Fächergruppen wie Ingenieurwissenschaften sind umfassende Mathematikkenntnisse von grundlegender Bedeutung für ein erfolgreiches Studium. Da viele Studienanfänger keine ausreichenden Kenntnisse für ein solches Studium mitbringen, ist es nicht verwunderlich, dass in technischen Fächern die Studienabbruchraten am höchsten sind (Heublein et al., 2012). Die unterschiedlichen Mathematikkenntnisse scheinen insbesondere durch schulische Bildungsverläufe begründet zu sein (Trautwein et al., 2007; Henn & Polaczek, 2007). Um diesem Problem zu begegnen, wird an der Westfälischen Hochschule das Unterstützungsformat „Einstiegsakademie“ angeboten. Dieses ist eine zweiwöchige Blockveranstaltung vor Studienbeginn, die darauf abzielt, vorhandene Defizite der Studienanfänger abzubauen. Das Format der Einstiegsakademie unterscheidet sich von üblichen Vorkursen an anderen Hochschulen insofern, als dass diese zusätzlich neben klassischen Vorkursen für Bachelor-Studiengänge als besonderes Förderprogramm stattfindet. Durch Teilnahme an der Einstiegsakademie wird den Studienanfängern in kleinen Gruppen die Gelegenheit gegeben, Methodenkompetenz zum effizienteren Lernen zu erwerben. Es wird versucht, Wissenslücken der Studienanfänger zu schließen und fehlende Mathematikkenntnisse aufzuarbeiten. Für die Teilnahme an der Einstiegsakademie müssen sich die Studienanfänger bewerben. Die Auswahl findet auf Grundlage der Ergebnisse eines vorherigen Orientierungstests statt. Wie wichtig Unterstützungsmaßnahmen am Studienanfang sind, zeigt eine Auswertung des Anteils bestandener Prüfungen von Studienanfängern im Wintersemester 2012/13 im 1. Fachsemester zwischen Teilnehmern und Nichtteilnehmern der Einstiegsakademie (nur 1. Prüfungsversuch, N[Prüfungen von Teilnehmern]=786, N[Prüfungen von Nicht-Teilnehmern]=2728). Die Auswertungsergebnisse zeigen, dass sich die Teilnahme an diesem Unterstützungsangebot positiv auf das Abschneiden in relevanten Prüfungen auswirkt. Der Anteil bestandener Prüfungen liegt bei Teilnehmern der Einstiegsakademie um 21% höher, als bei Nicht-Teilnehmern. Dieses Resultat lässt sich in fast allen betrachteten Studienfächern feststellen. Es können sehr starke Effekte bei Studienanfängern, die ihre Studienberechtigung an Gesamtschulen erworben haben, nachgewiesen werden. Während bei Studienanfängern, die ihre Hochschulzugangsberechtigung am Gymnasium erworben haben, unabhängig vom Besuch der Einstiegsakademie der Anteil bestandener Prüfungen auf sehr hohem Niveau liegt, liegt der Anteil bestandener Prüfungen bei früheren Gesamtschülern bei 83,6% (Teilnehmer) bzw. bei 58,6% (Nicht-Teilnehmer). Erste multivariate Auswertungsergebnisse bestätigen die Bedeutsamkeit von Unterstützungsmaßnahmen vor Studienbeginn. Die Untersuchung von Effekten durch Unterstützungsformate für Studienanfänger mit heterogenen Studieneingangsvoraussetzungen liefert einen Beitrag zu regional-spezifischen Analysen des Studienverlaufs in der Studienanfangsphase. Die Erfahrungen aus dem nördlichen Ruhrgebiet können Impulse für andere Hochschulen in Regionen mit einer nicht bildungsbürgerlich geprägten Schullandschaft liefern.

Henn, Gudrun; Polaczek, Christa (2007): Studienerfolg in den Ingenieurwissenschaften. In: Das Hochschulwesen, 55 (2007) 5, S. 144-147.

Heublein, Ulrich; Richter, Johanna; Schmelzer, Robert; Sommer, Dieter (2012): Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen. Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010. HIS: Forum Hochschule 3/2012.

IT NRW (2012): Studierende im Wintersemester 2010/11 nach Bundesland und Kreis der Hochschul-zugangsberechtigung. Sonderauswertung.

Köller, Olaf; Baumert, Jürgen; Cortina, Kai S.; Trautwein, Ulrich; Watermann, Rainer (2004): Öffnung von Bildungswegen in der Sekundarstufe II und die Wahrung von Standards. Analysen am Beispiel der Englischleistungen von Oberstufenschülern an integrierten Gesamtschulen, beruflichen und allgemein bildenden Gymnasien, Zeitschrift für Pädagogik 50 (2004) 5, S. 679-700.

Kottmann, Marcus; Kriegesmann, Bernd (2011): Mit FH-INTEGRATIV Talente entfalten – Ein Programm an der FH Gelsenkirchen. In: Heinrich-Böll-Stiftung (Hg.): Öffnung der Hochschule – Chancengerechtigkeit, Diversität, Integration, S. 52-58.

Kriegesmann, Bernd; Kottmann, Marcus (2010): Talente stärken – nicht die sozial Starken. In: VDI-Nachrichten. Quelle: <http://www.ingenieur.de/Arbeit-Beruf/Ausbildung-Studium/Talente-staerken-sozial-Starken> (letzter Zugriff am 30.10.2013).

Trautwein, Ulrich; Lüdtke, Oliver; Lehmann, Rainer; Köller, Olaf (2007): Öffnung von Bildungswegen und Vergleichbarkeit von Leistungsniveaus: Zusammenfassung und Ausblick. In: Ulrich Trautwein (Hg.): Schulleistungen von Abiturienten. Regionale, schulformbezogene und soziale Disparitäten. Münster, New York, München, Berlin: Waxmann, S. 161-170.



## **Lehrveranstaltungsevaluation als wirksames Instrument zur Qualitätsverbesserung? Studie zu möglichen Wirkmechanismen.**

Problemstellung: Studentische Lehrveranstaltungsevaluationen haben sich in der hochschulischen Realität etabliert (vgl. Schmidt & Loßnitzer, 2010). Lange Zeit wurde vor allem die Frage der Validität kontrovers diskutiert (z.B. Aleamoni, 1999; Hoffmann, 2009; Marsh, 1987; Rindermann, 2001; Spooren et al., 2013). Dringlich scheint darüber hinaus die Frage nach der Wirksamkeit. Für die oft beklagte Folgenlosigkeit von Lehrveranstaltungsevaluationen lassen sich empirische Belege nennen (z. B. Böttcher & Hogrebe, 2010; Rindermann, 2001), die entsprechende Wirkungsforschung lässt sich jedoch in zweierlei Hinsicht kritisieren. Einerseits verwendet sie einen relativ undifferenzierten Wirkungs- bzw. Nutzungsbegriff, während die neuere Evaluationsforschung von dem breiteren Begriff des Einflusses von Evaluationen ausgeht (vgl. z.B. Kirkhart, 2000). Andererseits werden Wirkungserwartungen in Bezug auf Evaluationen theoretisch oft sehr verkürzt modelliert, obwohl auch hierzu v.a. mit der theory of evaluation influence (Henry & Mark, 2003) differenziertere Ansätze verfügbar wären. Ziel der vorliegenden Studie ist, unter Verwendung dieser theoretischen Perspektiven die Rezeptions- und Wirkungsprozesse studentischer LVE auf Seiten der Dozierenden eingehender zu analysieren. Vorgestellt wird eine qualitative Vorstudie, auf der weitere quantitative Arbeiten aufbauen sollen.

Theoretischer Hintergrund: Die Nutzungsforschung nimmt einen wesentlichen Platz im Feld der Evaluationsforschung ein (vgl. z.B. Henry & Mark, 2003; Hense et al., 2013; Patton, 1997; Weiss, 1998). Im internationalen Diskurs identifizierte man drei Formen von Evaluationsnutzung: instrumentelle, konzeptionelle und symbolische Nutzung (vgl. Hense & Widmer, 2013). Patton (1998) ergänzte die Diskussion mit seiner Unterscheidung zwischen product use und process use, in Erweiterung des vorherrschenden, engen Nutzungsbegriff führte Kirkhart (2000) den Terminus influence ein. Erkenntnisse aus den empirischen Arbeiten zur Wirkungsanalyse studentischer Lehrveranstaltungsevaluation zeigen, dass die Wirkung von der Rückmeldeform beeinflusst wird (vgl. z.B. Böttcher & Grewe, 2010). Weiterhin gibt es jedoch keine hinreichende Erklärung, welche Faktoren auf welcher Ebene einen Einfluss auf die Nutzungen von Evaluationen haben (vgl. Sandermann, 2013). Zu möglichen Mechanismen, durch die evaluationsbasierte Veränderungsprozesse auf individueller, interpersoneller und organisationaler Ebene ausgelöst werden, legen Henry und Mark (2003) einen theoretischen Entwurf vor. Die Autoren verstehen Evaluation als Intervention und identifizieren aus einer psychologischen Perspektive mögliche kognitive und affektive sowie motivationale und verhaltensbasierte Mechanismen, durch welche Evaluationen Einfluss nehmen können.

Fragestellung: Die vorliegende Studie untersucht die Frage, inwiefern sich die von Henry und Mark (2003) angenommenen Wirkmechanismen auf individueller Ebene (Elaboration, Salienz, Wissenserwerb, Priming, Heuristics und Einstellungsänderung) bei der Rezeption von Lehrveranstaltungsevaluationen identifizieren lassen. Für die Studie wurden diese durch die von Viebahn (2004) angenommenen Mechanismen (Selbstbild, institutioneller Stellenwert der Lehre, Gütekriterien, subjektives Kontrollempfinden) ergänzt.

Methode: Als methodologischer Rahmen der Untersuchung wurde ein explorativ-qualitatives Vorgehen gewählt, um Mechanismen auf der Mikroebene in der angemessenen Inhaltstiefe zu erfassen und eine inhaltsvalide Basis für weitere Forschungsvorhaben anzulegen. In leitfadengestützten Interviews wurden die Mechanismen in Form von jeweils einer Interviewfrage operationalisiert und anhand des aktuellen Rückmeldeverfahrens einer psychologischen Fakultät abgefragt. Um Vergleichbarkeit herzustellen, wurden Personen gleicher Hierarchieebene befragt, zudem war es Voraussetzung, dass die Befragten in den letzten 24 Monaten Teil der Lehrveranstaltungsevaluation und entsprechenden Rückmeldeverfahrens waren.

Ergebnisse: Die Wirkmechanismen lassen sich mit unterschiedlicher Häufigkeit empirisch beobachten. Während der Mechanismus der Elaboration eine starke, Priming, Selbstbild, der institutionelle Stellenwert der Lehre, das subjektive Kontrollempfinden und die instrumentelle Nutzung eine mittlere bis starke Tendenz erkennen ließen, konnten die Mechanismen Gütekriterien, Salienz und Heuristics lediglich in einer schwachen Ausprägung beobachtet werden. Wissenserwerb und Einstellungsänderung konnten nicht festgestellt werden. Es stellt sich die Frage, ob dies an dem Rückmeldeverfahren liegt oder ob gar andere Wirkmechanismen zum Tragen kommen.

Im Rahmen der Folgestudie liegt der Fokus auf quantitativ-messbaren Indikatoren. Basierend auf den Ergebnissen der Interviewstudie soll untersucht werden, ob sich repräsentative Aussagen ableiten lassen. Im weiteren Ausblick stellt sich die Frage, welche möglichen Implikationen daraus sowohl für die Nutzungsforschung als auch auf dem hochschulpolitischen Sektor entstehen können.

Aleamoni, L.M. (1999). Student Rating Myths Versus Research Facts from 1924 to 1998. In: Journal of Personnel Education, 13 (2), S. 153-166.

Böttcher W., Dicke J.N., Hogrebe N. (2010) (Hrsg.). Evaluation, Bildung und Gesellschaft. Steuerungsinstrumente zwischen Anspruch und Wirklichkeit. Münster u.a.: Waxmann.

Böttcher W., Grewe, M. (2010). Eine Untersuchung zur Wirksamkeit der studentischen Lehrveranstaltungskritik am Beispiel der Westfälischen Wilhelms-Universität. In: Pohlenz, P., Oppermann, A., Lehre und Studium professionell evaluieren: Wie viel Wissenschaft braucht die Evaluation, S. 73-82. Bielefeld: UniversitätsVerlagWebler.

Böttcher, W., Hense, J., Keune, M. (2013). Schulinspektion als eine Form externer Evaluation – ein Forschungsüberblick. In: Hense, J., Rädiker, S., Böttcher, W., Widmer, T. Forschung über Evaluation. Bedingungen, Prozesse und Wirkungen, S. 231-250. Münster u.a.: Waxmann.

Hense, J., Rädiker, S., Böttcher, W., Widmer, T. (2013) (Hrsg.). Forschung über Evaluation. Bedingungen, Prozesse und Wirkungen. Münster u.a.: Waxmann.

Hense, J., Widmer, T. (2013). Ein Überblick zum internationalen Stand der Forschung über Evaluation. In: Hense, J., Rädiker, S., Böttcher, W., Widmer, T. Forschung über Evaluation. Bedingungen, Prozesse und Wirkungen, S. 251-278. Münster u.a.: Waxmann.

Henry, G. T. & Mark, M. M. (2003). Toward an agenda for research on evaluation. New Directions for Evaluation, 97, 69–80.

Hoffmann, A. (2009). Studentische Lehrveranstaltungsevaluation – Ein Zerrbild der Qualität akademischer Lehre? In: Empirische Pädagogik, 23 (1), S. 19-49.

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. New Directions for Evaluation (88), 5–23.

Marsh, H.W. (1987). Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research. In: International Journal of Educational Research, 11 (3), 253–388.

Patton, M. Q. (1997). Utilization-focused evaluation. Thousand Oaks, Calif. (u.a.): Sage.

Patton, M. Q. (1998). Die Entdeckung des Prozessnutzens - Erwünschtes und unerwünschtes Lernen durch Evaluation. In M. Heiner (Hrsg.), Experimentierende Evaluation: Ansätze zur Entwicklung lernender Organisationen (S. 55-66). Weinheim (u.a.): Juventa Verlag.

Rindermann, H. (2001). Die studentische Beurteilung von Lehrveranstaltungen. Forschungsstand und Implikationen. In: Spiel, C. (Hrsg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck, S. 61-88. Münster u.a.: Waxmann.

Sandermann, M. (2013). Wie bedeutsam sind soziale Kompetenzen für Evaluationsnutzungen? In: Hense, J., Rädiker, S., Böttcher, W., Widmer, T. Forschung über Evaluation. Bedingungen, Prozesse und Wirkungen, S. 143-164. Münster u.a.: Waxmann.

Schmidt, B., Loßnitzer, T. (2010). Lehrveranstaltungsevaluationen: State of the Art, ein Definitionsvorschlag und Entwicklungslinien. In: Zeitschrift für Evaluation, 9 (1), S. 49-72.

Spooren, P., Brockx, B., Mortelmans, D. (2013). On The Validity of Student Evaluation of Teaching: The State of the Art. In: Review of Educational Research, 83 (4), 1-45.

Viebahn, P. (2004). Hochschullehrerpsychologie. Theorie- und empiriebasierte Praxisanregungen für die Hochschullehre. Bielefeld: UniversitätsVerlagWebler.

Weiss, C. H. (1998). Have we learned anything new about the use of evaluation. American Journal of Evaluation, 19 (1), 21–33.



Chair(s): Susanne Maria Weber (Philipps-Universität Marburg)

Diskutant/in: Anna Brake (Universität Augsburg)

Das hier vorgeschlagene Symposium versammelt Beiträge aus dem BMBF-Programm „Forschung für den wissenschaftlichen Nachwuchs“ (FOWIN), die vor dem Hintergrund unterschiedlicher disziplinärer Zugänge, theoretischer Grundlagen und methodischer Vorgehensweisen ein gemeinsames Phänomen in den Blick nehmen: Die Konstituierung von Nachwuchskarrieren in der Wissenschaft in Abhängigkeit von je spezifischen institutionellen Strukturen. Ziel des Symposiums ist es nicht, bereits konkrete Untersuchungsergebnisse zu diskutieren, sondern die unterschiedlichen theoretischen und methodologischen Zugangsweisen in ihrer Relevanz für die Hervorbringung empirischer Ergebnisse als produktive Perspektivenvielfalt auf den gemeinsamen Gegenstand herauszuarbeiten.

Im Zentrum des Symposiums steht die Frage, welche Bedeutung institutionelle Strukturen für das Zustandekommen wissenschaftlicher Nachwuchskarrieren erlangen. Den gemeinsamen Ausgangspunkt bildet dabei die Annahme, dass sich die Karriereverläufe von NachwuchswissenschaftlerInnen nicht allein durch individuelle Faktoren erklären lassen, sondern dass die vielfältigen institutionellen Faktoren empirisch in den Blick genommen und in ihrem Zusammenspiel mit den individuellen Faktoren analysiert werden müssen. Von dieser gemeinsamen Grundannahme ausgehend soll im Symposium die Vielfalt sich darauf beziehender Forschungszugänge zur Diskussion gestellt werden. Diesen Focus verdeutlicht eingangs Chair Susanne Maria Weber (Erziehungswissenschaft).

In ihrem Beitrag fragen Heiner Minssen und Caroline Richter (Arbeitswissenschaft) in einer arbeitswissenschaftlichen Perspektive nach der Wirkung der strukturellen Organisation von Vertrauen auf die Karriereentwicklung des wissenschaftlichen Nachwuchses. Den empirischen Zugang bildet dabei ein qualitatives Forschungsdesign. Aus einer prospektiv-längsschnittlichen und psychologischen Perspektive analysieren Simone Kauffeld und Daniel Spurk (Psychologie) Laufbahnverläufe des wissenschaftlichen Nachwuchses, um für deren Austritt, Verbleib und Rückkehr soziale und organisationale Faktoren zu identifizieren. Schließlich präsentieren Jörg Schwarz (Erziehungswissenschaft) und Julia Elven (Soziologie) ein qualitatives Forschungsdesign im interdisziplinären Verbund, mit dem aus einer erziehungswissenschaftlichen Perspektive die Gelegenheiten rekonstruiert werden, welche institutionelle Kontexte NachwuchswissenschaftlerInnen zur Aneignung wissenschaftsbezogener Wissensbestände, Fähigkeiten und grundlegender Haltungen bieten.

Durch diese unterschiedlichen disziplinären Perspektiven und theoretischen Konzepte ergeben sich nicht nur differente Auffassungen darüber, was in der Untersuchung wissenschaftlicher Nachwuchskarrieren empirisch in den Blick genommen werden muss; vielmehr muss auch diskutiert werden, wie die institutionelle Strukturierung von Karriereverläufen mit empirischen Untersuchungsmethoden zugänglich gemacht werden kann. Die Beiträge zeigen die Multiperspektivität methodischer Zugriffe auf: Sie reicht von großangelegten und z.T. längsschnittlichen quantitativen Untersuchungen bis hin zu qualitativen Verfahren. Vor dem Hintergrund dieser Vielfalt wird im Symposium der Frage nachgegangen, wie sich die institutionelle Strukturierung wissenschaftlicher Nachwuchskarrieren theoretisch konzipieren und methodisch erfassen lässt. Die in den Symposiumsbeiträgen entfalteten Antworten werden durch Anna Brake (Soziologie, empirische Sozialforschung) aus methodologischer und methodischer Sicht kommentiert und diskutiert.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer



## Die Organisation von Vertrauen

Im Verbundprojekt „Vertrauen und Wissenschaftlicher Nachwuchs“ widmet sich der Lehrstuhl für Arbeitsorganisation und Arbeitsgestaltung von Prof. Dr. Heiner Minssen am Institut für Arbeitswissenschaft der Ruhr-Universität Bochum dem Thema Vertrauen und Karriereentwicklungen von Wissenschaftlichem Nachwuchs aus einer arbeits- und organisationssoziologischen Perspektive. Die Verbundpartnerin an der Universität zu Köln untersucht das Erleben von Vertrauen aus bildungssoziologischer Perspektive. Das Teilprojekt mit dem Schwerpunkt der Organisation von Vertrauen verfolgt die Fragestellung: Welchen Einfluss hat organisational-institutionalisiertes Vertrauen auf die Karriereentwicklung von Wissenschaftlichem Nachwuchs innerhalb des Hochschulsystems aus Perspektive von WiN und der Organisation Hochschule? Die forschungsleitende Annahme lässt dahingehend konkretisieren, dass Vertrauen und die Vertrauenswürdigkeit von Hochschule aus Perspektive von Nachwuchswissenschaftler/innen während der Promotion, vorwiegend aber in der Phase nach der Promotion/vor der Berufung maßgeblich deren Karrieren beeinflusst. Diese Annahme resultiert vorrangig aus der Vertrauensforschung, die die Relevanz von Vertrauen unter Bedingungen von Ungewissheit und Risiken als organisationaler Integrations- und Handlungskordinationsmechanismus ebenso betont wie die wichtige Funktion von institutionalisiertem Misstrauen (Sztompka, 1999; Endreß, 2012) als impliziten Aspekt (z.T. als unbeabsichtigte Nebenfolge) der Nachwuchsförderung. Vor dem Hintergrund des beschriebenen Forschungsinteresses bieten die Dimensionen „Person/System“ (Giddens, 1994, Luhmann, 1988, 2001) und „reflexives/habituell Vertrauen“ (Endreß, 2002, 2012) einen fundierten Zugang, anhand derer das komplexe Phänomen in Bezug auf die Fragestellung empirisch und analytisch zugänglich wird. Dabei sucht das Teilprojekt Antworten auf folgende Fragen:

Welche Anforderungen an die Vertrauenswürdigkeit von Hochschule und ihre Strukturen, Institutionen und Praktiken werden aus Perspektive von Wissenschaftlichem Nachwuchs adressiert, der durch strukturierte Programme mehr oder weniger eng an das Hochschulsystem angebunden ist?

Wie sind hochschulpolitische Instrumente zur Nachwuchsförderung aus Sicht der mit Nachwuchsförderung konfrontierten WiN gestaltet? Wie müssten diese gestaltet sein, damit sich Vertrauen sowohl auf Seiten der etablierten Wissenschaftler/innen wie auch auf Seiten der Nachwuchswissenschaftler/innen generieren kann?

Welche Bedeutung besitzt das Vertrauen der akademischen Lehrer/innen für die Nachwuchswissenschaftler/innen innerhalb von strukturierten Nachwuchsprogrammen, wie wird es wahrgenommen und interpretiert, wie wirkt es sich auf Selbstvertrauen, Handlungs- und Konkurrenzfähigkeit aus?

Im Fokus der Untersuchung steht die Organisation von Vertrauen a) aus Perspektive von Wissenschaftlichem Nachwuchs mit/nach Erfahrung innerhalb strukturierter Nachwuchsprogramme und b) aus Perspektive von Expert/innen des Hochschulsystems.

Für die Perspektive der Nachwuchswissenschaftler/innen liegen inhaltlich bislang keine belastbaren theoretischen Überlegungen oder empirischen Befunde vor, um das Untersuchungsfeld ex ante zu strukturieren; aktuelle Ergebnisse aus der allgemeinen Arbeitsforschung sind auf die besondere Situation der Hochschulen und der spezifischen Anforderungen der hier interessierenden Zielgruppe nur unzureichend übertragbar. Daher stellt das Forschungsdesign in einem ersten Schritt auf das Vertrauen aus Organisations- und Hochschulperspektive ab. Um diese Perspektive nachvollziehen zu können, werden 10 leitfadengestützte Expert/inneninterviews (Meuser & Nagel, 1991) mit Mitgliedern von Förderinstitutionen und Hochschulmanagement und berufungserfahrenen Professor/innen durchgeführt. Ziel ist es, Hinweise auf Zuschreibungen von Vertrauenswürdigkeit gegenüber Nachwuchs und gegenüber institutionnalisierter Nachwuchsförderung zu identifizieren, um weiterführende Untersuchungskategorien und Fragekonstruktionen festzulegen. Aufgrund der gewonnenen Erkenntnisse kann dann in einem zweiten Schritt die im Fokus des Vorhabens stehende Zielgruppe befragt werden. Um Aussagen zum Phänomen Vertrauen aus Perspektive von Wissenschaftlichem Nachwuchs zu Vertrauenskonzeptionen sowie Einschätzungen zu Hochschulkarriere, Vertrauen und Nachwuchsförderung zu erheben, sind 20 bildungs- und berufsbiographische Interviews mit erzählgenerierenden Fragen (Hopf, 1991) geplant. Für die Perspektive der Organisation von Vertrauen sind besonders Nachwuchswissenschaftler/innen unterschiedlicher Qualifikationsstufen aus ausgewählten Fachdisziplinen von Interesse, die über Erfahrung mit institutionalisierter Nachwuchsförderung verfügen. Als Ergebnis wird ein Deutungsrahmen vorliegen, der auf Aussagen zu Genese, Herstellung, Erscheinungsweisen, Objekten und Funktionen von Vertrauen/Misstrauen fokussiert werden kann, um den Einfluss von Vertrauen/Misstrauen auf Karrierebedingungen, -entwicklungen und -verläufe von und für WiN innerhalb der Hochschule zu verdeutlichen; die Analyse von Zuschreibungen an den Verbleib an Hochschule und an Exit-Optionen in hochschulunabhängige Beschäftigungsfelder knüpft unmittelbar daran.

Endreß, M. (2002): Vertrauen. Bielefeld.

Endreß, M. (2012): Vertrauen und Misstrauen. Soziologische Überlegungen, in: Schilcher, C. et al. (Hg.): Vertrauen und Kooperation in der Arbeitswelt, Wiesbaden: Springer VS: 81-102.

Giddens, A. (1994): Risk, trust, reflexivity; in: Beck, U.; Giddens, A.; Lash, S. (Eds.): Reflexive Modernization. Politics, Tradition and Aesthetics in the Modern Social Order: Stanford (Ca): Stanford University Press: 184-197.

Hopf, C. (1991): Qualitative Interviews in der Sozialforschung; in: Flick, U. et al. (Hg.): Handbuch qualitativer Sozialforschung, München: 177-182.

Luhmann, N. (2001/1988): Vertraulichkeit, Zuversicht, Vertrauen. Probleme und Alternativen; in: Hartmann, M.; Offe, C. (Hg.): Vertrauen. Die Grundlage des sozialen Zusammenhalts. Frankfurt/Main, New York: Campus: 143-160.

Meuser, M.; Nagel, U. (1991): ExpertInneninterviews – vielfach erprobt, wenig bedacht. Ein Beitrag zur qualitativen Methodendiskussion; in: Garz, D.; Kraimer, K. (Hg.): Qualitativ-empirische Sozialforschung, Opladen: 441-471.

Sztompka, P. (1999): Trust: A Sociological Theory. Cambridge.



## ***Quo vadis Post-Doc: Professur, Wirtschaft oder Sackgasse? Individuelle, soziale und organisationale Faktoren für die Laufbahnausrichtung und -entwicklung des wissenschaftlichen Nachwuchses (Pro-Wi+)***

Ziel des Projektes ist es Beschreibung, Analyse und Erklärung von Laufbahnverläufen des wissenschaftlichen Nachwuchses in Deutschland aus einer multifaktoriellen, prospektiv-längsschnittlichen, multi-methodalen und psychologischen Perspektive. Dabei sollen individuelle (z.B. Einstellungen, Kompetenzen), soziale (z.B. Netzwerke, Führung) und organisationale (z.B. Weiterbildungsmaßnahmen, Institutsmerkmale) Faktoren identifiziert werden, die a) den Austritt aus vs. den Verbleib im Wissenschaftssystem begünstigen, b) die Rückkehr in das Wissenschaftssystem begünstigen, nachdem bereits eine andere Tätigkeit in einem anderen Bereich verfolgt wurde, und c) erfolgreiche wissenschaftliche Laufbahnen prognostizieren können. Zusätzlich werden geschlechts- und fächerspezifische Einflüsse untersucht. Im Rahmen des Symposiums werden vornehmlich organisationale aber auch institutionell förderbare soziale Prädiktoren thematisiert. Bei organisationalen Faktoren sind zum einen objektive Variablen und zum anderen die subjektive Wahrnehmung dieser objektiven Variablen aus psychologischer Perspektive wichtig. Es zeigte sich beispielsweise, dass bereits die Teilnahme an Weiterbildungsmaßnahmen und das Prestige einer Arbeitseinheit für Nachwuchswissenschaftler wichtig sind (Judge, 2004; Louis et al., 2007; Hadani et al., 2012). In dem Projekt soll genauer untersucht werden ob und in welchem Umfang Weiterbildungsmaßnahmen von der Universität bzw. vom Lehrstuhl angeboten und letztlich auch von Post-Docs in Anspruch genommen werden. Neben der Abfrage, ob an bestimmten Maßnahmen teilgenommen wurde (Kauffeld, 2010), sollen insbesondere die Wahrnehmung von Mentoring- und Coaching-Prozessen im Sinne einer karriereförderlichen bzw. psycho-sozialen Unterstützung im wissenschaftlichen Kontext genauer untersucht werden (Blickle, Witzki, & Schneider, 2009; Kauffeld & Montasem, 2009). Weitere Faktoren, die untersucht werden sollen sind zum Beispiel die Größe der Arbeitseinheit, der Publikationsoutput und das Drittmittelvolumen des Lehrstuhls, das Organisationsklima, die Art und Träger der bewilligten Projekte als auch die Stellenstruktur (z.B. Juniorprofessur vs. Drittmittelstelle vs. Lehrstuhlstelle) auf der die Post-Doc Phase verbracht wird. Zudem sind insbesondere soziale Netzwerke ein entscheidender Karrierefaktor, der institutionell gefördert werden sollte und intensiv in dem Projekt untersucht wird.

Die primäre Zielgruppe des Projekts sind Personen, welche gerade promoviert haben und sich daraufhin im Wissenschaftssystem oder an dessen Schnittstelle bewegen. Diese Zielgruppe soll erstens quantitativ mittels einer großen, prospektiven Befragung drei Mal in einem Zeitraum von drei Jahren befragt werden (Projektbereich A). Zweitens sollen in einer Querschnittsuntersuchung Verhaltensbeobachtungen via Videoanalysen in Teammeetings sowie qualitative Interviews durchgeführt werden (Projektbereich B). Die Ergebnisse aus Projektbereich B werden zeitnah in Projektbereich A integriert. Zudem sollen Personen aus einem anderen BMBF-Projekt (ProNet+) in die Erhebung eingebunden werden. Da der Fokus des Projekts auf Laufbahnverläufen liegt, werden in der Längsschnittuntersuchung parametrische und nicht-parametrische Analysen durchgeführt, mit denen sich optimal zeitliche Abfolgen und Veränderungsprozesse abbilden lassen (Duncan, Duncan, & Strycker, 2006; Kline, 2005). Hierzu zählen beispielsweise latente Wachstumskurven- oder Survivalanalysen. Zudem werden bevorzugt konkurrierende Prädiktoren gleichzeitig als auch mögliche vermittelnde Mechanismen mittels indirekter Effektanalyse analysiert. Bei Daten auf Team- oder Abteilungsebene werden Mehrebenenmodelle berücksichtigt. Konkrete Merkmale der Tätigkeit, Anreiz- und Motivationsfaktoren sowie subjektiv bedeutsame Erfolgskriterien des wissenschaftlichen Nachwuchses werden mittels Interviews qualitativ untersucht (Mayring, 2010). Team- und Führungsprozesse als weitere institutionelle bzw. kontextuelle Ressourcen der Nachwuchswissenschaftler werden über Verhaltensanalysen in Teammeetings gemessen (Kauffeld, 2006). Die Ergebnisse dieses Projekts sollen sowohl für das Management von Forschungsabteilungen und Universitäten, als auch für bildungspolitische Entscheidungsträger, und für die individuelle Laufbahnplanung des wissenschaftlichen Nachwuchses verwertbar gemacht werden. Die Projektergebnisse können eine gute Grundlage für die Systematisierung unterschiedlicher Laufbahnverläufe darstellen und sollen unter anderem als Grundlage für Fort- und Weiterbildungsprogramme für den wissenschaftlichen Nachwuchs genutzt werden. So sollen strukturell erfolgreiche, gesunde und attraktive wissenschaftliche Laufbahnen unterstützt und aktiv begleitet werden.

Blickle, G., Witzki, A. H. & Schneider, P. B. (2009). Mentoring support and power. A three year predictive field study on protégé networking and career success. *Journal of Vocational Behavior*, 74, 181–189.

Duncan, T. E., Duncan, S. C. & Strycker, L. A. (2006). *An introduction to latent growth curve modeling - Concepts, issues, and applications*. London: Lawrence Erlbaum Associates.

Hadani, M., Coombes, S., Das, D. & Jalajas, D. (2012). Finding a good job: Academic network centrality and early occupational outcomes in management academia. *Journal of Organizational Behavior*, 33, 723–739.

Judge, T. A., Kammeyer-Mueller, J. & Bretz, R. D. (2004). A longitudinal model of sponsorship and career success: A study of Industrial-Organizational Psychologists. *Personnel Psychology*, 57, 271–303.

Kauffeld, S. (2006). *Kompetenzen messen, bewerten, entwickeln. Ein prozessanalytischer Ansatz für Gruppen*. Stuttgart: Schäffer-Poeschel.

Kauffeld, S. & Montasem, K. (2009). *Kompetenzmodell als Basis. Professionelle Video-Analyse im Coaching*. *Coaching-Magazin*, 4, 44–49.

Kauffeld, S. (2010). Das Kasseler-Kompetenz-Raster. In W. Sarges, H. Wottawa & C. Roos (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren*. Band II: Organisationspsychologische Instrumente (S. 165–173). Pabst Science Publisher.

Kauffeld, S. (2010). *Nachhaltige Weiterbildung. Betriebliche Seminare und Trainings entwickeln, Erfolge messen, Transfer sichern*. Berlin: Springer Verlag.

Louis, K. S., Holdsworth, J. M., Anderson, M. S. & Campbell, E. G. (2007). Becoming a scientist: The effects of WorkGroup size and organizational climate. *Journal of Higher Education*, 78, 311–336.

Kline, R. (2005). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (11. Aufl.). Weinheim: Beltz.



## **Passungsverhältnisse zwischen habituellen und institutionellen Strukturen im Kontext wissenschaftlicher Nachwuchskarrieren**

Der Beitrag diskutiert theoretische und methodische Zugangsweisen zur Untersuchung der Denk-, Wahrnehmungs- und Handlungsmuster von NachwuchswissenschaftlerInnen, welche in je spezifischem Verhältnis zu wissenschaftlichen bzw. universitären Institutionen stehen und somit verschiedene Möglichkeiten, aber auch Grenzen hinsichtlich der weiteren wissenschaftlichen Laufbahn herstellen. Dabei greift der Beitrag auf erste Ergebnisse des Verbundprojekts „Trajektorien im akademischen Feld – Ermöglichungskontexte und Entwicklungsverläufe wissenschaftlicher Nachwuchskarrieren“ zurück. Dieses Forschungsprojekt untersucht die Karrieren von NachwuchswissenschaftlerInnen als langfristige und pfadabhängige soziobiographische Verläufe („Trajektorien“) unter Bezugnahme auf Pierre Bourdieus Praxistheorie (Bourdieu, 1987, 1993, 2002) und basiert entsprechend auf der Annahme, dass Karriereverläufe im Zusammenspiel von institutionellen Strukturen und den Denk-, Wahrnehmungs- und Handlungsstrukturen der NachwuchswissenschaftlerInnen hervorgebracht werden. Vor diesem Hintergrund gliedert sich der Forschungsverbund in zwei Teilprojekte, welche die frühe Post-Doc-Phase mit einem multimethodischen qualitativen Forschungszugang aus verschiedenen Perspektiven untersuchen: Während das an der Philipps-Universität Marburg angesiedelte Teilprojekt „Institutionelle Ermöglichungskontexte wissenschaftlicher Nachwuchskarrieren“ unterschiedliche institutionelle Konstellationen untersucht, die den NachwuchswissenschaftlerInnen als Arbeits-, Lern- und Förderstrukturen begegnen, fokussiert das an der Universität Augsburg angesiedelte Teilprojekt „Habitus als Ermöglichungsstruktur wissenschaftlicher Nachwuchskarrieren“ die Rekonstruktion habitusverankerter Dispositionen der Promovierten. Die Habitus von NachwuchswissenschaftlerInnen bedingen die Art und Weise der alltäglichen wissenschaftlichen Arbeit, etwa des Verfassens von Texten, der Diskutierens, der Zeitplanung oder der Darstellung von Forschungsergebnissen. Dieser Modus wissenschaftlichen Denkens, Wahrnehmens und Handelns muss anschlussfähig sein an die in der Wissenschaft vorfindlichen institutionellen Kontexte – beispielsweise Tagungen, Stiftungen, Promotionskollegs, aber auch spezifischen Arbeitsformen wie Projekt- oder Gremienarbeit – damit sich die NachwuchswissenschaftlerInnen erfolgreich etablieren können. Der Beitrag wendet nun zunächst den Blick auf die theoretische Rekonstruktion der wissenschaftsbezogenen Dispositionen und auf die Passungen von habituellen Dispositionen und wissenschaftlichen Institutionen und kann hier bereits an Erkenntnisse aus der praxeologisch orientierten Hochschulforschung anknüpfen (vgl. Engler, 2000, 2001; Beaufays, 2003; Kraus, 2008). Hieran anschließend werden die im Projekt gewählten Forschungszugänge zur Analyse jener wissenschaftsbezogenen Dispositionen und habituell-institutionellen Passungsverhältnisse vorgestellt (vgl. Brake et al., 2013). Das empirische Vorgehen stützt sich insbesondere auf zwei methodische Elemente: Einen zentralen Zugang bilden 36 leitfadengestützte Einzelinterviews mit jüngst promovierten NachwuchswissenschaftlerInnen. Um wissenschafts- und wissenschaftskarriererelevante habituelle Muster aufspüren zu können, sind die Interviews so angelegt, dass sie Erzählungen zur wissenschaftlichen Alltagspraxis evozieren (Brake, 2006) und dadurch in besonderem Maße auf Denk-, Wahrnehmungs- und Handlungsstrukturen schließen lassen. Um einerseits eine große Reichweite, gleichzeitig aber auch Spezifität (Hopf, 1978) der Schilderungen wissenschaftlicher Alltagspraxis zu erreichen ist der Leitfaden offen angelegt, wobei dennoch zentrale Themenbereiche wissenschaftlicher Karrieren aufgegriffen werden, die eine themenbezogene Vergleichbarkeit ermöglichen. Die Interviews zielen auf Praxisschilderungen und Haltungen der NachwuchswissenschaftlerInnen und können Aufschluss darüber geben, inwieweit jene habitualisierten Dispositionen angeeignet bzw. konsolidiert wurden, die entlang der korrespondierenden institutionalisierten Anerkennungs- und Zuschreibungsprozesse ihre Zugehörigkeit zum wissenschaftlichen Feld begründen. Ein zweiter empirischer Zugang ermöglicht die Annäherung an diese Frage durch die Analyse visuellen Datenmaterials und fokussiert dabei die körperliche Manifestation wissenschaftsbezogener Habitusstrukturen. Hierzu werden die NachwuchswissenschaftlerInnen gebeten, sich innerhalb 3-5 Minuten als WissenschaftlerIn bei laufender Kamera vorzustellen (Fünf-Minuten-Selbstpräsentation, FMSP). Die Analyse des videoaufgezeichneten Materials richtet sich nicht nur auf den Inhalt des Gesagten, sondern vor allem darauf, inwieweit ein überzeugendes „mastering the informal aspects of cultural capital“ (Hansen & Mastekaasa, 2006) erkennbar wird: Wie glaubhaft gelingt es den NachwuchswissenschaftlerInnen mit ihrer gesamten Person einschließlich ihres körpersprachlichen Ausdrucks wissenschaftliche Autorität zu verkörpern? Aufbauend auf der theoretischen und empirischen Darlegung des Forschungsprojekts einschließlich erster Ergebnisse sollen Möglichkeiten diskutiert werden, gewonnene Erkenntnisse über die habituell verankerten, wissenschaftliche Alltagspraxis hervorbringenden Dispositionen auf die institutionellen Kontexte der Wissenschaft zu beziehen. Dabei sind beispielsweise Fragen nach Passungsverhältnissen zwischen Habitus und Institutionen oder nach den Formen institutionell beförderter wissenschaftlicher Sozialisation bedeutsam.

Beaufays, Sandra (2003): Wie werden Wissenschaftler gemacht? Beobachtungen zur wechselseitigen Konstitution von Geschlecht und Wissenschaft. Bielefeld: Transcript

Bourdieu, Pierre (1987): Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft. Frankfurt a. M.: Suhrkamp

Bourdieu, Pierre (1993): Sozialer Sinn. Kritik der theoretischen Vernunft. Frankfurt a. M.: Suhrkamp

Bourdieu, Pierre (2002): Homo academicus. Frankfurt a. M.: Suhrkamp

Brake, Anna; Bremer, Helmut; Lange-Vester, Andrea (Hg.) (2013): Empirisch arbeiten mit Bourdieu. Weinheim: Juventa Verlag

Brake, Anna (2006): Der Bildungsort Familie. Methodische Grundlagen der Untersuchung. In: Peter Büchner und Anna Brake (Hg.): Bildungsort Familie. Transmission von Bildung und Kultur im Alltag von Mehrgenerationenfamilien. Wiesbaden: VS, S. 49–79

Engler, Steffani (2000): Zum Selbstverständnis von Professoren und der Illusion des wissenschaftlichen Feldes. In: Beate Kraus (Hg.): Wissenschaftskultur und Geschlechterordnung. Über die verborgenen Mechanismen männlicher Dominanz in der akademischen Welt. Frankfurt a. M.: Campus, S. 121–151

Engler, Steffani (2001): "In Einsamkeit und Freiheit"? Zur Konstruktion der wissenschaftlichen Persönlichkeit auf dem Weg zur Professur. Konstanz: UVK

Hansen, Marianne Nordli; Mastekaasa, Arne (2006): Social Origins and Academic Performance at University. In: European Sociological Review 22 (3), S. 277–291

Hopf, Christel (1978): Die Pseudo-Exploration - Überlegungen zur Technik qualitativer Interviews in der Sozialforschung. In: Zeitschrift für Soziologie 7 (2), S. 97–115

Kraus, Beate (2008): Wissenschaft als Lebensform: Die alltagspraktische Seite akademischer Karrieren. In: Yvonne Haffner und Beate Kraus (Hg.): Arbeit als Lebensform? Beruflicher Erfolg, private Lebensführung und Chancengleichheit in akademischen Berufsfeldern. Frankfurt a. M., New York: Campus, S. 177–211



## ***Institutionelle Ermöglichungskontexte wissenschaftlicher Nachwuchskarrieren***

Der Beitrag diskutiert theoretische und methodische Zugangsweisen zur Untersuchung der institutionellen Strukturierung wissenschaftlicher Nachwuchskarrieren anhand erster Ergebnisse aus dem Verbundprojekt „Trajektorien im akademischen Feld – Ermöglichungskontexte und Entwicklungsverläufe wissenschaftlicher Nachwuchskarrieren“. Dieses Forschungsprojekt untersucht die Karrieren von NachwuchswissenschaftlerInnen als langfristige und pfadabhängige soziobiographische Verläufe („Trajektorien“) unter Bezugnahme auf Pierre Bourdieus Praxistheorie (Bourdieu, 1993). Dabei wird davon ausgegangen, dass Karriereverläufe in einem Zusammenspiel von institutionellen Strukturen mit den Denk-, Wahrnehmungs- und Handlungsstrukturen der NachwuchswissenschaftlerInnen hervorgebracht werden. Vor diesem Hintergrund gliedert sich der Forschungsverbund in zwei Teilprojekte, welche die frühe Post-Doc-Phase mit einem multimethodischen qualitativen Forschungszugang aus verschiedenen Perspektiven untersuchen: Während das an der Universität Augsburg angesiedelte Teilprojekt „Habitus als Ermöglichungsstruktur wissenschaftlicher Nachwuchskarrieren“ auf die Rekonstruktion habitusverankerter Dispositionen von NachwuchswissenschaftlerInnen zielt, untersucht das an der Philipps-Universität Marburg angesiedelte Teilprojekt „Institutionelle Ermöglichungskontexte wissenschaftlicher Nachwuchskarrieren“ die unterschiedlichen institutionellen Konstellationen, die ihnen z.B. als Arbeits-, Lern- und Förderstrukturen begegnen. Der Beitrag fokussiert zunächst die Frage, wie solche institutionellen Strukturen mit dem theoretischen Instrumentarium der bourdieuschen Praxistheorie analytisch gefasst werden können. Von zentraler Bedeutung ist dabei die Grundannahme, dass sich soziale Praxis im Zusammenwirken von subjektiven und objektiven Strukturen konstituiert. Während der Begriff des Habitus die subjektiven Strukturen umfasst, werden die objektiven Strukturen bei Bourdieu analytisch dem Feldbegriff (Bourdieu, 2004) zugeordnet: Soziale Felder stellen spezifische Praxisbereiche dar, die zwar einerseits in einem immerwährenden praktischen Herstellungsprozess begriffen sind, andererseits den AkteurInnen aber als objektiv gegebene und relativ dauerhafte Strukturen gegenüberreten. Diese Strukturen umfassen nicht nur die jeweiligen historisch gewachsenen Akteurskonstellationen als spezifische Sozialstruktur des Feldes, sondern ebenso die im Feld etablierten Selbstverständlichkeiten und Spielregeln sowie die geteilte Überzeugung vom Wert der im Feld umkämpften Gegenstände. An diese Theorieperspektive schließen sich unmittelbar erziehungswissenschaftliche Fragen an (Friebertshäuser, Rieger-Ladich, & Wigger, 2009). Denn mit dem Eintritt in ein spezifisches Feld eröffnen sich den AkteurInnen vielfältige Lerngelegenheiten, die zugleich meist auch einen Aufforderungscharakter gegenüber den NovizInnen entfalten: Um dauerhaft an der Praxis eines sozialen Feldes beteiligt sein zu können oder gar mächtige Positionen in diesem Feld zu besetzen ist es unerlässlich, sich die dort herrschenden Spielregeln anzueignen. Empirisch herauszuarbeiten, welche Gelegenheiten die institutionellen Strukturen des wissenschaftlichen Feldes (Bourdieu, 1992) den NachwuchswissenschaftlerInnen zur Aneignung von wissenschaftsbezogenen Wissensbeständen, Fähigkeiten und grundlegenden Haltungen über ihre fachliche Qualifikation hinaus bieten und welche Aneignungsleistungen sie gleichzeitig einfordern ist das Ziel des Projekts. Für diese Zielstellung wurde ein qualitatives Forschungsdesign gewählt, das in seiner Methodologie rekonstruktiv angelegt ist und sich an Bourdieus Praxistheorie orientiert (Brake, Bremer, & Lange-Vester, 2013). Dabei wird die frühe Post-Doc-Phase als neuralgischer Punkt der wissenschaftlichen Laufbahn fokussiert und in prozessualer Perspektive durch die Rekonstruktion der Promotionsphase erweitert. In die Analyse werden systematisch drei Promotionsmodi und zwei differente Forschungsfelder einbezogen. Für zwölf Kernfälle der durch das Augsburger Teilprojekt untersuchten NachwuchswissenschaftlerInnen erhebt das Marburger Teilprojekt die institutionellen Kontexte über drei Zugriffsebenen, mit denen sich jeweils ein spezifischer methodischer Zugang verbindet. Auf der Ebene der Organisation gilt es die formalen, programmatischen und praktischen Strukturierungen des Promotionsprozesses über Dokumentenanalysen zu rekonstruieren. Nach der Ausgestaltung der Betreuungspraxis wird auf der Ebene der Betreuenden gefragt, weiterhin wie darin ein spezifisches Betreuungsverständnis zum Ausdruck kommt, welche Auffassungen von Wissenschaft und wissenschaftlicher Karriere dem zu Grunde liegen und von welcher Position im Feld aus die Betreuenden agieren. Um diese Fragen zu untersuchen, werden teilstrukturierte Interviews geführt. Schließlich werden bezüglich der Peers Form und Funktion der Austauschbeziehungen mittels Gruppendiskussionen ebenso analysiert wie geteilte Erfahrungshorizonte.

Ausgehend von der Annahme, dass sich eine erfolgreiche Nachwuchskarriere im prozessualen Zusammenspiel von individuellen und institutionellen Faktoren herstellt, erfolgt die Rekonstruktion der institutionellen Ermöglichungskontexte durch gemeinsame Analysen eng verzahnt mit dem Teilprojekt „Habitus als Ermöglichungsstruktur wissenschaftlicher Nachwuchskarrieren“ der Universität Augsburg. Es werden erste Ergebnisse dieser multiperspektivischen und multimethodischen empirischen Untersuchung vorgestellt.

Bourdieu, Pierre (2004): Über einige Eigenschaften von Feldern. In: Pierre Bourdieu: Soziologische Fragen. Frankfurt a.M.: Suhrkamp, S. 107–114.

Bourdieu, Pierre (1993): Sozialer Sinn. Kritik der theoretischen Vernunft. Frankfurt a.M.: Suhrkamp.

Bourdieu, Pierre (1992): Homo academicus. Frankfurt a.M.: Suhrkamp.

Brake, Anna; Bremer, Helmut; Lange-Vester, Andrea (2013): Empirisch arbeiten mit Bourdieu. Theoretische und methodische Überlegungen Konzeptionen und Erfahrungen. Weinheim, Basel: Juventa.

Friebertshäuser, Barbara; Rieger-Ladich, Markus; Wigger, Lothar (Hg.) (2009): Reflexive Erziehungswissenschaft. Forschungsperspektiven im Anschluss an Pierre Bourdieu. 2., durchgesehene und erweiterte Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.



**Soziale Unterstützung in der Promotionsphase – Analyse von Netzwerkkarten**

In der Promotionsphase stellen soziale Unterstützungsbeziehungen neben fachlichen und methodischen Kenntnissen wichtige Ressourcen für die Qualifizierung des wissenschaftlichen Nachwuchses dar. Der Zugang zu diesen Ressourcen und ihre Nutzung werden durch soziale Beziehungen vermittelt. Dieses ‚Sozialkapital‘ beschreibt Bourdieu (1983) als theoretisches Konstrukt zur Analyse der Reproduktion sozialer Ungleichheiten. Demnach können persönliche Kontakte zum Informationsaustausch, zur Nutzung impliziten Wissens und zur emotionalen Unterstützung genutzt werden. Bourdieu leitet daraus die Hypothese ab, dass der Umfang des Netzwerks positiv mit den daraus resultierenden Erträgen (z.B. Anerkennung, Unterstützung, Prestige) zusammenhängt. Dies konnten Hennig und Kohl (2011) empirisch durch soziale Netzwerkanalysen von Sozialkapital nachweisen. Gleichzeitig bilden beruflich-soziale Netzwerke der Promotionsphase den sozialen Kontext, in welchem situiertes Lernen (Lave & Wenger, 1991) stattfindet. Dieser soziale Kontext beeinflusst die Handlungsmöglichkeiten von Akteur/innen und umgekehrt, weshalb Informationsaustausch, Wissensvermittlung und Lernprozesse während der Promotionsphase durch das soziale Netzwerk und das Lernumfeld unterstützt werden können (Rehrl, Palonen & Gruber 2006, 2010). Die Unterstützung der Promovierenden und ihre fachliche und methodische Qualifizierung für die empirische Bildungsforschung wurden im letzten Jahrzehnt oftmals bemängelt (Huber, 2012; Merkt 2008; Weishaupt et al., 2008). In der vorgestellten Studie werden soziale Unterstützungsbeziehungen in der Promotionsphase analysiert, um die Bedeutung des Sozialkapitals für wissenschaftliche Karrieren einzuschätzen. Im Fokus des Beitrags steht die Forschungsfrage nach den Typen von Unterstützungsnetzwerken der Promovierenden. Je nach Netzwerktypus werden unterschiedliche Erträge des Sozialkapitals gewonnen und fördern die Qualifizierung in der Promotionsphase bzw. die berufliche Entwicklung stärker oder schwächer. Die empirische sozialwissenschaftliche Methode der sozialen Netzwerkanalyse (SNA) (Jansen, 2006; Schnegg & Lang, 2002) verwendet relationale Daten als zentrale Analyseeinheiten. Die qualitative SNA eignet sich, um Deutungsmuster von Akteur/innen zu erfassen, damit die Wechselwirkungen zwischen Einschätzungen und Handlungen ersichtlich werden (Hollstein, 2006).

Basierend auf der Population der quantitativen SNA-Studie (Kenk, 2012; Kenk eingereicht) über wissenschaftliche Netzwerke aus Publikations-, Forschungs- und Betreuungsrelationen im Zeitraum von 2000 bis 2010 wurden für die qualitative SNA-Studie 16 ehemalige Promovierende eines interdisziplinären Forschungsverbands der Bildungsforschung ausgewählt. Das theoretische Sampling stellt die Merkmalsverteilung bzgl. Disziplin, Geschlecht und erreichter beruflicher Position sicher. Für die egozentrierte SNA wurden die persönlichen Beziehungen in der computergestützten persönlichen Befragung mithilfe von Netzwerkkarten (Software Vennmaker, Gamper & Kronenvett, 2012) erhoben. Die Leitfragen der Befragung lauteten: - „Welche Personen haben Sie während Ihrer Promotionsphase unterstützt?“ - „Wie wichtig war die Person für Sie während der Promotionsphase?“ - „Welche Art von Unterstützung haben Sie geleistet bzw. erhalten? Neuigkeiten, Fachliches, oder Rat?“

Das Datenmaterial umfasst Informationen über die Befragten, ihre Netzwerkpersonen (z.B. Geschlecht, Disziplin, berufliche Position) und ihre Beziehungen, grafische Abbildungen und Filmsequenzen der Netzwerkkarten sowie Transkriptionen aufgezeichneter Gespräche. Für diesen Beitrag werden quantitative Daten der Netzwerkkarten mit Fokus auf die persönlichen Beziehungen pseudonymisiert ausgewertet. Die Typisierung der Fälle erfolgt die zwei Dimensionen der ‚Unterstützung‘ (instrumentelle, kognitive, emotionale Hilfe) (Diewald, 1991) sowie des ‚Lernumfelds‘ (Beziehungsqualität zu Promotionsbetreuungspersonen sowie zur peer-group). Die Ausprägungen beider Merkmale bilden das Typentableau, das als Basis der Typencharakterisierung dient. Erste Auswertungen zeigen zwei kontrastive Typen: kleinere Netzwerke mit starken Beziehungen zu Promotionsbetreuer/innen gegenüber größeren Netzwerken mit intensiven peer-group-Beziehungen. Zur Präsentation der Ergebnisse wird als drittes Merkmal die ‚Interdisziplinarität‘ herangezogen, um die Verbindung disziplinärer Perspektiven in den sozialen Netzwerken aufzuzeigen. Zwar sind Einschränkungen der Studie zu nennen: Die Ergebnisse sind aufgrund der spezifischen Population und des theoretischen Samplings nicht repräsentativ und somit nicht übertragbar auf beispielsweise zunehmend verbreitete strukturierte Promotionsprogramme (Herz & Korff, 2013). Dennoch liefert die Studie wichtige Hinweise auf nützliche Unterstützungsbeziehungen aus einer rekonstruktiven Perspektive einer älteren Promovierendengeneration. Die Relevanz der Studie für die empirische Bildungsforschung liegt im Fokus auf eine tertiäre Lernphase sowie auf interdisziplinäre Zusammenarbeit in Forschungsprojekten. Die Erkenntnisse zur Unterstützung in der Promotionsphase können zur Verbesserung der Situation des wissenschaftlichen Nachwuchses anregen.

Bourdieu, Pierre (1983): Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In: Reinhard Kreckel (Hg.): Soziale Ungleichheiten. Göttingen: Schwartz, S. 183–198.

Diewald, Martin (1991): Soziale Beziehungen: Verlust oder Liberalisierung? Soziale Unterstützung in informellen Netzwerken. Berlin: Edition Sigma.

Gamper, Markus; Kronenvett, Michael (2012): Visuelle Erhebung von egozentrierten Netzwerken mit Hilfe digitaler Netzwerkkarten. In: Sabrina Kulin, Keno Frank, Detlef Fickermann und Knut Schwippert (Hg.): Soziale Netzwerkanalyse. Theorie, Methoden, Praxis. Münster: Waxmann, S. 151–166.

Hennig, Marina; Kohl, Steffen (Hg.) (2011): Rahmen und Spielräume sozialer Beziehungen. Zum Einfluss des Habitus auf die Herausbildung von Netzwerkstrukturen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Herz, Andreas; Korff, Svea (2013): Promovieren in Programmen sturkturierter Promotion aus Sicht der AdressatInnen - Ergebnisse der standardisierten Online-Befragten. In: Svea Korff und Navina Roman (Hg.): Promovieren nach Plan? Wiesbaden: Springer Fachmedien Wiesbaden, S. 75–116.

Hollstein, Betina (2006): Qualitative Methoden und Netzwerkanalyse - ein Widerspruch? In: Betina Hollstein und Florian Straus (Hg.): Qualitative Netzwerkanalyse. Konzepte, Methoden, Anwendungen. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 11–35.

Huber, Nathalie (Hg.) (2012): Der Dokortitel zwischen Status und Qualifikation. Bonn: iFQ (iFQ-Working Paper, 12).

Jansen, Dorothea (2006): Einführung in die Netzwerkanalyse. Grundlagen, Methoden, Forschungsbeispiele. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kenk, Martina (2012): Publikations- und Forschungsnetzwerke von Nachwuchswissenschaftlerinnen und -wissenschaftlern in der empirischen Bildungsforschung. In: Sabrina Kulin, Keno Frank, Detlef Fickermann und Knut Schwippert (Hg.): Soziale Netzwerkanalyse. Theorie, Methoden, Praxis. Münster: Waxmann, S. 267–278. Online verfügbar unter [http://www.pedocs.de/volltexte/2013/5769/pdf/Kenk\\_Publikations\\_und\\_Forschungsnetzwerke\\_D\\_A\\_.pdf](http://www.pedocs.de/volltexte/2013/5769/pdf/Kenk_Publikations_und_Forschungsnetzwerke_D_A_.pdf).

Kenk, M. (eingereicht, 2014): Wissenschaftliche Netzwerke in der Bildungsforschung. In: Kolleck, Nina, Zander, Lysann, Hannover, Bettina (Hg.) Social Network Analysis in Educational Research and Educational Policy“ Sonderheft der Zeitschrift für Erziehungswissenschaft.

Lave, Jean; Wenger, Etienne (1991): Situated Learning. Legitimate peripheral participation. Cambridge: Cambridge University Press.

Merkt, Marianne (2008): Die Entwicklung der Doktorandenausbildung in Deutschland und daraus entstehende, zukünftige Aufgaben für hochschuldidaktische Zentren. In: Erziehungswissenschaft 19 (36), S. 57–76.

Rehrl, Monika; Palonen, Tuire; Gruber, Hans (2006): Expertise development in science. In: H. P. A. Boshuizen (Hg.): Lifelong learning of professionals: Exploring implications of a transitional labour market. Proceedings of the 3rd EARLI SIG Professional Learning and Development Conference. EARLI SIG Professional Learning and Development Conference. Heerlen. Open University of the Netherlands. Heerlen: Open University of the Netherlands.

Rehrl, Monika; Palonen, Tuire; Gruber, Hans (2010): Professionelle Vernetzung in der Wissenschaft. In: Unterrichtswissenschaft 38 (4), S. 334–353.

Schnegg, Michael; Lang, Hartmut (2002): Netzwerkanalyse. Eine praxisorientierte Einführung. In: Methoden der Ethnographie (1).

Weishaupt, H.; Kraul, M.; Böhm-Kasper, C.; Schulzeck, U.; Zügenrucker, I. (2008): Zur Situation der Bildungsforschung in Deutschland. Bonn: BMBF.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer





## **Die Bedeutung von Selbsterkundungsverfahren bei der Studienwahl: Welche Effekte haben sie zu welchem Zeitpunkt des Entscheidungsprozesses?**

In den letzten Jahren sind zahlreich sogenannte Selbsterkundungsverfahren (SEV) veröffentlicht worden. Diese sollen Schülerinnen und Schüler (SuS) der Sekundarstufe II bei der Studien- und Berufswahl unterstützen, indem sie zu der Frage Stellung nehmen, welche Studiengänge bzw. Berufe zu den Kompetenzen und Interessen der SuS passen (Kanning, 2013). Viele der entsprechenden Tests orientieren sich in ihrem Aufbau an der Passungstheorie von Holland (1997). Darüber hinaus können die Selbsterkundungsverfahren von Universitäten als sogenannte fachspezifische Studierfähigkeitstests zur Ergänzung für die Studierendenauswahl genutzt werden (gemäß 7. Novelle des HRG; Bundestag, 2004; Klusmann, Köller, & Kunter, 2011). In beiden Einsatzbereichen wird das Ziel verfolgt die Passung zwischen Merkmalen der Studierenden und Anforderungen des Studiengangs bzw. Berufs zu erhöhen und somit einen Wechsel oder Abbruch des Studiums zu vermeiden. Die aktuelle Studie untersucht als erstes die Frage, nach dem Effekt der SEV auf die Selbstwirksamkeit bei der Studien- bzw. Berufsentscheidung. Zur Beantwortung dieser Frage wurden in einer größeren quasiexperimentellen Studie zwei fachübergreifende Selbsterkundungsverfahren, der EXPLORIX® (Joerin Fux, Stoll, Bergmann, & Eder, 2012) und „was-studiere-ich.de“ (Hell, Päßler, & Schuler, 2009) mit einer standardisierten allgemeinen Studienberatung verglichen. Die ausgewählten Selbsteinschätzungsverfahren erfassen die Ausprägungen auf den sechs Dimensionen des beruflichen Interesses (Holland, 1997) und listen anschließend dazu passende Studienfächer und/oder Berufe auf. Die allgemeine Studienberatung wurde von Psychologiestudentinnen im Klassenverband durchgeführt und umfasste die Themen der Informationssuche, mögliche Wege nach dem Abitur (Freiwilligendienste, Auslandsaufenthalte, Ausbildung, Studium und -fächer) und einen Zeitplan für die Termine ein Jahr vor dem Erwerb der allgemeinen Hochschulreife. Die Selbstwirksamkeit der Studien- bzw. Berufsentscheidung wurde mit den Skalen der Career Decision Self-Efficacy Scale – Short Form (CDSE-SF; Betz, Klein, & Taylor, 1996) erfasst. Eine Erhöhung des Gesamtwertes der CDSE-SF bedeutet inhaltlich einen Anstieg der subjektiven Kompetenzerwartung für den Bereich der Studien- und Berufswahl. Die zweite Fragestellung der Studie ist, inwieweit die Effekte der SEV von dem Level der beruflichen Entscheidung, den kognitiven berufsspezifischen Charakteristika (Klarheit über sich selbst, Wissen über Berufe, Entscheidungsfähigkeit und Wichtigkeit der Berufsentscheidung) und der affektiven Komponente, wie wohl man sich mit dem aktuellen Stand der Berufswahl fühlt, beeinflusst werden. Der Level der beruflichen Entscheidung wurde mit der ersten Dimension der Entschiedenheit des Career Decision Profile (CDP; Jones, 1989) erhoben. Die kognitiven auf die berufliche Entscheidung bezogenen Charakteristika wurden mit der dritten Dimension und die affektive Komponente mit der zweiten Dimension des CDP erfasst. Diese Studie fand als Prä-Post-Vergleich im Sommer 2013 in den zehnten und elften Jahrgängen an 10 Gymnasien in Schleswig-Holstein statt. Insgesamt nahmen 529 SuS (54% weiblich) teil. Den SuS wurde zwischen der Bearbeitung der Prä- und Post-Fragebögen randomisiert eines von vier Verfahren (CCT – für diese Studie nicht relevant, „was-studiere-ich.de“, EXPLORIX® oder eine allgemeine Studienberatung) zugeteilt. Die SEV wurden einzeln am PC bearbeitet, die Studienberatung fand für die zugeteilten SuS gemeinsam in einem separaten Raum statt. Bezogen auf die erste Fragestellung deuten erste ANOVA-Analysen darauf hin, dass sich hinsichtlich der Selbstwirksamkeit bzgl. der Studien- bzw. Berufsentscheidung keine signifikanten Unterschiede zwischen den Verfahren ergeben. Für die zweite Fragestellung zeigen sich nach ersten ANCOVA-Analysen auch unter Kontrolle des aktuellen Entscheidungslevels und der kognitiven und affektiven Entscheidungskomponenten bzw. –charakteristika (Kovariaten) keine signifikanten Unterschiede hinsichtlich der Effekte der Selbsterkundungsverfahren. Diese ersten Ergebnisse lassen zusammenfassend vermuten, dass Gymnasiasten der Sekundarstufe II von studien- und berufswahlunterstützenden Verfahren im Allgemeinen zu profitieren scheinen, egal in welcher Phase des beruflichen Entscheidungsprozesses sie sich aktuell befinden.

- Betz, N. E., Klein, K. L., & Taylor, K. M. (1996). Evaluation of a short form of the Career Decision-Making Self Efficacy Scale. *Journal of Career Assessment*, 4, 47-57.
- Bundestag (2004). Siebtes Gesetz zur Änderung des Hochschulrahmengesetzes (7. HRGÄndG). 2298 Bundesgesetzblatt Jahrgang 2004 Teil I Nr. 47. Bonn.
- Griepentrog, M. (2001). Qual der Wahl oder Perspektivlosigkeit? Berufswahl und Berufsberatung in der dritten industriellen Revolution. In J. Mausel, W. Schweins & M. Ulbrich-Herrmann (Hrsg.), *Zukunftsperspektiven Jugendlicher - wirtschaftliche und soziale Entwicklung als Herausforderung und Bedrohung für die Lebensplanung* (S. 117–128). Weinheim und München: Juventa.
- Heine, C., Willich, J., & Schneider, H. (2010). Informationsverhalten und Entscheidungsfindung bei der Studien- und Berufswahl. *Studienberechtigte 2008 ein halbes Jahr vor dem Erwerb der Hochschulreife*. Hannover: Hochschul-Informations-System.
- Hell, B., Päßler, K., & Schuler, H. (2009). "was-studiere-ich.de" - Konzept, Nutzen und Anwendungsmöglichkeiten. *Zeitschrift für Beratung und Studium*, 4 (1), 9–14.
- Helmken, K. (2008). Studentische Schülerberatung - Evaluation einer Praxisforschung. *Psychologie in Erziehung und Unterricht*, 55, 61–67.
- Hirschi, A., & Läge, D. (2007). The Relation of Secondary Students' Career-Choice Readiness to a Six-Phase Model of Career Decision Making. *Journal of Career Development*, 34 (2), 164–191.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3. Aufl.). Eaglewood Cliffs, NJ: Prentice Hall.
- Joerin Fux, S., Stoll, F., Bergmann, C., & Eder, F. (2012). EXPLORIX - Das Werkzeug zur Berufswahl und Laufbahnplanung. Bern: Huber.
- Jones, L. K. (1989). Measuring a Three-Dimensional Construct of Career Indecision Among College Students: A Revision of the Vocational Decision Scale - The Career Decision Profile. *Journal of Counseling Psychology*, 36 (4), 477–486.
- Kanning, U. P. (2013). Berufsorientierungstests. In T. Brüggemann & S. Rahn (Hrsg.), *Berufsorientierung. Ein Lehr- und Arbeitsbuch* (S. 236–250). Ascheberg: Waxmann.
- Klusmann, U., Köller, M., & Kunter, M. (2011). Anmerkungen zur Validität eignungsdiagnostischer Verfahren bei angehenden Lehrkräften. *Zeitschrift für Pädagogik*, 57 (5), 711-721.



## Warum und wozu studiere ich mein Fach? Entwurf eines integrativen Motivationskonzepts zur Studienfachwahl

Einleitung: Was motiviert die Studienfachwahl? Vor dem Hintergrund aktueller Studienabbruchquoten um 28% geraten Motivationsfaktoren der Studienfachwahl verstärkt in den Fokus der wissenschaftlichen Auseinandersetzung (z.B. Heublein et al., 2012). In diesem Poster wird ein Modell vorgestellt, das mehrere motivationstheoretische Konzepte integriert und die Grundlage für die Entwicklung und Validierung eines Instrumentariums zur Studienfachwahlmotivation bildet.

Theoretischer Hintergrund: Ausgangspunkt des hier zugrunde gelegten theoretischen Rahmens bilden die Überlegungen von Alfred Schütz, der Handeln als Prozess menschlichen Verhaltens, der einem vorgefertigten Entwurf folgt, erklärt (Schütz, 1951, 1971). Nach Schütz beeinflussen zwei zeitliche Dimensionen den Entwurf als motivationales Verhalten: vergangene Erfahrungen und zukünftige Ziele. Erstere determinieren den Handlungsentwurf selbst – sie sind der objektive Grund einer Handlung und lassen sich als Weil-Motive fassen. Zweitere werden Um-zu-Motive genannt. Sie verweisen in die Zukunft und sind das Ziel, das durch ein Handeln erreicht werden soll. Das hier zugrunde gelegte Rahmenmodell erweitert diese Überlegung um Grundannahmen des Achievement-Goal Ansatzes (z.B. Harackiewicz, J. et al., 2000) und des Erwartungs-Wert-Ansatzes (z.B. Eccles & Wigfield, 1995). Der hier verfolgte Ansatz unterteilt die verschiedenen Konzepte einerseits danach, ob sie zukunftsorientiert (Um-zu-Motive) oder vergangenheitsbasiert (Weil-Motive) sind (Dimension A) und andererseits danach, ob sie ‚fremdbezogen‘, also ohne direkten eigenen Einfluss des Akteurs auf das Ziel und mit externem Bezug oder ‚selbstbasiert‘, d.h. mit direktem Einfluss auf das Ziel, sind (Dimension B). Zudem unterscheiden wir zwischen mittelfristigen (universitäts-bezogen) und langfristigen Um-zu-Motiven (berufsbezogen).

<http://www.myimg.de/?img=KonzeptStudienfachwahld67c0.jpg>

Abbildung 1: Konzept der Studienfachmotivation

Forschungsfrage: Lassen sich die theoretischen Differenzierungen zwischen Weil- und Um-zu-Motiven bzw. zwischen selbstbasierten und fremdbezogenen Motiven und die jeweils integrierten Konstrukte empirisch bestätigen?

Methode: Das oben skizzierte Modell bildete die Operationalisierungsgrundlage für die Entwicklung eines neuen Instrumentes zur Studienfachwahlmotivation. Die Itementwicklung basierte auf der sprachlichen Differenzierung zwischen Weil- und Um-Zu-Motiven nach Schütz, z.B. ‚Ich habe mein Studienfach gewählt,[...] um später einen stressfreien Beruf zu haben oder [...] weil es meinen Fähigkeiten entspricht‘. Das hierbei entstandene Fragebogeninstrument besteht aus 8 Skalen mit insgesamt 55 Items. Die Validierungsstudie wurde mit Studierenden im ersten Semester an verschiedenen staatlichen Hochschulen in Nordrhein-Westfalen durchgeführt (N=1518).

Erhebungszeitraum ist vom 01.08.2013-30.11.2013.

Ergebnisse: In einer exploratorischen Faktorenanalyse (Hauptkomponentenanalyse, Rotation: Oblimin) wurden die Items auf 38 reduziert und 6 Faktoren ermittelt (erklärte Gesamtvarianz: 58,2%). KMO-Test und MSA-Werte weisen auf eine gute Stichproben- und Itemeignung hin. Die Faktoren spiegeln das theoretisch entworfene Modell größtenteils wider. Dimension A (Weil-Motive vs. Um-Zu-Motive) sowie Dimension B (‚fremdbezogen‘ vs. ‚selbstbasiert‘) finden sich nahezu eindeutig in den Faktoren wieder. Der Faktor Leistungs- und Erfolgsorientierung im Studium (1) verbindet Annäherungs-Leistungs-Ziele mit studienbezogenen Nutzen-Werten. Einen gemeinsamen Faktor (2) bilden die Weil-Motive ‚Interesse‘ und ‚Fähigkeits-Selbstkonzept‘. Der Faktor Einkommen und Aufstieg (3) bildet die berufsbezogene Nutzen-Komponente ab. Der Arbeitsvermeidungs-Faktor (4) hebt die Trennlinie Studium und Beruf auf. Ebenfalls eindeutig extrahiert sich der Faktor Altruismus (5) mit Berufsbezug. Soziale Einflussfaktoren als Weil-Motive (6) bilden ebenfalls einen eigenen Faktor.

Die formalen Einflussfaktoren (z.B. NC-Einschränkungen) wurden über die Datenreduzierung vollständig aus dem Modell eliminiert.

Die interne Konsistenz der extrahierten Faktoren ist gut bzw. sehr gut (Cronbachs Alpha: .81 - .89).

Weitere erste Analysen zeigen signifikante Unterschiede in der Faktorenstruktur der Studienfächer. Die Bildung eindeutiger Motivationsprofile für die Studienfächer bestätigt den Einfluss der ermittelten Faktoren auf die Studienfachwahl. Die Ergebnisse legen nahe, dass das entwickelte Instrumentarium geeignet ist, um Studienmotivation im Sinne der theoretischen Annahmen zu erfassen. Die integrative Perspektive auf die Studienfachwahl erweist sich demnach fruchtbar.

Ausblick: Es sind tiefergehende Analysen geplant (u.a. konfirmatorische Faktorenanalysen, weitere Stichproben, latente Strukturgleichungsmodellierungen), u.a. um die extrahierte Faktorenstruktur zu manifestieren. Langfristig soll der Zusammenhang zwischen der Studienfachwahlmotivation und Studienzufriedenheit bzw. -abbruch tendenz untersucht werden.

Eccles, J.S., & Wigfield, A. (1995). The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs. *Personality And Social Psychology Bulletin*, Vol. 21, 215-225.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology*, Vol. 92 (2), 316-330.

Heublein, U., Richter, F., Schmelzer, R., & Sommer, D. (2012). Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen. *HIS Forum Hochschule*, 3/2012.

Schütz, A. (1951). Chosing among projects of action. *Philosophy and Phenomenological Research*, Vol. 12 (2), 161-184.

Schütz, A. (1971). *Gesammelte Aufsätze Bd. 1. Das Problem der sozialen Wirklichkeit*, Den Haag: Martinus Nijhoff.

Das Netzwerk, das im Dezember 2013 gegründet wurde, setzt an der Einzigartigkeit der Berliner Region an, in der mehrere Forschungseinrichtungen angesiedelt sind, die sich mit Themen der Bildungsforschung befassen. Ziel des Netzwerkes ist es, Nachwuchswissenschaftler/-innen in Berlin und Umgebung die Gelegenheit zu bieten, sich regelmäßig über ihre Einrichtung und Disziplin hinaus, zu treffen und auszutauschen. Um Bildungsprozesse umfassend zu erforschen, ist der Austausch zwischen den Disziplinen gerade in der Bildungsforschung von hoher Bedeutung. An der Freien Universität Berlin, der Technischen Universität und der Humboldt Universität zu Berlin, der Hertie School of Governance, der Universität Potsdam sowie zahlreichen Forschungseinrichtungen, welche vielfach der Leibniz-Gemeinschaft oder der Max-Planck-Gesellschaft zugehörig sind, wird Bildungsforschung betrieben. Mit diesen Einrichtungen sind für die Bildungsforschung zentrale Disziplinen, wie die Erziehungswissenschaft, Psychologie, Ökonomie und Soziologie sowie die bildungsgeschichtliche Forschung im Netzwerk vertreten. An all diesen Einrichtungen arbeiten Nachwuchswissenschaftler/-innen auf unterschiedlichen Ebenen an Themen der Bildungsforschung.

Nachwuchswissenschaftler/-innen aus unterschiedlichen Disziplinen nachhaltig vernetzen

Das Akronym des Netzwerkes ist durchaus mit bedacht gewählt, denn es geht darum, dass sich eine Gemeinschaft von NachwuchswissenschaftlerInnen im Bereich der Bildungsforschung vernetzt und wie der Bien damit Fähigkeiten entwickelt, die ein einzelner Nachwuchswissenschaftler nicht beherrscht. Das Netzwerk setzt sich zum Ziel, Nachwuchswissenschaftler/-innen aus unterschiedlichen Disziplinen systematisch miteinander zu vernetzen. Die räumliche und teilweise auch institutionelle Nähe in Berlin und Umgebung ist ein Standortvorteil des Netzwerkes. Die Vielzahl der beteiligten Einrichtungen und Disziplinen macht deutlich, dass im Netzwerk Themen der Bildungsforschung über den gesamten Lebensverlauf behandelt werden. So bietet das Netzwerk die Chance, sich als Nachwuchswissenschaftler/-in über die eigene Disziplin hinweg und den eigenen Forschungsfokus hinaus zu informieren und weiter zu qualifizieren. Zentrale Aktivitäten des Netzwerkes sind das monatlich stattfindende Lunchtime Seminar. Darüber hinaus sind pro Jahr eine Jahrestagung und zwei Methodenworkshops geplant, die auch einen Beitrag darstellen, um die fachlichen und methodischen Kenntnisse der beteiligten Nachwuchswissenschaftler/-innen weiter zu entwickeln.

Organisation und Koordination

Zum Aufbau des BIEN-Netzwerkes ist am DIW Berlin eine Koordinatorenstelle angesiedelt. Hier werden die Veranstaltungen des Netzwerkes organisiert und stattfinden. Renommierete Bildungsforscher/-innen an den unterschiedlichen Forschungseinrichtungen agieren als Schirmdamen und Schirmherren für das Netzwerk. Auf der Ebene der Nachwuchswissenschaftler/-innen gibt es an allen beteiligten Einrichtungen eine Ansprechperson für das Netzwerk.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer

Miriam Schaffer (Johannes Gutenberg-Universität Mainz), Christiane Kuhn (Johannes Gutenberg-Universität Mainz), Susanne Schmidt (Johannes Gutenberg-Universität Mainz), Olga Zlatkin-Troitschanskaia (Johannes Gutenberg-Universität Mainz), Sebastian Brückner (Johannes Gutenberg-Universität Mainz), Corinna Lautenbach (Humboldt-Universität zu Berlin), Sigrid Blömeke (Humboldt-Universität zu Berlin)

## **Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor (KoKoHs) – Herausforderungen eines interdisziplinär angelegten BMBF-Forschungsprogramms**

Die aktuellen Differenzierungsprozesse im Hochschulsystem finden im Kontext einer Steuerungspolitik statt, welche im Anschluss an die internationale Bildungspolitik den output bzw. outcome von Lernprozessen an Hochschulen fokussiert. Für nachhaltige Entwicklungs- und Optimierungsmaßnahmen im zunehmend differenzierteren Hochschulsystem sind evidenzbasierte Aussagen auf der Basis von Wirksamkeitsprüfungen erforderlich, die Hinweise auf individueller, hochschuldidaktischer und struktureller Ebene liefern können. Die Erfahrungen aus den aktuellen Umstrukturierungs- und Umbruchprozessen (wie die Umstellung auf das Bachelor-Master-Modell) sowie internationale Studien (vgl. z. B. die OECD-Machbarkeitsstudie AHELO, Tremblay, 2011) zeigen, dass die Entwicklung und Realisierung von erfolgsversprechenden Konzepten in der Hochschule einer theoretischen und empirischen Fundierung bedürfen. Nicht zuletzt die jüngst veröffentlichten Befunde der OECD-Studie PIAAC zur Ausprägung von grundlegenden Kompetenzen bei Erwachsenen können als aktuellen Hinweis dafür gesehen werden, dass die Untersuchung von Kompetenzen über den bislang fokussierten Schulbereich hinaus von hoher Brisanz ist (Rammstedt, 2013). Um die damit verbundenen Herausforderungen speziell im Hochschulbereich auf nationaler Ebene systematisch angehen zu können und international anschlussfähig zu sein, wurde das BMBF-Forschungsprogramm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor (KoKoHs)“ initiiert ([www.kompetenzen-im-hochschulsektor.de](http://www.kompetenzen-im-hochschulsektor.de)). Im interdisziplinär sowie methodenintegrativ angelegten Forschungsprogramm KoKoHs sind 23 Forschungsverbände vereint, die mit rund 70 Einzelprojekten in vierzehn Bundesländern vertreten sind. Sie alle verbindet das übergeordnete Ziel, zentrale Impulse zur grundlagenorientierten, akademischen Kompetenzforschung in Deutschland zu liefern und dabei an die internationale Forschung anzuknüpfen. Im Detail wird in jedem Projektverbund den Fragen nachgegangen, (1) wie generische und domänenspezifische Kompetenzen von Studierenden in unterschiedlichen Studienfachbereichen modelliert werden können und (2) wie geeignete Verfahren zur Erfassung dieser Kompetenzen entwickelt und validiert werden können. Neben dem Bereich der akademischen Forschungskompetenzen und selbstregulativen Fähigkeiten (als bedeutsame fachübergreifende Kompetenzen) werden die Studienfachbereiche Bildungswissenschaften, Ingenieurwissenschaften, Lehrerbildung für MINT-Fächer sowie Wirtschafts- und Sozialwissenschaften verstärkt in den Blick genommen. Die Projektverbände stellen sich dabei der Herausforderung, die zunehmend differenzierteren Bedingungen des Hochschulsystems (u. a. durch die inter- und intranationale Vielfalt an Studienmodellen, Strukturen und Lehrpraktiken) umfassend zu berücksichtigen. Auch die Frage nach einem geeigneten bzw. angemessenen Kriterium (z. B. zukünftige Berufsanforderungen), anhand dessen der Kompetenzerwerb eingeschätzt werden kann, stellt eine Herausforderung dar. Die systematische Verbindung der unterschiedlichen Perspektiven (bedingt durch die Vielfalt der fokussierten Studien(fach)bereiche sowie durch weitere verbundene Disziplinen wie insbesondere Fachdidaktiken, Psychologie und Statistik) und die damit einhergehende standortübergreifende und internationale Vernetzung der Projekte sind durch die Einrichtung einer Koordinierungsstelle mit Standorten in Berlin und Mainz über einen vierjährigen Zeitraum (2011-2015) gewährleistet. Ein besonderes Anliegen ist dabei, den wissenschaftlichen Nachwuchs nachhaltig zu fördern sowie die gewonnenen internationalen Kontakte in Form eines zukunftssträchtigen disziplinenübergreifenden kommunikativen Austauschs für das nationale Forschungsprogramm weiter zu nutzen und auszubauen. Im Rahmen der Tagung sollen theoretische und methodische Entwicklungsansätze zur Modellierung und Erfassung von akademisch erworbenen Kompetenzen aus Sicht der verschiedenen Studien(fach)bereiche aufgezeigt und hinsichtlich der Implikationen in einem zunehmend differenzierteren Hochschulsystem diskutiert werden. Ein Schwerpunkt soll dabei auf den Herausforderungen liegen, die mit der Zusammenführung der im KoKoHs-Programm aufeinander treffenden unterschiedlichen Perspektiven einhergehen, sowie auf Ansätzen, wie diese Herausforderungen im Rahmen des Programms systematisch angegangen werden.

Blömeke, S., Zlatkin-Troitschanskaia, O., Kuhn, C. & Fege, J. (Eds.) (2013). Modeling and Measuring Competencies in Higher Education. Rotterdam: Sense Publishers.

Rammstedt, B. (Hrsg.) (2013). Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012. Münster: Waxmann. Online verfügbar unter [http://www.gesis.org/fileadmin/piaac/Downloadbereich/PIAAC\\_Ebook.pdf](http://www.gesis.org/fileadmin/piaac/Downloadbereich/PIAAC_Ebook.pdf) am 16.10.2013.

Tremblay, K. (2013). OECD Assessment of Higher Education Learning Outcomes (AHELO): Rationale, Challenges and Initial Insights from the Feasibility Study. In S. Blömeke, O. Zlatkin-Troitschanskaia, C. Kuhn & J. Fege (Hrsg.) Modeling and Measuring Competencies in Higher Education (S. 113-126). Rotterdam: Sense Publishers.

Zlatkin-Troitschanskaia, O. & Kuhn, C. (2010). Messung akademisch vermittelter Fertigkeiten und Kenntnisse von Studierenden bzw. Hochschulabsolventen – Analyse zum Forschungsstand. Johannes Gutenberg-Universität Mainz: Arbeitspapiere Wirtschaftspädagogik, 56.

Ying Zhang (Universität des Saarlandes), Eric Klopp (Universität des Saarlandes), Heike Dietrich (Universität Heidelberg), Roland Brünken (Universität des Saarlandes), Ulrike-Marie Krause (Universität Oldenburg), Birgit Spinath (Universität Heidelberg), Robin Stark (Universität des Saarlandes), Frank M. Spinath (Universität des Saarlandes)

## **Zur Untersuchung inkrementeller Validität von Gewissenhaftigkeit und Leistungsmotivation über kognitive Prädiktoren hinaus: in der Prognose des Studienerfolgs bei Psychologiestudenten**

Die Erklärung und Vorhersage von Studienerfolg im Hochschulkontext ist eine aktuelle Forschungsfrage sowohl in der pädagogischen als auch in der differentiellen Psychologie. Sowohl Abiturnote als auch Intelligenz sind als gewichtige Prädiktoren zur Vorhersage des Studienerfolgs bekannt (Kappe & van der Flier, 2012; Steyer et al., 2005).

Gewissenhaftigkeit und Leistungsmotivation leisten ebenfalls einen wichtigen Erklärungsbeitrag zum Studienerfolg (Busato et al., 2000; Chamorro-Premuzic & Furnham, 2003; Farsides & Woodfield, 2003; Kappe & van der Flier, 2012; Richardson et al., 2012) wenngleich die Effektstärken hier geringer ausfallen. Das Ziel dieser Studie besteht darin, zunächst die prädiktive Validität von Abiturnote und Intelligenz für Studienerfolg zu überprüfen und in einem weiteren Schritt die inkrementelle Validität von Gewissenhaftigkeit und Leistungsmotivation zu untersuchen. Es wurden folgende Hypothesen aufgestellt:

H1: Abiturnote und Intelligenz weisen prädiktive Validität für die Vorhersage des Studienerfolgs auf.

H2: Gewissenhaftigkeit weist über die Abiturnote und Intelligenz hinaus inkrementelle Validität für die Vorhersage des Studienerfolgs auf.

H3: Leistungsmotivation weist über die Abiturnote und Intelligenz hinaus inkrementelle Validität für die Vorhersage des Studienerfolgs auf.

Die Stichprobe umfasst N=109 Bachelor Studierende im Fach Psychologie an der Universität des Saarlandes (M= 21.6 Jahre, SD= 3,66 Jahre), die im Rahmen des Projektes „WiKom-Sowi: Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern“ (Zhang et al., 2013) befragt wurden. Erfasst wurden die durchschnittliche Abiturnote und Intelligenz welche über die drei Subskalen Analogien, Zahlenreihen und Matrizen aus dem Intelligenz-Struktur-Test 2000R (IST 2000R; Liepmann et al., 2007) operationalisiert wurde. Gewissenhaftigkeit wurde mit der Kurzversion des Big Five Inventory (BFI-K; Rammstedt & John, 2005) erfasst. Die Kurzversion des Leistungsmotivationsinventar (LMIK; Schuler & Prochaska, 2001) diente zur Erfassung der Leistungsmotivation. Der Studienerfolg wurde über den Mittelwert der Noten in den Prüfungsmodulen Quantitative Methoden I und II ermittelt.

Zur Überprüfung der ersten Hypothese wurde eine lineare Regressionsanalyse mit den unabhängigen Variablen Abiturnote und Intelligenz durchgeführt, um deren prädiktive Validität auf die Studienleistung zu untersuchen. Darauf aufbauend wurde eine zwei-schrittige hierarchische Regressionsanalyse zur Überprüfung der zweiten Hypothese unter der Hinzunahme der Gewissenhaftigkeit durchgeführt; zur Überprüfung der dritten Hypothese wurde parallel dazu eine zwei-schrittige hierarchische Regressionsanalyse unter der Hinzunahme der Leistungsmotivation zur Überprüfung der dritten Hypothese berechnet. Die erste lineare Regressionsanalyse unter Verwendung der Prädiktoren Intelligenz und Abiturnote erbrachte eine Varianzaufklärung von 14,3 % auf ( $F(2,106) = 9.99, p < .001$ ). Darauf basierend wurde im zweiten Schritt die inkrementelle Validität der Gewissenhaftigkeit sowie der Leistungsmotivation zur Testung der zweiten und dritten Hypothesen mittels hierarchischer Regressionsanalyse separat voneinander untersucht: Mit der Aufnahme von Gewissenhaftigkeit klärte das Gesamtmodell 16,5 % Varianz auf ( $F(3,105) = 8.10, p < .001$ ). Die inkrementelle Validität von Gewissenhaftigkeit betrug 2,9 % ( $\Delta R^2 = .029, \Delta F(1, 105) = 3.75, p = .055$ ) und war marginal signifikant. Mit der Aufnahme von Leistungsmotivation klärte das Gesamtmodell keine zusätzliche Varianz auf. Interessanterweise wies die Subskala der Leistungsmotivation Engagement zusätzliche Varianz auf. Die inkrementelle Validität von Engagement betrug 2,5 % ( $\Delta R^2 = .025, \Delta F(1,105) = 3.16, p = .078$ ) und war marginal signifikant.

Zusammengefasst wurde die prädiktive Validität von Abiturnote und Intelligenz für Studienerfolg gewährleistet. Darüber hinaus war die inkrementelle Validität von Gewissenhaftigkeit marginal signifikant. Entgegen der Erwartung zeigte Leistungsmotivation keine inkrementelle Validität für Studienerfolg.

Unsere Befunde sprechen für die Bedeutsamkeit der kognitiven Prädiktoren wie Abiturnote und Intelligenz. Besonders bei Untersuchungen zur Bedeutsamkeit der non-kognitiven Prädiktoren wie Persönlichkeit und Motivation sollen die kognitiven Variablen aufgrund der hohen Varianzerklärung mitberücksichtigt und kontrolliert werden. Die zusätzliche Varianzerklärung von 2,5% durch Engagement spricht dafür dass die kurze Skala der Leistungsmotivation als eine gesamte Skala für die Vorhersage des Studienerfolgs im Fach quantitative Methoden eventuell zu breit ist.

Busato, V.V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and individual differences* 29(6), 1057-1068.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality*, 17, 237-250.

Farsides T., & Woodfield R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and individual differences* 34(2), 1225-1243.

Kappe, R., & van der Flier, H. (2012). Predicting academic success in higher education: what's more important than being smart? *European journal of psychology of education*, 27(4), 605-619.

Liepmann, D., Beauducel, A., Brocke, B., & Rudolf, A. (2007). Intelligenz-Struktur-Test 2000R (I-S-T 2000R). (2., erweiterte und überarbeitete Auflage). Göttingen: Hogrefe.

Rammstedt, B., & John, O.P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit. *Diagnostica*, 51 (4), 195-206.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353-387.

Schmidt-Atzert, L. (2005). Prädiktion von Studienerfolg bei Psychologiestudenten. *Psychologische Rundschau*, 56 (2), 131-133.

Schuler, H., & Prochaska. (2001) Leistungsmotivationsinventar. Dimensionen berufsbezogener Leistungsbeurteilung. Göttingen: Hogrefe.

Steyer, R., Yousfi, S., & Würfel, K. (2005). Prädiktion von Studienerfolg: Der Zusammenhang zwischen Schul- und Studiennoten im Diplomstudiengang Psychologie. *Psychologische Rundschau*, 56 (2), 129-131.

Zhang, Y., Klopp, E., Dietrich, H., Brünken, R., Krause, U.-M., Spinath, B., Stark, R., & Spinath, F. M. (2013). Modeling and Measuring Scientific Competencies in Social Sciences. In S. Blömeke & O. Zlatkin-Troitschanskaia (Eds.), *The German funding initiative "Modeling and Measuring Competencies in Higher Education": 23 research projects on engineering, economics and social sciences, education and generic skills of higher education students (KoKoHS Working Papers, 3)*. Berlin & Mainz: Humboldt University & Johannes Gutenberg University.





## Nutzen und Effektivität von Studienhilfen im Mathematik-Nebenfach-Studium

Heublein, Richter, Schmelzer und Sommer (2012) berichten in ihrer Analyse über Studienabbruchquoten an deutschen Hochschulen eine Durchschnittsabbruchquote für Bachelorstudiengänge von 35%, für die Mathematik sind es sogar 55%. Heublein, Hutzsch, Schreiber, Sommer und Besuch (2010) ermittelten folgende Ursachen: Leistungsprobleme (33%), mangelnde Studienmotivation (25%) und Studienbedingungen (13%). Die Studie von Voelkle und Sander (2008) zeigte anhand eines SEM-Modells, dass eine schlechte Hochschulzugangsberechtigungsnote nicht direkt auf eine hohe Abbruchrate wirkt, sondern diese zu schlechten Studienleistungen führt, die dann einen Studienabbruch zur Folge haben. Auf alle Studiengänge bezogen fanden Schiefele, Streblow und Brinkmann (2007) als ungünstige Voraussetzung für ein erfolgreiches Studium eine schlechte Abiturnote, ungünstige Lernstrategien und fehlende Motivation. Aufgrund dieser schlechten Prognosen für Studienanfänger der Mathematik wollten wir explorieren, welche Voraussetzungen günstig sind, damit Mathematik-Nebenfach-Studierende erfolgreich die Abschlussklausur absolvieren. Da in Mathematik an der Universität Frankfurt Tutorien und e-learning eingesetzt werden, um Studierende zu unterstützen, wird zugleich deren Effektivität in Bezug auf die Klausurnote untersucht. Daraus ergaben sich drei Fragestellungen: 1. Welche Probanden brechen das Tutorium ab und welche Probanden nehmen an der Klausur teil? 2. Welche Voraussetzungen sind günstig für eine gute Klausurleistung? Aus der Literatur werden folgende günstige Prädiktoren angenommen: gute Schulmathematiknote, günstige Lernstrategien, günstige Motivation. 3. Haben Studierende, die das e-learning-Angebot häufiger nutzen, eine bessere Klausurnote?

An der Studie nahmen 156 Studierende (davon 40 weiblich, 15 ohne Angabe, Alter:  $M=20.5$  ( $SD=2.25$ )) des Bachelor-Studiengangs Meteorologie ( $n=47$ ) und Physik ( $n=103$ ) im WS 2012/13 teil, die das Mathematik-Nebenfach bestehen müssen. Insgesamt gab es drei Messzeitpunkte: am Semesterbeginn, drei Wochen vor Semesterende und nach der Klausur. Am Semesterbeginn wurden unter anderem die Mathematikschulnote, die Lernstrategien mit dem AssisT von Entwistle, Tait und McCune (2000) und die Motivation mit dem Interest/Enjoyment-Fragebogen von Williams und Deci (1996) erhoben. Auch drei Wochen vor Semesterende wurden die Lernstrategien und die Motivation erfasst. Am Ende des Semesters wurde das e-learning-Angebot ausgewertet. Hier wurde die Häufigkeit des Zugriffs auf Musteraufgaben und/oder Vorlesungsaufzeichnungen registriert. Das erste Ziel war zu untersuchen, welche Studierenden die Klausur mitschrieben und welche nicht. Es ließen sich vier Gruppen unterscheiden, nämlich Studierende, die an dem Tutorium teilnahmen und die Klausur mitgeschrieben haben (KI+Tu+,  $n = 78$ ), solche, die am Tutorium teilnahmen, aber nicht die Klausur mitschrieben (KI+Tu-,  $n = 15$ ), solche, die das Tutorium abbrachen, aber an der Klausur teilnahmen (KI+Tu-,  $n = 14$ ) und Studierende, die weder das Tutorium zu Ende führten, noch an der Klausur teilnahmen (KI-Tu-,  $n = 49$ ). KI-Tu- unterscheidet sich von allen anderen durch die schlechteste Mathematiknote, die geringste Motivation und eher oberflächenbetontes Lernen. Die Studierenden ohne Tutorium, aber mit Klausur (KI+Tu-) schneiden vergleichbar mit KI+Tu+ ab und haben auch in Bezug auf Oberflächenstrategien und Motivation vergleichbare Werte. Ihre guten Mathematikkenntnisse (Schulmathematiknote) und das e-learning-Angebot haben anscheinend das fehlende Tutorium kompensiert. Die KI-Tu+-Gruppe unterscheidet sich von den Klausurteilnehmenden durch eine geringere Motivation. Das zweite Ziel war, gute Prädiktoren für die Klausurleistung festzustellen. Eine gute Mathematikschulnote, günstige Lernstrategien und eine günstige Motivation sollten wenige Wochen vor der Klausur einen Einfluss auf den Erfolg haben. Allerdings ergab sich nur ein Einfluss der Mathematiknote und dem Ablehnen von Oberflächen-Lernen. Das vierte Ziel war, dass wir die Nutzung des e-learning-Angebots untersuchen wollten. Insgesamt haben nur 53% das e-learning-Angebot genutzt, von diesen haben 98% Vorlesungsaufzeichnungen angeschaut und 90% Musteraufgaben heruntergeladen. Eine Regression belegte, dass die Gesamt-Online-Zugriffe auch bei Konstanthalten der Schulmathematiknote zu einer besseren Klausurleistung führen.

Insgesamt zeigt die Studie, dass Studierende in den Tutorien auf den Einfluss von Lernstrategien aufmerksam gemacht werden sollten. Die Aufmerksamkeit der Studierenden sollte zudem stärker auf das hilfreiche e-learning-Angebot gerichtet werden.

Entwistle, N. J., Tait, H., & McCune, V. (2000). Patterns of response to an approaches to studying inventory across contrasting groups and contexts. *European Journal of the Psychology of Education*, 15(1), 33-48.

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen (HIS. Forum Hochschule 2, 2010). Hannover: Hochschul-Informationssystem.

Heublein, U., Richter, J., Schmelzer, R., & Sommer, S. (2012). Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen (HIS. Forum Hochschule 3, 2012). Hannover: Hochschul-Informationssystem.

Schiefele, U., Streblow, L., & Brinkmann, J. (2007). Aussteigen oder Durchhalten. Was unterscheidet Studienabbrecher von anderen Studierenden? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39(3/9), 127-140.

Voelkle, M. C., & Sander, N. (2008). University dropout. A structural equation approach to discrete-time survival analysis. *Journal of Individual Differences*, 29(3), 134-147.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology*, 70, 115-126.

## **Experimentelle Manipulation von Ergebnismeldungen – Einfluss qualitativer und quantitativer Komponenten der Lehrevaluationsrückmeldung auf die Urteilsbildung von Lehrenden**

Die Rückmeldung, die Lehrende aus den Evaluationen durch ihre Studierenden erhalten, soll von ihnen für die Optimierung der eigenen Lehre genutzt werden. Inwiefern Lehrende das Feedback tatsächlich zu diesem Zweck nutzen und welche Rückmeldeformate die Wahrscheinlichkeit günstig beeinflussen, dass das Feedback von Lehrenden nicht nur verstanden, sondern auch in Handlungsintentionen umgesetzt wird, ist weitgehend unbekannt. Der vorliegende Beitrag untersucht die Frage, ob Rückmeldungen in Form statistischer Aussagen – also mittlere Angaben von Studierenden auf vorgegebenen Antwortskalen – und Rückmeldungen in Form offener Kommentare einen differentiellen Einfluss auf die Beurteilung des Feedbacks durch den jeweiligen Lehrenden haben. Wir erwarten, dass qualitative Rückmeldungen stärker gewichtet werden als quantitative, obwohl letztere eigentlich aussagekräftiger sind. Denn Lehrende verwenden bei der Interpretation ihrer Evaluationsergebnisse vermutlich Urteilsheuristiken als "mentale Daumenregeln" (Tversky & Kahneman, 1973, 1983). Wird beispielsweise die Repräsentativitätsheuristik angewendet, wird die Wahrscheinlichkeit von Ereignissen auf der Grundlage des Ausmaßes ihrer Übereinstimmung mit bestimmten prototypischen Ereignissen beurteilt – ein offener Kommentar, der repräsentativ für eine gute oder schlechte Lehrveranstaltung ist, könnte also möglicherweise einen besonders großen Einfluss auf die Interpretation der Evaluationsergebnisse haben. Auch die Anwendung der Verfügbarkeitsheuristik (Tversky & Kahneman, 1973) kann zu einer Übergewichtung offener Kommentare relativ zu Aussagen in Form statistischer Kennwerte führen, da lebhaftere und vielleicht auch ungewöhnliche Kommentare leichter aus dem Gedächtnis abgerufen werden können, also stärker verfügbar sind, als ein statistischer Kennwert auf einer Urteilsskala. In einer Vorstudie wurden quantitative Rückmeldungen und qualitative Kommentare, die aus einer Lehrevaluationsrückmeldung stammen könnten, Thurstone-skaliert (Thurstone, 1928). Dazu wurden 12 Personen gebeten, verschiedene fiktive quantitative Ausprägungen auf verschiedenen Lehrevaluationskriterien und verschiedene fiktive offene Kommentare hinsichtlich ihrer Valenz von sehr negativ bis sehr positiv zu bewerten. Aus den so gewonnenen Bewertungen wurden Ergebnismeldungen (Vignetten) generiert, in denen quantitative und qualitative Rückmeldungen mit unterschiedlicher Valenz jeweils kombiniert wurden. In einem 3x4-Design wurden insgesamt 234 Lehrende verschiedene Vignetten (von sehr negativ über neutral bis sehr positiv) präsentiert. Sie sollten angeben, wie eine entsprechende Ergebnismeldung über eine selbst gehaltene Lehrveranstaltung von ihnen bewertet würde (Manipulation Check), wie stark sie sich durch sie veranlasst sehen würden, etwas an ihrer Lehre zu verändern und für wie valide und glaubwürdig sie die Rückmeldung halten würden.

Erste varianzanalytische Analysen bestätigten u.a. einen besonders starken Einfluss qualitativer Rückmeldungen: Lehrende, die eine sehr positive quantitative, dazu aber eine einzige sehr negative qualitative Rückmeldung erhalten haben, waren deutlich unzufriedener mit der Rückmeldung als Lehrende, die ebenfalls eine sehr positive quantitative Rückmeldung und eine einzige positive qualitative Rückmeldung erhalten hatten. Ähnliche Erkenntnisse ließen sich beim Handlungsimpuls erkennen. Hier gaben Lehrende den stärksten Handlungsimpuls an, wenn sie eine negative qualitative Rückmeldung und sehr positive quantitative Rückmeldungen bekommen hatten. Diese Befunde bestätigen unsere Vermutung, dass offene Kommentare unter Anwendung von Urteilsheuristiken einen stärkeren Einfluss auf die Beurteilung einer Lehrevaluationsrückmeldung nehmen als rational ist.

Aufbauend auf dieser Erkenntnis werden in der Diskussion Möglichkeiten erarbeitet, wie es im Rahmen der Ergebnisdarstellung von Lehrevaluationen gelingen kann, eine angemessene Beurteilung bei den Lehrenden zu unterstützen; wie also verhindert werden kann, dass sie sich von extremen Einzelurteilen zu stark beeinflussen lassen. Abschließend soll kritisch geprüft werden, wie eine optimale Passung zwischen (durch die Studierenden rückgemeldetem) Änderungsbedarf und (durch die Lehrenden initiierten) Anpassungsprozessen erreicht werden kann.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529-54.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 42, 207-232.

Tversky, A. & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293-315.

## **Stand der hochschuldidaktischen Forschung zur Unterrichtsqualität an Hochschulen**

Die Qualität der Lehre wird in der Öffentlichkeit viel diskutiert und rückt aktuell verstärkt in den Fokus der Hochschulforschung. Innerhalb der hochschuldidaktischen Forschung existieren bereits viele Beiträge, die sich zum Teil jedoch sehr spezifisch auf einzelne Studiengänge beziehen. Insgesamt mangelt es an Metaanalysen und Überblicksartikeln. Zur Unterrichtsqualität an Schulen liegen dagegen umfassende Forschungsergebnisse vor. Ziel des vorliegenden Beitrages ist es, den aktuellen Stand der hochschuldidaktischen Forschung aufzuarbeiten und Implikationen für die Übertragbarkeit von Modellen zur Unterrichtsqualität an Schulen auf die Hochschullehre abzuleiten. Um den Stand der Forschung aufzubereiten, wurde zunächst mit allgemein gehaltenen Schlagworten wie „higher education didactics“ Literatur über die Plattform scopus recherchiert. Die Abstracts von  $N = 71$  Zeitschriften-Manuskripten wurden in einem ersten Schritt kategorisiert, um einen Überblick über den Stand der hochschuldidaktischen Forschung zu gewinnen. Für die Kategorienbildung wurden die fünf Fortbildungsbereiche der Deutschen Gesellschaft für Hochschuldidaktik als Hauptkategorien gesetzt (Arbeitsgemeinschaft für Hochschuldidaktik, 2005). Dabei wurde die Hauptkategorie „Lehren und Lernen“ in Anlehnung an den Fragebogen „LeKo“ zur Lehrkompetenz weiter ausdifferenziert (Thiel, Blüthmann, & Watermann, 2012). Die vorab gesetzten Hauptkategorien konnten weitgehend beibehalten werden. Nur schwach besetzt waren die Hauptkategorien „Prüfen“ und „Innovatives Entwickeln von Studium und Lehre“, gar nicht besetzt die Kategorie „Beraten“. In der sich daran anschließenden Literaturrecherche wurde insbesondere zu den schwach besetzten Kategorien gezielt nachrecherchiert. Das weitere Vorgehen bei der Übertragbarkeit von Modellen zur Unterrichtsqualität an Schulen auf die Hochschullehre unter Einbezug von Experteninterviews wird zur Diskussion gestellt.

Arbeitsgemeinschaft für Hochschuldidaktik. (2005). Leitlinien zur Modularisierung und Zertifizierung hochschuldidaktischer Weiterbildung. Retrieved from <http://www.dghd.de/download.php?f=affa3cfb4a405681a9b89dee617eeba3>

Thiel, F., Blüthmann, I., & Watermann, R. (2012). Konstruktion eines Fragebogens zur Erfassung der Lehrkompetenz. In B. Berendt, H.-P. Voss & J. Wildt (Eds.), Neues Handbuch Hochschullehre (Vol. Loseblattsammlung. 55. Ergänzungslieferung, Beitrag I 1.13, 27 S.). Stuttgart: Raabe Verlag.

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer

Olga Troitschanskaia (JGU Mainz)

## ***Kompetenzmodellierung und Kompetenzmessung in den Fachdomänen Wirtschaftswissenschaften, Sozialwissenschaften und Ingenieurwissenschaften - nationale und internationale Perspektiven***

Chair(s): Olga Troitschanskaia (JGU Mainz), Reinhold Nickolaus (Universität Stuttgart)

Diskutant/in: Klaus Beck (JGU Mainz)

In letzter Zeit lassen sich auf dem Gebiet der Kompetenzmodellierung auch für den Hochschulbereich verstärkte Forschungsbemühungen beobachten. Sie zielen darauf, theoretische und methodische Konzepte zu entwickeln, die es erlauben, akademische Kompetenzen zu beschreiben und sie messbar zu machen. In einem BMBF geförderten Forschungsprogramm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor“ (KoKoHs) findet dieses Interesse sichtbaren Ausdruck. Weit fortgeschritten sind inzwischen die Arbeiten zur Kompetenzmodellierung, -messung und -entwicklung im Bereich der Lehrerbildung. Andere große Hochschuldomänen wurden bislang in der Kompetenzforschung in Deutschland vernachlässigt. Abgesehen von diesen inhaltlichen Lücken zeigen sich auf dem Weg zu einer tragfähigen Modellierung und zuverlässigen Erfassung akademisch vermittelter Kompetenzen über alle Fächergrenzen hinweg erhebliche konzeptionelle und methodische Herausforderungen. Das gleiche gilt für die Modellierung der Prozesse des Erwerbs von Kompetenzen einschließlich der Berücksichtigung der relevanten Entwicklungsbedingungen. Eine erfolgreiche Bearbeitung dieser Forschungsdesiderate setzt einen gezielt systematischen sowie fachlich und methodisch interdisziplinären Austausch voraus. Aufgrund der nationalen Forschungsdefizite sowie auch zur Sicherung der internationalen Anschlussfähigkeit ist es dabei zudem erforderlich, die internationale Perspektive zu berücksichtigen. Dafür spricht u.a. die Fokussierung auf die Studiendomänen „Business and Economics“ und „Engineering“, welche auch im Rahmen der ersten internationalen Machbarkeitsstudie zur vergleichenden Erfassung akademischer Kompetenzen herangezogen wurden (s. die OECD-Studie „Assessment of Higher Education Learning Outcomes“). Vor diesem Hintergrund sollen die Vorträge dieses Symposiums einen interdisziplinären Überblick zu einem sich international neu konstituierenden Forschungsfeld bieten, das für die akademische Ausbildung bzw. die evidenzbasierte Hochschullehre/Hochschulpolitik von zunehmender Relevanz ist. An der defizitären Forschungslage zur Kompetenzmodellierung in den jeweiligen Domänen ansetzend bieten die Vorträge neben den Ansätzen, die zur Lösung der oben angedeuteten konzeptionellen und methodischen Herausforderungen entwickelt wurden, auch Befunde aus ausgewählten Studiendomänen. Davon fokussieren zwei Vorträge die internationale, ländervergleichende Perspektive der Kompetenzmessung im Hochschulbereich (Musekamp et al., Happ et al.) und zwei weitere präsentieren die nationalen Forschungsperspektiven (Behrendt/Dammann, Wolter/Schiener). Die Vorträge von Musekamp et al. sowie Behrendt/Dammann können im Bereich der Ingenieurwissenschaften verortet werden. Der Vortrag von Happ et al. repräsentiert die wirtschaftswissenschaftliche Fachdomäne, der Vortrag von Wolter/Schiener die Fachdomäne Soziologie. Insgesamt zeigen die Beiträge, dass für diese Fachdomänen trotz der enormen theoretischen und empirischen Herausforderungen bereits beachtliche Fortschritte erreicht wurden – dies sowohl hinsichtlich der theoretischen Modellierung als auch mit Blick auf die Entwicklung von geeigneten Messinstrumenten, von denen interessante Impulse auf die Kompetenzmodellierung in anderen Studiendomänen ausgehen.

**GEBF** TAGUNG  
2014

O P1

Poster-  
Symposium

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer

Wirtschaftspädagogik

2

## **Ingenieurwissenschaftliche Grundkompetenzen, Modellierungen und erste Erklärungsmodelle**

Die Technische Mechanik ist sowohl in Maschinenbau- als auch in Bauingenieurwesen-Studiengängen eine zentrale Thematik, auf der auch viele weitergehende Lehrveranstaltungen, z.B. Konstruktionslehre, aufbauen. Damit werden Modellierungen dieser Kompetenzfacette zu einem zentralen Bezugspunkt ingenieurwissenschaftlicher Kompetenzmodellierungen in diesen Domänen. Da die Grundlagen zum Aufbau dieser Kompetenzen (z.B. Newtonsche Axiome der Mechanik) schon in der Schullaufbahn gelegt werden, scheint es erstrebenswert, nicht nur die im Studium erworbenen Kompetenzen zu messen, sondern, um erste Erklärungsmodelle generieren zu können, auch die zu Beginn des Studiums (noch) vorhandenen Kompetenzen zu berücksichtigen. Die zentralen Fragestellungen des Projekts sind auf die Klärung der Kompetenzstrukturen zu Beginn des Studiums und nach Absolvieren der betroffenen Lehrveranstaltungen sowie die Generierung von Erklärungsmodellen zum Kompetenzerwerb unter Berücksichtigung der kognitiven Eingangsvoraussetzungen sowie struktureller und qualitativer Merkmale der Hochschullehre gerichtet.

Zur Kompetenzstruktur kann im Anschluss an Modellierungsergebnisse aus nichtakademischen Berufsfeldern (im Überblick Nickolaus & Seeber, 2013) unterstellt werden, dass sich die strukturellen Ausdifferenzierungen vor allem entlang von Inhaltsbereichen vollziehen. Als inhaltliche Bezugspunkte kommen sowohl Tätigkeitsbereiche/Merkmale von Ingenieuren als auch curriculare Bezugspunkte in Frage, wobei letztere trotz erheblicher standortspezifischer Varianzen große Überschneidungen aufweisen, die eine übergreifende Modellierung vertretbar machen. Da u.W. keine Kompetenzmodelle zur Technischen Mechanik existieren, wurden in Anlehnung an die Erkenntnisse aus berufsfachlichen Modellierungen und der Mathematik eigene Modelle generiert, die hauptsächlich nach inhaltlichen Aspekten gegliedert sind. Hierbei wird im Eingangstest unterschieden zwischen curricular verankerten Themen und mechanischen Grundvorstellungen, die i.A. nicht Teil des Curriculums sind. Für den Abschlusstest wird von einem dreidimensionalen Modell - Statik, Elastostatik und Dynamik – ausgegangen, wobei geprüft werden soll, ob die Dynamik sich weiter unterteilt in Kinematik und Kinetik. Detaillierte Überlegungen zur Modellierung finden sich in Stefanica (2011).

In den vorliegenden Erklärungsmodellen zur berufsfachlichen Kompetenzen (z.B. Lehmann & Seeber, 2007; Nickolaus et al., 2012) erweisen sich vor allem die kognitiven Eingangsvoraussetzungen sowie die Motivation, die ihrerseits von den Qualitätsmerkmalen des Lehrangebots beeinflusst wird, als prädiktiv. Im Rahmen des längsschnittlich angelegten Projekts KoM@ING werden deshalb u.A. ein Eingangstest zur Erhebung physikalisch/technischer Kompetenzen zum Studienbeginn, ein Abschlusstest zur Erhebung der Fachkompetenzen in den Teilgebieten Statik und Elastostatik am Ende des zweiten Fachsemesters und ein Abschlusstest zur Erhebung der Fachkompetenzen im Teilgebiet Dynamik am Ende des dritten Fachsemesters eingesetzt. Des Weiteren wurden Qualitätsmerkmale der Lehre erhoben. Für den Eingangstest wurden insgesamt 28 Aufgaben aus einem vorpilotierten Itempool (Hindenach, 2010) verwendet. Im Abschlusstest zu den Teilgebieten Statik und Elastostatik finden sich 91 Aufgaben, die im Rahmen des Projektes entwickelt und pilotiert wurden. Diese wurden auf 13 Testhefte à 28 Aufgaben verteilt. Beide Pools wurden unter Berücksichtigung curricularer Analysen (vgl. z.B. Hindenach, 2010; Behrendt, 2009), unter Einbezug von Bildungsplänen bzw. Modulhandbüchern, Vorlesungsunterlagen und Klausuren, generiert. Einbezogen wurden sowohl ausgewählte Universitäten als auch ausgewählte Fachhochschulen mit den Studiengängen Maschinenbau und Bauingenieurwesen. Zum jetzigen Zeitpunkt können für den Eingangstest N=1258 Probanden (1005 Universität, 223 Fachhochschule) und für den Abschlusstest (Teilgebiete Statik und Elastostatik) N=249 Probanden (148 Universität, 96 Fachhochschule) in die Analyse einbezogen werden. In einem echten Längsschnitt liegen gegenwärtig Daten von ca. 100 Probanden vor. Die Analysen erfolgen im Rückgriff auf IRT-basierte Modellierungen und Strukturgleichungsmodelle.

Der Beitrag präsentiert die Modellierungsansätze in Technischer Mechanik und erste Ergebnisse zu den Kompetenzmodellen und Erklärungsmodellen zur Kompetenzentwicklung. Im Eingangstest zeichnet sich eine eindimensionale Kompetenzstruktur ab, die sowohl dem oben skizzierten zweidimensionalen Ansatz als auch einem vierdimensionalen Ansatz, in dem das curricular verankerte Fachwissen weiter unterteilt wird, überlegen ist.

Im Abschlusstest erweist sich ein mehrdimensionales Modell, das in die zwei Dimensionen Statik und Elastostatik mit jeweils den zwei Subdimensionen Einwirkungen und Tragwerkseigenschaften zerfällt, anderen Modellierungen überlegen.



## **Technische Mechanik als Gegenstand der Messung von ingenieurwissenschaftlichen Kompetenzen im nationalen und internationalen Kontext**

Sowohl innerhalb des BMBF Forschungsprogramms „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor (KoKoHs)“ als auch in der OECD Machbarkeitsstudie „Assessment of Higher Education Learning Outcomes (AHELO)“ gibt es derzeit große Anstrengungen, Kompetenzen in den Ingenieurwissenschaften reliabel und valide zu messen. Beide Ansätze unterscheiden sich dabei beträchtlich in Zielstellung, Operationalisierung und Methoden der Kompetenzerfassung. Während im KoKoHs-Projekt „Modellierung und Messung von Kompetenzen der Technischen Mechanik in der Ausbildung von Maschinenbauingenieuren (KOM-ING)“ im Rahmen des Grundstudiums erworbene Kompetenzen in Fach der Technischen Mechanik (TM) erfasst werden (Musekamp et al. [im Druck]), wird mit der OECD Machbarkeitsstudie im Bereich der Ingenieurwissenschaften die Kompetenz erfasst, nach dem Bachelorstudium Leistungen in professionellen Kontexten zu erbringen (Readiness for Workplace Performance, Hadgraft et al., 2012). Ein Bindeglied zwischen zwei so verschiedenen Ansätzen ist die TM als das fachliche Einmaleins für (angehende) Ingenieure: Die TM gilt deshalb als wichtiger Baustein des ingenieurwissenschaftlichen Grundstudiums (siehe KOM-ING), weil sie elementarer Bestandteil des international etablierten Kanons ingenieurwissenschaftlichen Professionswissens ist (Ferguson, 2006). Sie spielt damit auch in AHELO eine zentrale Rolle und mit Bezug zur TM wird ein Vergleich beider Ansätze auf Itemebene möglich. Daraus ergeben sich Einsichten zu zentralen Fragen der Kompetenzmessung im Hochschulbereich, insb. zur Abgrenzung zwischen curricularen und beruflicher Validität von Instrumenten und zur Einbettung international einheitlicher Fachinhalte in international adaptierte Kontexte. Sowohl im KOM-ING als auch im AHELO Projekt ist der Ausgangspunkt für die Kompetenzmessung ein Kompetenzbegriff, dessen konstitutives Merkmal im anwenden Können von psychischen Potenzialen auf jeweils neue Situationen liegt (vgl. Klieme & Hartig, 2007, S. 14; Rychen & Salganik, 2001). Situationen sind wiederum gekennzeichnet durch Umweltbedingungen, die sich den Handelnden als objektive Gegebenheiten präsentieren (Kontext der TM). Unterschiede ergeben sich hinsichtlich der spezifischen Gliederung beider Kompetenzkonzepte. Das KOM-ING Modell unterscheidet einerseits nach den drei TM Inhaltsbereichen Statik, Festigkeitslehre und Dynamik, sowie andererseits nach den vier kognitiven Prozessen des mechanischen Modellierens: reales Objekt zu mechanischem Modell abstrahieren (1), formulieren des mechanischen Modells als Gleichungssystem (2), Lösen des Gleichungssystems (Rechnen, 3) und Ergebnisse der Gleichungen evaluieren (4). Bei AHELO wird Ingenieurkompetenz globaler unterschieden in ‚Basic/Engineering Science‘ (1), ‚Engineering Processes‘ (2) und ‚Generic Skills for Engineering‘ (3). Während das KOM-ING Modell damit ausschließlich die TM Kompetenz umfasst, wird TM im AHELO Ansatz in die arbeitsbezogenen Dimensionen integriert. Vor dem Hintergrund dieser konzeptionellen Ansätze wird die Frage beantwortet, wie sich die Testaufgaben hinsichtlich der ‚TM‘ in beiden Projekten unterscheiden. Zudem wird beantwortet, welche Konsequenzen sich aus Gemeinsamkeiten und Unterschieden in den Testaufgaben in Bezug auf Validierungsstrategien und Validierungsergebnisse sowie in Bezug auf die internationale Einsetzbarkeit der Instrumente ergeben. Im KOM-ING- wie im AHELO Projekt werden Methoden der Item-Response-Theorie (IRT) zur psychometrischen Validierung der angenommenen Kompetenzstrukturen angewandt. Dies umfasst im AHELO Projekt insbesondere die Prüfung auf einheitliche Interpretierbarkeit der gewonnenen Skalen in unterschiedlichen Ländern (Differential Item Functioning, z. B. Osterlind & Everson, 2009) als auch die Prüfung der gewonnenen Daten auf die angenommene mehrdimensionale Struktur (Multidimensional Item-Response Modeling, z. B. Reckase, 2009). Neben Strukturprüfungen wurde im KOM-ING Projekt zusätzlich die Bedeutung von TM-spezifischen Aufgabenmerkmalen für deren empirische Schwierigkeit analysiert. Im Vortrag werden erste Analysen aus KOM-ING und AHELO Projekt präsentiert, die auf eine gute Passung der erhobenen Daten und den jeweils angenommenen Kompetenzstrukturen hindeuten. Sowohl Annahmen zur curricularen Bedeutung im Falle von KOM-ING als auch zur internationalen Vergleichbarkeit im Falle von AHELO scheinen sich zu bestätigen. Ausgehend von den TM bezogenen Itemanalysen werden diese Ergebnisse hinsichtlich ihrer curricularen bzw. berufsbezogenen Validität interpretiert und diskutiert.

Ferguson, Clive (2006): Defining the Australian mechanical engineer. In: European Journal of Engineering Education 31 (4), S. 471–485. Online verfügbar unter [http://pdfserve.informaworld.com/534691\\_907399526\\_747701701.pdf](http://pdfserve.informaworld.com/534691_907399526_747701701.pdf), zuletzt geprüft am 16.12.2010.

Hadgraft, Roger; Pearce, Jacob; Edwards, Daniel; Fraillon, Julian; Coates, Hamish (2012): Assessing Higher Education Learning Outcomes in Civil Engineering: the OECD AHELO Feasibility Study. Proceedings of the 2012 AAEE Conference. Melbourne. Online verfügbar unter <http://www.aaee.com.au/conferences/2012/documents/abstracts/aaee2012-submission-143.pdf>, zuletzt geprüft am 14.10.2013.

Klieme, Eckhard; Hartig, Johannes (2007): Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In: Manfred Prenzel, Ingrid Gogolin und Heinz-Hermann Krüger (Hg.): Kompetenzdiagnostik. Zeitschrift für Erziehungswissenschaft. Sonderheft 8 | 2007. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden (Springer-11776 /Dig. Serial]), S. 11–29.

Musekamp, Frank; Mehrafza, Mostafa; Heine, Jörg-Henrik; Schreiber, Britta; Saniter, Andreas; Spöttl, Georg et al. (im Druck): Formatives Assessment fachlicher Kompetenzen von angehenden Ingenieuren. Validierung eines Kompetenzmodells für die Technische Mechanik im Inhaltsbereich Statik. In: Lehrerbildung auf dem Prüfstand (LpP).

Osterlind, Steven J.; Everson, Howard T. (2009): Differential item functioning. 2nd. Thousand Oaks, Calif: Sage.

Reckase, Mark D. (2009): Multidimensional item response theory. [Online-Ausg.]. New York, NY [u.a.]: Springer.

Rychen, Dominique Simone; Salganik, Laura Hersh (Hg.) (2001): Defining and selecting key competencies. Seattle: Hogrefe & Huber.

## **Vergleich des ökonomischen Wissensstandes von Studierenden in Deutschland und Japan – Implikationen für die vergleichende internationale Kompetenzforschung**

Dass vergleichende internationale Studien für den Hochschulbereich von zunehmender Relevanz sind, zeigen nicht zuletzt auch die Arbeiten zu der Feasibility Study „Assessment of Higher Education Learning Outcomes (AHELO)“ der OECD, die u. a. auch die Studienfachdomäne der Volkswirtschaftslehre in den Blick nimmt (vgl. OECD, 2011). Allerdings stellen solche Studien die internationale Hochschulforschung vor große Herausforderungen, insbesondere wenn hier Länder mit unterschiedlichen kulturellen Hintergründen gegenübergestellt werden. Dieses ist bei einem Vergleich von Japan und Deutschland gegeben. Zwar lassen sich historisch bedingt durchaus einige Parallelen zwischen den Hochschulsystemen dieser beiden Länder identifizieren (vgl. Grewe, 1984; Mützenich, 2011). Allerdings können in beiden Ländern auch kulturelle Spezifika festgestellt werden, welche in ihrem Einfluss auf die Rahmenbedingungen für den ökonomischen Wissenserwerb nicht zu unterschätzen sind (vgl. Schlieper, 1997; Takanashi, 1998). Die kritische Prüfung der Machbarkeit von Vergleichsstudien aufgrund solcher länderspezifischen Differenzen wird in der internationalen Hochschulforschung jedoch nicht selten vernachlässigt. An diesem Forschungsdesiderat knüpft der vorliegende Vortrag an, in dem die Machbarkeit von vergleichenden Analysen der ökonomischen Wissenstestdaten aus dem Hochschulbereich zwischen Deutschland und Japan umfassend geprüft und durch erste empirische Analysen gestützt wird. In beiden Ländern wurden Adaptationen des Test of Understanding College Economics (TUCE4; Walstad, Watts, & Rebeck, 2007) zur Erfassung des volkswirtschaftlichen Wissensstandes von Studierenden eingesetzt. Bevor derartige Vergleiche auf Datenbasis aus den beiden Ländern durchgeführt wurden, erfolgte zunächst eine kritische Prüfung der curricularen und institutionellen länderspezifischen Bedingungsfaktoren als grundlegende Voraussetzung für eine vergleichende Studie. In diesem Kontext wurde auch überprüft, ob die im TUCE thematisierten VWL-Inhalte curricular valide für die Erfassung des ökonomischen Wissens sowohl in Deutschland als auch in Japan sind. Weiterhin wurde betrachtet, ob sich für den Adaptionsprozess für den TUCE (USA – Deutschland vs. USA – Japan) Unterschiede finden lassen. Aufbauend auf diesen Betrachtungen konnten gerichtete Hypothesen über die Unterschiede in den länderspezifischen Ausprägungen des ökonomischen Fachwissens der Studierenden entwickelt werden. Auf Basis dieser Befunde wird in dem Vortrag das folgende Teilziel fokussiert: Es wird geprüft, in wie weit sich die theoretisch modellierten Hypothesen zu Unterschieden und Gemeinsamkeiten zwischen den beiden Ländern, die aus der Analyse der Strukturen und Curricula der beiden Hochschulsysteme und der ermittelten Unterschiede im Adaptionsprozess gewonnen wurden, anhand empirischer Daten aus dem TUCE4 bestätigen lassen. Dem empirischen Vergleich wird mittels gewonnener Testdaten nachgegangen. Die deutschen TUCE-Daten wurden im Jahr 2013 an 23 Hochschulen bundesweit erhoben und umfassen eine Stichprobengröße von ca. 1.000 Studierenden der Wirtschaftswissenschaften. Ebenso wurde der TUCE 4 bei insgesamt ca. 1.100 Studierenden an 10 Hochschulen in Japan eingesetzt (Yamaoka, Asano, & Abe 2010). Erste Ergebnisse aus dem Vergleich der Hochschulsysteme zeigen, dass die Unterschiede zwischen den beiden Ländern doch relativ gering sind (Arimoto, 2007; Hirowatari, 2002). In den Analysen zum Adaptionsprozess zeichnet sich anhand einer Rückübersetzung (Japan – Deutsch) und beim Vergleich der deutschen mit der japanischen Testversion ab, dass der Adaptionsprozess des amerikanischen Testinstrumentes in beiden Ländern ähnlich erfolgte. Welche spezifischen Inhaltsbereiche bzw. Testaufgaben in den adaptierten Testversionen zu Differenzen führen könnten, kann jedoch erst nach weiterführenden Analysen der Items identifiziert werden. Abschließend wird mittels empirischer Daten geprüft, ob es Unterschiede im volkswirtschaftlichen Wissen zwischen den Studierenden aus den beiden Ländern zu finden sind und wie diese Unterschiede im Hinblick auf die Dokumentenanalyse zu Bildungssystemen der beiden Länder sowie Curriculaanalysen interpretiert werden können. Insbesondere erweist sich bei der empirischen Datenprüfung von Interesse, ob sich die Unterschiede in den Bildungssystemen auch zu signifikanten Unterschieden in den Lösungshäufigkeiten der als kritisch identifizierten Aufgaben niederschlagen. Als Fazit werden Implikationen zu Grenzen und Chancen solcher Vergleichsstudien aufgezeigt.

- Arimoto, A. (2007). Japan: Origins, History and Transition to a Universal Higher Education System. In W. Locke & U. Teichler (Eds.), *The Changing Condition for Academic Work and Careers in select Countries*. Werkstattberichte 66. Kassel: Verlag Jenior.
- Grewe, W. (1984). Japan und Deutschland nach dem Kriege – Parallelen und Divergenzen. In: Josef Kreiner (Hrsg.), *Deutschland und Japan. Historische Kontakte* (S. 285-300). Bonn: Bouvier.
- Hambleton, R. K. (2001). The Next Generation of the ITC Test Translation and Adaption Guidelines. *European Journal of Psychological Assessment*, 17(3), 164-172.
- Hirowatari, S. (2002). Hochschulsysteme in Japan und Deutschland. In *Japanisch-Deutsches Zentrum Berlin [jdzb]* (Hrsg.), *Symposium. Hochschulentwicklung und -reform in Japan und Deutschland 05.-06.12.2001* (Band 49, S. 37-38). Berlin: Veröffentlichungen des Japanisch-Deutschen Zentrums Berlin.
- Mützenich, R. (2011). Japan und Deutschland: Gemeinsamkeiten und Unterschiede. In E. G. Fritz & D. Sankyo (Hrsg.), *Zeitenwende in Japan?* (S. 13-20) Oberhausen: Athena.
- OECD (2011). *Tuning-AHELO Conceptual Framework of Expected and Desired Learning Outcomes in Economics*. OECD Education Working Papers, No. 59, OECD Publishing.
- Schlieper, A. (1997). *Die Nähe fremder Kulturen. Parallelen zwischen Japan und Deutschland*. Frankfurt/Main; New York: Campus Verlag.
- Takanashi, A. (1998). Einführung. In A. Takanashi & U. Teichler (Hrsg.), *Berufliche Kompetenzentwicklung im Bildungs- und Beschäftigungssystem in Japan und Deutschland* (S. 11-12). Baden-Baden: Nomos Verlagsgesellschaft.
- Walstad, W. B., Watts, M. & Rebeck, K. (2007). *Test of understanding in college economics: Examiner's manual* (4th ed.). New York NY: National Council on Economic Education.
- Yamaoka, M., Asano, T. & Abe, S. (2010). The present state of Economic Education in Japan. *The Journal of Economic Education* 41(4), 448-460.

## ***Auf dem Weg zum „Hochschul-PISA“? Ergebnisse zweier Pilotstudien zur Messung von fachlichen und methodischen Kompetenzen in der Soziologie***

Im Gegensatz zum allgemeinbildenden schulischen Bereich steckt die Kompetenzdiagnostik im Hochschulbereich immer noch in den Kinderschuhen. Zugleich ist Evaluation an Hochschulen immer noch entweder Input-orientiert, also gerichtet auf Strukturen und Prozesse von Studium und Lehre wie z. B. die sachliche und personale Ausstattung von Instituten, oder stützt sich auf subjektive Indikatoren wie selbsteingeschätzte Lernerfolge von Studierenden oder deren Beurteilungen von Dozenten, Veranstaltungen usw. Ihre Output-Orientierung anhand objektiver Indikatoren ist dagegen schwächer ausgeprägt – valide Messinstrumente zum tatsächlichen Kompetenzerwerb von Studierenden stehen bisher nicht oder nur eingeschränkt zur Verfügung. Dies gilt auch für die Soziologie, wobei dies verwunderlich ist, da gerade in dieser Disziplin die Methodenkompetenz und das Bewusstsein für die Notwendigkeit valider Indikatoren vergleichsweise hoch sein dürften. Neben der evaluatorischen Motivation für eine Kompetenzmessung im Hochschulbereich ist diese auch grundsätzlich für inhaltliche soziologische Fragestellungen von einigem Interesse, wenn es beispielsweise um Determinanten und Erträge des Bildungserwerbs jenseits von Noten oder Zertifikaten geht. So stellt sich für diesen Beitrag die Frage, inwieweit die Methodologie der Kompetenzdiagnostik, die in den letzten Jahren vor allem anhand der großen Schulleistungsvergleiche (TIMSS, PISA usw.) entwickelt wurde, auf spezifische Domänen fachlicher Kompetenz im Hochschulbereich – in diesem Fall fachliche und methodische Kompetenzen in der Soziologie – übertragbar ist. Wir präsentieren Ergebnisse zweier Pilotstudien, die sich der Modellierung und Messung von „soziologischen Kompetenzen“ respektive von „Kompetenzen in Methoden der empirischen Sozialforschung“ widmen. Kompetenzen werden in beiden Studien als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme & Leutner, 2006, S. 879, Herv. im Orig.) verstanden, in weitere Subdimensionen untergliedert, mit entsprechenden Testaufgaben operationalisiert und mit den Verfahren der Item-Response-Theorie (Rasch bzw. Birnbaummodell) skaliert. Zusätzlich zum Kompetenztest umfassen die Erhebungen Fragen zum bisherigen Studienverlauf und -erfolg, zur studienbezogenen Selbstwirksamkeitsüberzeugung, zur Soziodemographie und zum sozialen Hintergrund. Die erste Erhebung zu „soziologischen Kompetenzen“ wurde im Herbst 2012 unter n=540 Soziologiestudierenden an der Johannes Gutenberg-Universität Mainz durchgeführt; die zweite Erhebung zu „Kompetenzen in Methoden der empirischen Sozialforschung“ befindet sich aktuell (Oktober 2013) im Feld, angestrebt wird die Befragung von mehreren Hundert Studierenden der Sozialwissenschaften an mehreren deutschsprachigen Universitäten. Die Ergebnisse der ersten Studie (vgl. dazu Wolter/Schiener 2013) demonstrieren nicht nur die grundsätzliche Durchführbarkeit vergleichbarer Vorhaben, sondern weisen auch auf gute Skaleneigenschaften des hier vorgestellten Kompetenzindikators hin. In ausgewählten inhaltlichen Analysen zeigt sich zudem eine hohe externe Validität des Indikators. Die Befunde der zweiten Studie liegen zum aktuellen Zeitpunkt noch nicht vor, sollen jedoch im Rahmen des Vortrags im März 2014 präsentiert werden. Hierbei wird eingegangen auf praktische erhebungstechnische Aspekte, auf die Skalierbarkeit der entsprechenden latenten Konstrukte sowie auf Analysen zur externen Validität der Kompetenzindikatoren. Erste Ergebnisse eines Pretests (Juli 2013) zeigen hierbei positive Ergebnisse und einige Verbesserungen im Vergleich zur ersten Pilotstudie. Insgesamt sprechen die bisher vorliegenden Ergebnisse dafür, die Idee einer standardisierten Kompetenzmessung in der Soziologie in Zukunft weiterzuentwickeln und auszubauen.

- Braun, Edith / Gusy, Burkhard / Leidner, Bernhard / Hannover, Bettina 2008: Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp), in: Diagnostica 54, S. 30–42.
- Hartig, Johannes / Klieme, Eckhard 2006: Kompetenz und Kompetenzdiagnostik, in: Schweizer, Karl (Hg.): Leistung und Leistungsdiagnostik, Heidelberg: Springer Medizin Verlag, S. 127–143.
- Klieme, Eckhard / Artelt, Cordula / Hartig, Johannes / Jude, Nina / Köller, Olaf / Prenzel, Manfred / Schneider, Wolfgang / Stanat, Petra (Hg.) 2010: PISA 2009. Bilanz nach einem Jahrzehnt, Münster: Waxmann.
- Klieme, Eckhard/Leutner, Detlev 2006: Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG, in: Zeitschrift für Pädagogik 52 (6): S. 876–903.
- Müller-Benedict, Volker / Tsarouha, Elena 2011: Können Examensnoten verglichen werden? Eine Analyse von Einflüssen des sozialen Kontextes auf Hochschulprüfungen, in: Zeitschrift für Soziologie 40, S. 388–409.
- Schmidt, Uwe 2009: Evaluation an deutschen Hochschulen. Entwicklung, Stand und Perspektiven, in: Widmer, Thomas / Beywl, Wolfgang / Fabian, Carlo (Hg.): Evaluation. Ein systematisches Handbuch, Wiesbaden: VS Verlag, S. 163–169.
- Weinert, Franz E. 2001: Concept of Competence: A Conceptual Clarification, in: Rychen, Dominique Simone / Salganik, Laura Hersh (Hg.): Defining and Selecting Key Competencies, Seattle: Hogrefe & Huber, S. 45–65.
- Wissenschaftsrat 2012: Prüfungsnoten an Hochschulen im Prüfungsjahr 2010. Arbeitsbericht mit einem wissenschaftspolitischen Kommentar des Wissenschaftsrates, Hamburg: Geschäftsstelle Wissenschaftsrat.
- Zlatkin-Troitschanskaia, Olga / Blömeke, Sigrid / Kuhn, Christiane / Buchholtz, Christiane 2012: Wirksamkeitsprüfungen im Hochschulbereich. Aufgaben und Herausforderungen des BMBF-Forschungsprogramms "Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor". in: Zeitschrift für Evaluation 11. S. 95–103.

## **Modellierung und Gewichtung von kognitiven und nicht-kognitiven Facetten domänenspezifischer Problemlösekompetenz von Industriekaufleuten**

Gängige Ansätze der Kompetenzmodellierung und -messung postulieren, dass das Konstrukt Kompetenz neben kognitiven Facetten wie Wissen, Fertigkeiten, Fähigkeiten etc. auch nicht-kognitive Facetten wie Interessen, Einstellungen, Motive, Selbstkonzepte umfasst (bspw. Weinert, 2001; Erpenbeck & von Rosenstiel, 2003). In empirischen Studien fällt jedoch auf, dass Letztere zwar in frühen Stufen der Modellierung mitgedacht werden, in der empirischen Modellierung und Operationalisierung aber meist eine untergeordnete Rolle spielen. Üblicherweise werden einer sehr elaborierten Testentwicklung im Bereich des deklarativen und prozeduralen Wissens einige eher globale Selbstberichtsskalen zu nicht-kognitiven Facetten wie Interessen und Selbstwirksamkeitserwartungen zur Seite gestellt, wobei Letztere meist Randnotizen bleiben. Angemessener erschien uns dagegen eine integrierte Erhebung eines komplexen Kompetenzkonstrukts, da kognitive Kompetenzfacetten zwar notwendige, aber nicht hinreichende Prädiktoren der Performanz im realen Arbeitskontext darstellen (Sembill, Rausch, & Kögler, 2013). Das Projekt „Domänenspezifische Problemlösekompetenz bei Industriekaufleuten“ (DomPL-IK) im Rahmen der BMBF-Initiative „Technologie-orientierte Kompetenzmessung in der Berufsbildung“ (ASCOT) greift die skizzierten Defizite auf und zielt auf ein ganzheitliches Kompetenzkonstrukt ab, das sich konsequent am Problemlöseprozess orientiert. Domänenspezifische Problemlösekompetenz umfasst demnach im (a) kognitiven Segment die Facetten (1) Handlungsbedarfe und Informationsquellen identifizieren, (2) Informationsverarbeitung mittels domänenspezifischer Algorithmen und Werkzeuge, (3) Entscheidungen treffen und begründen sowie (4) Entscheidungen angemessen kommunizieren. Im (b) metakognitiven Segment werden (5) planvolles (strukturiertes) Vorgehen, (6) Persistenz (fokussiertes Vorgehen) und (7) retrospektive Handlungskontrolle berücksichtigt. Im nicht-kognitiven Bereich schließlich sind zum einen das (c) domänenspezifische Selbstkonzept als Bündel aus (8) angemessenem Vertrauen in die eigene Kompetenz, (9) Unsicherheits-/Ungewissheitstoleranz und (10) angemessenem Vertrauen in die eigene Lösung zu nennen. Zum anderen wird das (d) domänenspezifische Interesse als Bündel aus (11) inhaltlichem Interesse am Problem, (12) positiver Aktiviertheit bei der Problembearbeitung und (13) Interesse am Fortgang des Problems und an potenziellen Lernchancen modelliert. Zur Erhebung der 13 Facetten wurden komplexe Problemszenarien aus der kaufmännischen Domäne entwickelt, deren Bearbeitung ca. 30 Minuten in Anspruch nimmt. Neben der Bewertung der Endprodukte werden Daten zum nicht-kognitiven Bereich mittels eingebetteter Erlebensstichproben („Embedded Experience Sampling“; Rausch & Seifried, in prep.) gesammelt. Auf die Frage nach der domänenspezifischen Problemlösekompetenz einer Person oder einer Personengruppe mit der Angabe von 13 Testscores für die einzelnen Facetten zu antworten, scheint aus mehreren Gründen unangebracht. Zwar kann eine solch differenzierte Wiedergabe bspw. zur Identifikation spezifischen Förderbedarfs geeignet sein, doch wünscht man sich in der Regel einen integrierten Score, der „die“ Kompetenz widerspiegelt. Mit der Entscheidung für ein Mehrfacettenmodell der Kompetenz stellt sich daher sogleich die Frage nach der Gewichtung der Facetten. Um diese Frage aus Sicht verschiedener Beteiligter zu beleuchten, findet im November 2013 eine Online-Befragung verschiedener Personengruppen statt. Hierzu werden die Facetten des Kompetenzmodells alltagsprachlich formuliert und Beispiele angegeben. Danach werden die Teilnehmenden gebeten, verschiedene Gewichtungen sowohl zwischen den 13 Facetten (siehe Ziffern 1 bis 13) als auch innerhalb und zwischen den vier Kompetenzbereichen (siehe A bis D) vorzunehmen. Der Fragebogen erfasst, für wie wichtig die Teilnehmenden die betreffende Kompetenzfacette oder den betreffenden Kompetenzbereich bei der Bewältigung alltäglicher, fachlicher Probleme in der Domäne halten. Als Teilstichproben vorgesehen sind jeweils etwa 20 Fachkräfte aus der Praxis, 20 Experten aus dem Bereich beruflicher Kompetenzmessung (z.B. innerhalb ASCOT-Initiative), 20 Mitglieder in betreffenden Prüfungsausschüssen der Industrie- und Handelskammern/ Aufgabenstelle für kaufmännische Abschluss- und Zwischenprüfungen, 20 Lehrkräfte und 20 Auszubildende. Es wird erwartet, dass den kognitiven Bereichen (A) und (B) zwar das höhere Gewicht beigemessen wird, aber auch die Facetten der nicht-kognitiven Bereiche (C) und (D) als wichtig erachtet werden. Deskriptive Befunde sowie varianzanalytisch abgesicherte Gruppenunterschiede werden im Rahmen des Vortrags berichtet. Zwar gibt es in zahlreichen Projekten zur Kompetenzmessung mehrdimensionale Kompetenzmodellierungen, doch wird der Gewichtung üblicherweise kaum Beachtung geschenkt. Stattdessen wird oft eine unbegründete Gleichgewichtung vorgenommen. Mit unserem Beitrag möchten wir diese Vorgehensweise problematisieren und mögliche Auswege diskutieren.

Erpenbeck, J., & von Rosenstiel, L. (2003): Einführung. In: Erpenbeck, J. & von Rosenstiel, L. (Hrsg.): Handbuch Kompetenzmessung – erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Stuttgart: Schäffer-Poeschel, IX-XL.

Rausch, A., & Seifried, J. (in prep.). Embedded Experience Sampling (EES)—An integrated approach of measuring non-cognitive facets of competence.

Sembill, D., Rausch, A., & Kögler, K. (2013). Non-cognitive facets of competence—Theoretical foundations and implications for measurement. In K. Beck & O. Zlatkin-Troitschanskaia (Eds.), From Diagnostics to Learning Success: Proceedings in Vocational Education and Training (pp. 199-212). Rotterdam: Sense.

Weinert, F. E. (2001): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, Basel, Bonn: Beltz, 17-31.





Eveline Wuttke (Goethe-Universität Frankfurt), Jürgen Seifried (Universität Mannheim), Christin Siegfried (Goethe-Universität Frankfurt), Andreas Rausch (Universität Bamberg), Steffen Brandt (Art of Reduction)

## **Operationalisierung und Skalierung kognitiver Facetten domänenspezifischer Problemlösekompetenz angehender Industriekaufleute**

In der Tradition von Chomsky (1965/1970) wird Kompetenz in der Regel als hypothetisches, d.h. nicht direkt messbares Konstrukt aufgefasst, das die Qualität beobachtbarer, manifester Performanzen bestimmt und nur über ebendiese Performanzen erschließbar ist (Shavelson, 2010). Kompetenzmodelle im weiteren Sinn umfassen daher sowohl Aussagen darüber, welche Situationen z.B. im beruflichen Umfeld mit Hilfe der betreffenden Kompetenz bewältigt werden können, als auch Aussagen über die korrespondierenden Testsituationen, auf Basis derer die Kompetenz erschlossen wird (Sembill, Rausch, & Kögler, 2013). Die Attribution von Kompetenz erfolgt mittels IRT-Skalierung der aus Testsituationen gewonnenen Performanzdaten. Die zuverlässige Schätzung ist dabei jedoch an einige Voraussetzungen gebunden. In diesem Zusammenhang ist u. a. die lokale stochastische Unabhängigkeit von Bedeutung, die – verkürzt dargestellt – fordert, dass unter Konstanthalten des Personenparameters die Performanz bei einem bestimmten Item eines Tests unabhängig ist von der Performanz bei allen anderen Items des Tests. Dies führte dazu, dass Testinstrumente neueren Datums aus einer Sammlung vieler, in sich abgeschlossener Testfragen bestehen. Testökonomische Anforderungen von Large-scale-Studien führten zudem zu einer Bevorzugung geschlossener Antwortformate. Die Messung komplexerer Personeneigenschaften wie Problemlösekompetenz sollte jedoch konsequenterweise anhand der in komplexen Problemsituationen gezeigten Performanzen beurteilt werden. Weder die Forderung nach geschlossenen Items noch die lokale stochastische Unabhängigkeit sind hierbei ohne Weiteres gewährleistet. Im Rahmen dieses Beitrags möchten wir unsere Vorgehensweise bei der Messung kognitiver Facetten domänenspezifischer Problemlösefähigkeit darstellen, die den testtheoretischen Herausforderungen gerecht wird, ohne die Komplexität der Testsituation zu beschneiden. Das Projekt „Domänenspezifische Problemlösekompetenz bei Industriekaufleuten“ (DomPL-IK; Projekt im Rahmen der ASCOT-Initiative des BMBF) zielt auf die Messung von Problemlösekompetenz angehender Industriekaufleute im Bereich Controlling / Beschaffung ab. Die Teilnehmenden bearbeiten drei etwa 30-minütige authentische Problemszenarien, bei denen Aufträge von einem fiktiven Vorgesetzten übergeben und in einer technologiebasierten Testumgebung mittels umfangreicher Informationen (Geschäftsbriefe, Stücklisten, Rechnungen, Belege etc.) und domänenspezifischer Werkzeuge (Tabellenkalkulation, Taschenrechner, Notizblock etc.) bearbeitet werden. Das Ergebnis ist der vorgesetzten Person in einer Antwortmail zu übersenden und zu erläutern. Die in dieser Antwortmail, den bearbeiteten Dokumenten sowie in sonstigen Dokumenten manifestierten Performanzen werden hinsichtlich folgender kognitiver Kompetenzfacetten ausgewertet: (1) Handlungsbedarfe und Informationsquellen identifizieren, (2) Informationsverarbeitung mittels domänenspezifischer Algorithmen und Werkzeuge, (3) Entscheidungen treffen und begründen sowie (4) Entscheidungen angemessen kommunizieren. Ziel ist es, eine Antwortkodierung zu entwickeln, die eine reliable Messung jeder dieser vier Facetten erlaubt. Zur Kodierung der produzierten Texte, verwendeten Formeln, errechneten Werte etc. wird dazu in einem ersten Schritt für jedes der drei Szenarien ein Kodierleitfaden erstellt, der je Kompetenzfacette bis zu zehn Kategorien enthält nach denen die Antworten der Probanden beurteilt werden (in der Regel dichotom). In einem zweiten Schritt werden dann je Szenario die Kategorien einer Facette zu einem einzelnen Partial Credit Item verdichtet. Hierzu werden die auftretenden Kombinationen der Kodierungen innerhalb einer Facette in Tabellenform dargestellt und per Expertenrating dann in eine Reihenfolge hinsichtlich ihrer Güte gebracht. Diese Reihenfolge ist dann Basis für die Definition des Partial Credit Items. Das beschriebene Verfahren hat den Vorteil, dass die grundlegende Kodierung auf vergleichsweise wenig komplexen, klar voneinander abgrenzbaren dichotomen Kodierungen basiert. Das anschließende Rating der Antwortkombinationen über die Kategorien gewährleistet dabei, dass die Aggregation der Kodierungen eindimensional im Sinne der von den Experten definierten Kompetenzfacette ist. Weiterhin liefert das Verfahren auf diese Weise je Szenario und Facette genau ein Partial Credit Item, das in die IRT-Skalierung eingeht. Die Gefahr fehlender lokaler stochastischer Unabhängigkeit von Items innerhalb eines Szenarios wird auf diese Weise umgangen. Ein weiterer Schritt im Rahmen der Kodierung wird sein, Möglichkeiten einer (teil-)automatisierten Auswertung der Antworten (die Kodierung der oben beschriebenen Kategorien) zu untersuchen. Im Rahmen des Beitrags werden wir die ersten beiden Schritte (inhaltsanalytische Auswertung und Verdichtung zu Partial Credit Items) vorstellen und anhand der auf diese Weise generierten Ergebnisse diskutieren.

Chomsky, N. (1965/1970): Aspekte der Syntax-Theorie. Frankfurt/Main, Berlin: Suhrkamp.

Sembill, D., Rausch, A., & Kögler, K. (2013). Non-cognitive facets of competence—Theoretical foundations and implications for measurement. In Shavelson, R. J. (2010). The Quest to Assess Learning and Hold Higher Education Accountable. National Research Council.

Shavelson, R. J. (2010). On the measurement of competency. Empirical Research in Vocational Education and Training, 2(1), 2010, 41-63.

GEBF TAGUNG  
2014

O P3

Postergruppe

Montag,

03.03.2014,

17:30 - 18:30,

Casino-Gebäude

im Foyer

Wirtschaftspädagogik





## **Fachkompetenzentwicklung bei Industriemechanikern vom Ausbildungsbeginn bis zur Zwischenprüfung**

Der Beruf des Industriemechanikers ist in Deutschland einer der am häufigsten gewählten Ausbildungsberufe. Die dreieinhalbjährige Ausbildung erfolgt dual, ausgerichtet auf die Förderung beruflicher Handlungskompetenzen, deren Kern die Fachkompetenz darstellt. Im Rahmen des vom BMBF geförderten Promotionsprojekts wird der erste Ausbildungsabschnitt, vom Beginn der Ausbildung bis zur Zwischenprüfung nach 1,5 Jahren, genauer beleuchtet wobei die Entwicklung der Fachkompetenz besonders im Fokus steht. Ziel des Projektes ist es ein Erklärungsmodell für die Fachkompetenzentwicklung bei Industriemechanikern bis zur Zwischenprüfung zu entwickeln. Notwendige Voraussetzung hierfür ist die Entwicklung eines Instrumentariums, mit dem die Fachkompetenz bei Industriemechanikern bis zur Zwischenprüfung erfasst werden kann.

Theoretischer Hintergrund: Wie sich in mehreren anderen Studien im beruflichen Bereich gezeigt hat, sind neben fachspezifischen Vorkenntnissen, Basiskompetenzen und allgemeinen kognitiven Voraussetzungen, vor allem die wahrgenommene betriebliche und schulische Ausbildungsqualität so wie die Motivation von entscheidender Bedeutung für die Entwicklung der Fachkompetenz (Nickolaus, Geschwendtner, & Geißel, 2008; Lehmann & Seeber, 2007; Nickolaus et al., 2010). Die Vermutung liegt nahe, dass sich für Industriemechaniker ähnliche Einflussfaktoren herauskristallisieren. In diesem Projekt werden daher ergänzend Einflussfaktoren wie der IQ, fachspezifische Vorkenntnisse, Mathematikkenntnisse, Lesekenntnisse und Motivationale Aspekte, so wie wahrgenommene Qualitätsmerkmale in der Schule und dem Betrieb miterfasst, um deren Einfluss auf die Kompetenzentwicklung empirisch zu prüfen. Neben den Einflussfaktoren wurden in mehreren Berufen (Industriekaufleute, Bankkaufleute, Bürokaufleute, KFZ-Mechatroniker und Elektroniker für Energie- und Gebäudetechnik) bereits mehrdimensionale Kompetenzstrukturen der Fachkompetenz empirisch bestätigt (Winther & Achtenhagen, 2009; Rosendahl & Straka, 2011; Seeber, 2008; Geschwendtner, 2011; Nickolaus et al., 2011). Auch bei den Industriemechanikern sind solche Strukturen zu vermuten, was im Rahmen der Studie genauer geprüft werden soll.

Fragestellung: Im Rahmen der Tagung ist vorgesehen ausgewählte Ergebnisse des Projektes vorzustellen. Schwerpunktmäßig soll der Frage nachgegangen werden, inwieweit sich die Basiskompetenzen (Lesen/Mathematik) und die fachspezifischen Vorkenntnisse zu Beginn der Ausbildung als relevante Prädiktoren für das erreichte Fachkompetenzniveau zum Schuljahresende erweisen. Darüber hinaus soll geprüft werden inwieweit sich schulische und betriebliche Ausbildungsqualitäten als relevante Prädiktoren für die Entwicklung der Fachkompetenz darstellen.

Methode: Die Daten der Längsschnittstudie wurden in 22 Industriemechanikerklassen an 10 Berufsschulen in Baden-Württemberg erhoben. Insgesamt beteiligten sich ca. 500 Schüler an der Studie, die mit drei Messzeitpunkten (zu Beginn der Ausbildung, zum Ende des ersten Lehrjahres und zum Zeitpunkt der Zwischenprüfung) durchgeführt wurde. Die eingesetzten Tests wurden als Paper-Pencil-Test realisiert. Zum dritten Messzeitpunkt wurde im Fachtest ein Multi-Matrix-Design verwendet. Für die Beantwortung der im Vortrag behandelten Fragestellung werden verschiedene mit Mplus erstellte Strukturgleichungsmodelle präsentiert.

Ergebnisse: Erste Analysen der durchgeführten Tests ergaben zufriedenstellende bis gute Testgüten für die entwickelten Instrumente zu Erfassung des Fachwissens. Darüber hinaus zeigen sich erwartungskonforme Bezüge zwischen den untersuchten Einflussfaktoren, welche im Bezug auf ihren Einfluss auf die Fachkompetenzentwicklung in weiteren Auswertungen noch weiter analysiert und im Rahmen des Tagungsbeitrags vorgestellt werden.

GSCHWENDTNER, T. (2011): Die Ausbildung zum Kraftfahrzeugmechatroniker im Längsschnitt. Analysen zur Struktur von Fachkompetenz am Ende der Ausbildung und Erklärung von Fachkompetenzentwicklungen über die Ausbildungszeit. In: Nickolaus, R./Pätzold, G. (Hrsg.): Lehr-Lernforschung in der gewerblich-technischen Berufsbildung. Beiheft zur Zeitschrift für Berufs- und Wirtschaftspädagogik 4/09, S.557-578

LEHMANN, R./ SEEBER, S. (Hrsg.) (2007): Untersuchungen von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen (ULME III). Behörde für Bildung und Sport Hamburg.

NICKOLAUS, R./ GEIßEL, B./ ABELE, S./NITZSCHKE, A. (2011): Fachkompetenzmodellierung und Fachkompetenzentwicklung bei Elektronikern für Energie- und Gebäudetechnik im Verlauf der Ausbildung – Ausgewählte Ergebnisse einer Längsschnittstudie. In: Seifried, J. u.a.(Hrsg.): Beiheft zur Zeitschrift für Berufs- und Wirtschaftspädagogik. Stuttgart: Franz Steiner Verlag

NICKOLAUS, R./ STRAKA, G. A./ FEHRING, G./ GSCHWENDTNER, T./ GEIßEL, B./ ROSENDAHL, J. (2010): Erklärungsmodelle zur Kompetenz- und Motivationsentwicklung bei Bankkaufleuten, Kfz-Mechatronikern und Elektronikern. In: Seifried, J./Wuttke, E. (Hrsg.): Lehr-Lern-Prozesse in der kaufmännischen Bildung gestalten und erforschen (Sonderheft der Zeitschrift für Berufs- und Wirtschaftspädagogik)

NICKOLAUS, R./ GSCHWENDTNER, T./ GEIßEL, B. (2008): Entwicklung und Modellierung beruflicher Fachkompetenz in der gewerblich-technischen Grundbildung. In: Zeitschrift für Berufs- und Wirtschaftspädagogik (ZBW), 104. Bd., H.1, S. 48-73

ROSENDAHL, J./ STRAKA G.A. (2011): Analysen zur wirtschaftlichen Fachkompetenz angehender Bankkaufleute. In: Zeitschrift für Berufs- und Wirtschaftspädagogik.

SEEBER, S. (2008): Ansätze zur Modellierung beruflicher Fachkompetenz in kaufmännischen Ausbildungsberufen. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, Band 104, Heft 1, S. 74-97.

WINTHER, E./ ACHTENHAGEN, F. (2009): Skalen und Stufen kaufmännischer Kompetenz. In: Zeitschrift für Berufs- und Wirtschaftspädagogik. Bd. 105, Heft 4, S. 521-556.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
im Foyer



Die Signifikanz von financial literacy gewinnt im Rahmen des nationalen und internationalen Forschungskontexts zunehmend an Bedeutung (Aprea, 2012). Nicht zuletzt durch die Finanzkrise ist es unabdingbar, dass Individuen in einer Gesellschaft eigenständig und sicher mit Finanzprodukten agieren können (Schürkmann & Schuhen, 2013). Die Bedeutung von financial literacy wird zudem durch den facettenreichen Gebrauch des Begriffs in unterschiedlichen Forschungskontexten verstärkt. Viele Studien versuchen financial literacy bei Erwachsenen zu messen. Schülerorientierte Studien nehmen im Forschungsfeld nur einen geringen Anteil ein und werden erst durch die Zusatzstudie zur financial literacy in PISA 2012 gestärkt (OECD, 2012). Trotz der zunehmenden Bedeutung haben sich einige Länder nicht an der Zusatzstudie beteiligt. Dabei wäre es interessant im Rahmen der Schülerbildung zu erkennen, welche Fähigkeiten Schülerinnen und Schüler im Bereich des Finanzwesens besitzen, um im momentanen und späteren Alltag finanziell handlungsfäh zu sein. Hierbei stellt sich zunächst die Frage, welche Inhalte einer financial literacy zugesprochen werden (hierzu: Schlösser et al., 2011; Macha & Schuhen, 2011; Kaminski & Friebel, 2012). Viele Studien verknüpfen den Begriff financial literacy mit einzelnen Aspekten zum Thema Sparen, Kreditkarten, Vorsorge oder Verschuldung (bspw. Lusardi & Mitchell, 2006; Lusardi & Tufano, 2009; Klapper & Lusardi, 2012; Panos, 2012; NFCC, 2013; Disney & Gathergood, 2012; Guiso & Tullio, 2008). Studien, die financial literacy als Gesamtkomposition verstehen und inhaltlich umfassende differenzierte Kategorien beinhaltet, konnten bisher nicht gefunden werden. Doch nicht nur die Frage nach relevanten Inhaltsfeldern spielt im Kontext von financial literacy eine forschungsrelevante Rolle. Auch die Frage nach direkt messbaren Handlungen ist ein Bestandteil einer financial literacy. Zunehmend mehr Finanzinstitute bieten Online-Tools als Hilfestellung und Informationsbeschaffung an. Kunden können Kreditraten, Sparraten und Renten individuell online berechnen (hierzu: Krottsch & Locher, 2012). Die Fähigkeit mit diesen Online-Tools umgehen zu können, wird als weiterer Teil einer financial literacy angesehen (Schürkmann & Schuhen, 2013).

Im Rahmen dieses Problemfelds fokussiert der Beitrag einen innovativen Weg in der Auswertungsmethodik innerhalb der Kompetenzmessung am Beispiel des Projektes FILS (Financial Literacy Study). FILS ist eine schülerorientierte Studie, die computerbasiert an verschiedenen Schulen in Nordrhein-Westfalen, Rheinland-Pfalz und Baden-Württemberg im Jahr 2012 durchgeführt wurde. Financial literacy wird im Projektkontext als ein Konstrukt angesehen, welches die Studienstruktur und die Konzeption der einzelnen Aufgaben berücksichtigt, wodurch eine Kompetenzstruktur (Koeppen et al., 2008) für FILS entwickelt werden konnte. Im Kontext der ökonomischen Bildung wurden für FILS fünf Inhaltsfelder kategorial hergeleitet (Schulden, Vermögensbildung, Versichern und Steuern, Zahlungsverkehr sowie Geldpolitik), die zusammen mit der Handlungsvariable zum Umgang mit Online-Tools den Kern des Konstrukts zur financial literacy bilden (Schürkmann & Schuhen, 2013). Die erhobenen Daten dieser sechs Felder werden mit Hilfe des Raschmodells ausgewertet und (Ayala, 2009) gemeinsam mit weiteren Variablen aus einem Einstellungstest zum Umgang mit Geld (vgl. hierzu Barry & Breuer, 2012; zurückzuführen auf Yamauchi & Templer, 1982; Furnham, 1984; Tang, 1992) als Inputvariablen für ein Konstrukt einer financial literacy dargestellt. Die Konstruktvalidität, die für diese Analyse eine notwendige Bedingung darstellt, wird mit Hilfe eines Strukturgleichungsmodells durch die Software Mplus (Muthén & Muthén, 2012) geprüft. Dabei stellen die Items des Einstellungstests das exogene manifeste Messmodell und die inhalts- bzw. handlungsspezifischen Items das endogene latente Messmodell dar. Zusammen erklären die Items das Konstrukt financial literacy für die FILS-Studie (RMSEA = 0.028, CFI = 0.937,  $\chi^2 = 49.901$ ,  $df = 39$ ,  $p = 0.1133$ ). Die Intention des Beitrags zur Tagung fokussiert sich auf die Entwicklung eines neuen Zugangs zur Auswertung von Kompetenzstudien am Beispiel von FILS. Die Implementierung raschmodellierter Werte (bspw. Rost, 2004) in den konzeptionellen Aufbau der Strukturgleichungsmodelle (vgl. u.a. Weiber & Mühlhaus, 2010) lässt eine Validitätsprüfung durch spezifische Gütekriterien innerhalb der Kompetenzmessung zu. Dadurch können die Beziehungen zwischen den differenziert gemessenen Kompetenzen innerhalb eines heterogenen Konstrukts dargestellt und auf weitere Faktoren bezogen werden.

Aprea, Camela (2012): Messung der Befähigung zum Umgang mit Geld und Finanzthemen: Ausgewählte Instrumente und alternative diagnostische Zugänge. Hg. v. (Berufs- und Wirtschaftspädagogik online) Herausgeber: Tade Tramm Susan Seeber & H.-Hugo Kremer bwp@ (Nr.22). Online verfügbar unter [http://www.bwpat.de/ausgabe22/aprea\\_bwpat22.pdf](http://www.bwpat.de/ausgabe22/aprea_bwpat22.pdf).

Ayala, R. J. de (2009): The theory and practice of item response theory. New York: Guilford Press (Methodology in the social sciences).

Barry, Daniela/ Breuer, Klaus (im Druck) (2012): Die Einstellung zu Geld bei Jugendlichen und jungen Erwachsenen - Entwicklung eines Instruments in deutscher Sprache. Discussion paper number 1202. (Hg.): Gesellschaftliche Teilhabe trotz Schulden C.W. Hergenröder. Wiesbaden. Online: [http://wiwi.uni-mainz.de/Dateien/DP\\_1202.pdf](http://wiwi.uni-mainz.de/Dateien/DP_1202.pdf) (30.08.12).

Disney, Richard / Gathergood, John (2012): Financial Literacy and Consumer Credit Use; Working Paper 12/01 Centre for Finance and Credit Markets (CFCM), Nottingham

Furnham, Adrian (1984): Many Sides of the Coin: The Psychology of Money Usage, in Personality and Individual Difference 5, 501-509

Guiso, Luigi / Jappelli, Tullio (2008): Financial Literacy and Portfolio Diversification; European University Institute, Departement of Economics; EUI Working Papers, ECO 2008/31

Kaminski, Hans / Friebel Stephan (2012): Arbeitspapier "Finanzielle Allgemeinbildung als Bestandteil der ökonomischen Bildung". Online: [http://www.ioeb.de/sites/default/files/img/Arbeitspapier\\_Finanzielle%20Allgemeinbildung\\_Juli\\_2012.pdf](http://www.ioeb.de/sites/default/files/img/Arbeitspapier_Finanzielle%20Allgemeinbildung_Juli_2012.pdf) (11.09.13).

Klapper, Leora F. / Lusardi, Annamaria / Panos, Georgios A. (2012): Financial Literacy and the financial crisis; National Bureau of Economic Research (NBER) Working Paper 17930; Cambridge

Koeppen, Karoline / Hartig, Johannes/ Klieme, Eckhard / Leutner, Detlev (2008): Current Issues in Competence - Modeling and Assessment. In: Zeitschrift für Psychologie / Journal of Psychology 2008; Vol. 216(2):61-73

Krottsch, S. / Locher, C. (2012): Banking News - Datenbasiertes Multikanal - Marketing: Signale für den Vertrieb. In: Bank - Zeitschrift für Bankpolitik und Bankpraxis, (11), 30

Lusardi, Annamaria / Mitschell, Olivia S. (2006): Financial Literacy and Planning: Implications for Retirement Wellbeing; National Bureau of Economic Research (NBER) Working Paper 17078; Cambridge

Lusardi, Annamaria / Tufano, Peter (2009): Debt literacy, financial literacy, and overdebtiness; National Bureau of Economic Research (NBER) Working Paper 14808; Cambridge

Macha, Klaas/ Schuhen, Michael (2012): Kompetenzmessung in der Domäne Ökonomie. In: Schuhen, Michael/ Wohlgemuth, Michael/ Müller, Christian (Hrsg.) (2012): Ökonomische Bildung und Wirtschaftsordnung. (Lucius&Lucius) Stuttgart, S. 125-138.

Muthén, L.K. and Muthén, B.O. (1998-2012): Mplus User's Guide, Seventh Edition. Los Angeles, CA: Muthén & Muthén

NFCC (National Foundation for Credit Counseling) (2013): The 2013 Consumer Financial Literacy Survey; Prepared By: Harris Interactive Inc. Public Relations Research; [http://www.nfcc.org/newsroom/FinancialLiteracy/files2013/NFCC\\_NBPCA\\_2013%20FinancialLiteracy\\_survey\\_datashet\\_key%20findings\\_032913.pdf](http://www.nfcc.org/newsroom/FinancialLiteracy/files2013/NFCC_NBPCA_2013%20FinancialLiteracy_survey_datashet_key%20findings_032913.pdf)

Organization For Economic Cooperation and Development (OECD) (2013): Pisa 2012 Assessment and Analytical Framework. Mathematics, Reading, Science, Problem Solving and Financial Literacy, OECD Publishing. <http://dx.doi.org/10.1787/9789264190511-en>

Rost, Jürgen (2004): Lehrbuch Testtheorie - Testkonstruktion. 2. Aufl. Bern [u.a.]: Huber (Psychologie Lehrbuch).

Schlösser, Hans Jürgen / Neubauer, Maria / Tzanova, Polia (2011): Finanzielle Bildung. In: Aus Politik und Zeitgeschichte. 12/2011, 21.03.2011, 21-27.

Schürkmann, Susanne / Schuhen, Michael (2013): Kompetenzmessung im Bereich financial literacy Ergebnisse zum Umgang mit Online-Rechnern aus der FILS-Studie. In: Zeitschrift für Ökonomische Bildung, Ausgabe 1, 73-89.

Tang, Thomas Li-Ping (1992): The meaning of money revisited, in Journal of Organizational Behavior 13, 197-202.

Weiber, Rolf; Mühlhaus, Daniel (2010): Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS. Berlin, Heidelberg: Springer (Springer-Lehrbuch).

Yamauchi, Kent T. / Templer, Donald I. (1982): The Development of a Money Attitude Scale. In: Journal of Personality Assessment 46.5, 522-528.

Montag,  
03.03.2014,  
17:30 - 18:30,  
Casino-Gebäude  
im Foyer



## ***Financial Literacy: ein facettenreiches Konstrukt und eine zentrale Kompetenz im 21. Jahrhundert***

Chair(s): Gerhard Minnameier (Goethe Universität Frankfurt)

Diskutant/in: Eveline Wuttke (Goethe Universität Frankfurt)

Financial Literacy kann verstanden werden als Fähigkeit und Bereitschaft, verantwortungsbewusst mit eigenen finanziellen Ressourcen umzugehen und darüber hinaus seine gesamtgesellschaftliche Rolle als mündiger Wirtschaftsbürger wahrzunehmen. Dazu ist (zumindest) finanzbezogenes Wissen nötig, aber auch zielführende Einstellungen und Interessen gegenüber finanzbezogenen Themen sind unabdingbar. In den vergangenen Jahren rückte Financial Literacy zunehmend in den Fokus institutionalisierter Bildungsprozesse und darauf bezogener Forschungsinteressen. Gerade auch die Gruppe der Jugendlichen und jungen Erwachsenen in der beruflichen Bildung ist von Interesse, da sie an der Schwelle zu finanzieller Autonomie steht und – teilweise zum ersten Mal – eigenständige Finanzentscheidungen treffen muss.

Im Symposium werden verschiedene Facetten von Financial Literacy aufgegriffen. Im ersten Beitrag wird am Beispiel der Schweizer Berufsbildung untersucht, wie und in welchem Umfang Financial Literacy curricular verankert ist, d.h. welchen Stellenwert schulische Curricula dieser Kompetenz zumessen. Der zweite Beitrag nimmt ebenfalls junge Erwachsene in den Blick und untersucht, welches Verständnis diese Zielgruppe von Finanzsystemen hat. Die Analysen beschränken sich hier nicht darauf, Financial Literacy als eine auf den persönlichen Lebensbereich ausgerichtete Fähigkeit abzubilden, sondern richten den Blick auch auf die gesellschaftliche Rolle.

Der dritte Beitrag analysiert Finanzwissen und Einstellungen gegenüber Geld bei verschiedenen Gruppen junger Erwachsener. Dies geschieht vor dem Hintergrund einer zunehmenden Verschuldung oder Überschuldung gerade dieser Gruppe.

Während sich die ersten drei Beiträge auf junge Erwachsene mit erster Verantwortung im Bereich des finanziellen Handelns konzentrieren, nimmt der vierte Beitrag Investoren in den Blick und geht der Frage nach, ob und wie sie aus vergangenen Investmententscheidungen lernen.

Methodisch spannen die Beiträge des Symposiums den Bogen von Curriculumanalysen über qualitative Interviews und Fragebogenerhebungen bis hin zu einer Feldstudie mit dem Einsatz verschiedener Treatments bei 4 Gruppen. Das Konstrukt Financial Literacy wird dabei unter wirtschaftspädagogischer, bildungspolitischer und wirtschaftswissenschaftlicher Perspektive sowie bei verschiedenen Zielgruppen beleuchtet.

Im Symposium werden Ergebnisse aus vier Studien präsentiert. Dabei sollen Unterschiede und Gemeinsamkeiten aufgezeigt und diskutiert werden.

## **Curriculare Repräsentationen von Financial Literacy: Das Beispiel der Schweizerischen Berufsbildung**

**Fragestellung:** Die Förderung von Financial Literacy, verstanden als Befähigung und Bereitschaft, in verständiger Art und Weise über den Einsatz finanzieller Ressource zu entscheiden, wird in der bildungspolitischen wie auch der bildungspraktischen Diskussion zunehmend als wichtige Zielstellung institutionalisierter Bildungsprozesse angesehen. Dies gilt auch für Lernende in der Berufsbildung, die aufgrund ihres Ausbildungsgehalts an der Schwelle zur finanziellen Autonomie stehen und sich damit unmittelbar mit der Notwendigkeit konfrontiert sehen, kompetent mit ihren finanziellen Mitteln umzugehen. Eine wesentliche Voraussetzung dafür, dass adäquate Lernangebote entwickelt werden können, ist die Berücksichtigung finanzbezogener Themen in den entsprechenden curricularen Ordnungsmitteln. In diesem Beitrag wird daher der Frage nachgegangen, in welchem Umfang und in welcher Form finanzbezogene Themen in Curricula der beruflichen Bildung repräsentiert sind. Die Ausführungen erfolgen am Beispiel der Schweizerischen Berufsbildung, in der drei verschiedene Anspruchsniveaus unterschieden werden, nämlich zwei-, drei-, und vierjährige Ausbildungsprogramme. Ergänzend zur bereits genannten Fragestellung wird dementsprechend außerdem eruiert, inwiefern Unterschiede im Hinblick auf die curricularen Repräsentationen finanzbezogener Themen in Ausbildungsprogrammen verschiedener Anspruchsniveaus vorzufinden sind.

**Theoretischer Hintergrund:** Zur Beantwortung der beiden Fragen wurde ein Analyseraster entwickelt, welches auf zwei Rahmenkonzepten beruht, nämlich erstens, dem PISA 2012 Financial Literacy Assessment Framework und zweitens, dem sozio-ökonomisch orientierten Ansatz von Davies (2012). Im PISA 2012 Financial Literacy Assessment Framework werden unter anderem die Dimensionen „Inhalte“ und „Prozesse“ unterschieden. Die Dimension „Inhalte“ umfasst verschiedene, vorwiegend haushaltsökonomisch orientierte Wissensbestandteile von Financial Literacy. Im Einzelnen handelt es sich dabei um: Geld und Transaktionen, Planung und Verwaltung von Finanzen, Risiko und Ertrag sowie Finanzprodukte. Bei der Prozessdimension geht es um mentale Vorgänge zur Verarbeitung dieser Inhalte, wobei zwischen Identifikation, Analyse, Evaluation und Anwendung von Informationen unterschieden wird. Der Ansatz von Davies (2012) stützt sich auf internationale Arbeiten aus dem Bereich der Economics and Citizenship Education und bezieht neben individuellen Finanzentscheidungen auch umfassendere systemische Aspekte wie z.B. „(Fehl-)Verhalten von Banken“ und „Finanzgebaren von Regierungen“ als inhaltliche Bestandteile von Financial Literacy ein.

**Methode:** In die Curriculumanalyse wurden der Rahmenlehrplan für den Allgemeinbildenden Unterricht in der Schweizerischen Berufsbildung sowie 21 Schullehrpläne aus ausgewählten Kantonen der Deutschschweiz einbezogen. Bei der Auswahl dieser Schullehrpläne wurde die Repräsentation aller deutschsprachigen Kantone sowie die Abdeckung verschiedener Ausbildungsberufe und Schulgrößen berücksichtigt. Die Analyse der entsprechenden Dokumente, welche von zwei Personen durchgeführt wurde, folgte ausgehend von den oben genannten Analysekategorien einem an Standards der Inhaltsanalyse (e.g. Krippendorff, 2012) orientierten Verfahren, welches die folgenden Schritte vorsieht: (1) Etablierung eines geteilten Verständnisses der Analysekategorien; (2) Individuelles Lesen der Lehrpläne und Exzerpieren der Ziele und Inhalte; (3) Individuelle Kategorisierung gemäß Analyseraster; (4) Diskussion der Ergebnisse und Klärung von Nichtübereinstimmungen; (5) Re-analyse und Ergebnisinterpretation.

**Ergebnisse:** In Bezug auf die Inhaltsdimension gemäß PISA (2012) Framework wurden Unterschiede im Hinblick auf die verschiedenen Anspruchsniveaus identifiziert, die sich insbesondere im Gesamtumfang und der Tiefe der Abdeckung finanzbezogener Inhalte manifestieren. Demgegenüber wurden hinsichtlich der anteilmäßigen Verteilung der vier Inhaltsbereiche nur geringe Unterschiede festgestellt. Die meisten Lernziele und Inhalte entfallen in allen drei Ausbildungsprogrammen auf den Bereich „Planung und Verwaltung von Finanzen“, gefolgt von „Geld und Transaktionen“, „Risiko und Ertrag“ sowie „Finanzprodukte“. Hinsichtlich der Prozessdimension lässt für alle Anspruchsniveaus eine starke Überrepräsentation der eher anspruchssarmen Kategorie „Identifikation von Informationen“ konstatieren. Dies trifft vor allem auf die Lehrpläne der zweijährigen Ausbildungsprogramme zu, bei denen mehr als die Hälfte der Lernziele dieser Kategorie zuzuordnen sind. Sozio-ökonomische Aspekte im Sinne von Davies (2012) werden in den Curricula nur sehr selten und nur in drei- und vierjährigen Ausbildungsprogrammen berücksichtigt. Sofern vorhanden, betreffen sie vor allem die Rolle und die Verantwortung des Staates, und sind auf einem sehr hohen Abstraktionsniveau formuliert.

Die Ergebnisse der Curriculumanalyse werden mit Blick auf Implikationen für die unterrichtliche Umsetzung und allfällige Konsequenzen für Curriculumrevisionen diskutiert.

Davies, P. (2012). Financial Literacy in the context of democracy. Paper presented at the ECER Conference, Cadiz, 22nd-25th September.

Krippendorff, K. (2012). Content analysis (3rd ed.). Newsbury Park, CA: Sage.

PISA (2012). PISA 2012 Financial Literacy Assessment Framework. Downloaded from <http://www.oecd.org/pisa/pisaproducts/46962580.pdf>

## ***Students' and Trainees' Financial Knowledge and Attitudes towards Money***

Within the international discussion of 21st-century learning, this contribution addresses the skill of financial literacy (Trilling & Fadell, 2009). The research focuses on the money management of adolescents, especially their financial knowledge and their attitudes towards money. These young people are confronted with the challenge of a self-regulated lifestyle and of economic responsibility. Most adolescents master this challenge without major problems. Nevertheless, latest studies in Germany show that the rate of personal insolvencies in this group is continuously increasing. As a result, the public request for better financial education is growing, which mostly addresses better financial knowledge. From an educational perspective, such a one-dimensional approach in respect to knowledge only cannot be sufficient. Financial Literacy is defined as “a combination of awareness, knowledge, skill, attitude and behavior necessary to make sound financial decisions and ultimately achieve individual financial wellbeing” (Atkinson & Messy, 2012). Following their understanding of Financial Literacy, we consider financial knowledge and attitudes towards money to be, among others, important factors that determine money management.

We capture financial knowledge and attitudes towards money within four sub-groups of young adults between 18 and 25 years. Those are Bachelor students in business administration versus technical disciplines against students within a vocational training program again within the financial versus the technical domain. These two factors (students versus apprentices and financial versus technical inclination) are supplemented by a third factor, which is time of direct experiences in coping with financial requirements. The operationalization of this third factor is based on the differentiation between incoming (first-semester) and graduating (last-semester) students respectively apprentices. The two points of measurement (in a cross-sectional survey) differentiate at least two and a half years of self-regulated financial practice both for students as well as apprentices. The basic assumption underlying the design is that different inclinations towards domains, different exposure to financial contents and the duration of self-regulated experiences in handling money will result in significant differences of financial knowledge and attitudes towards money.

To gather the level of financial knowledge, we used Lusardi and Mitchells' (2007) scale on financial literacy. For capturing the attitudes towards money we use the new scale developed by Barry (2013). For analyzing the differences between the groups, we use a multivariate variance analysis. Sample-sizes for each of the eight sub-groups are approximately 200 participants. The raw data collected throughout Germany in total covers more than 2.500 individual cases.

Results show a significant increase of financial knowledge of Bachelor students in economics/ business administration and students within a vocational training program again within the financial sector (banker and insurance intermediaries). The level of financial knowledge of Bachelor students and students within a vocational training program in technical domains do not reveal significant increases. Looking at the attitudes towards money, no important differences between the groups could be found. The assumption in literature that gaining new financial knowledge would lead to a different attitude towards money cannot be confirmed.

Atkinson, A. & Messy, F. (2012). Measuring Financial Literacy: Results of the OECD / International Network on Financial Education (INFE) Pilot Study. OECD Working Papers on Finance, Insurance and Private Pensions, No. 15, OECD Publishing. Gefunden am 05.03.2013 unter <http://www.oecd-ilibrary.org/docserver/download/5k9cfs90fr4.pdf?expires=1370513095&id=id&acname=guest&checksum=0509908118DA747FFA44B5B1C9F29317>

Barry, D. (2013). Die Einstellung zu Geld bei jungen Erwachsenen – Eine Grundlegung aus wirtschaftspädagogischer Sicht. Dissertation am Fachbereich Rechts- und Wirtschaftswissenschaften der Johannes Gutenberg-Universität Mainz.

Lusardi, A. & Mitchell, O. S. (2007). Financial Literacy and Retirement Planning: New Evidence from the Rand American Life Panel. Gefunden am 08.10.2013 unter <http://www.mrrc.isr.umich.edu/publications/papers/pdf/wp157.pdf>

Trilling, B. & Fadel, C. (2009). 21st Century Skills. Learning for Life in our Times. San Francisco: Jossey-Bass.



Patricia Alexander (University of Maryland)

***Information Management versus Knowledge Building: What Studies of Multiple Source Use Reveal about Students' Learning and Performance***

Chair(s): Johannes Hartig (DIPF)

Today's students live and learn in a world that could not have been imagined a generation or even a decade ago — a world where a deluge of information continues unabated. Today virtual “friends” can number in the hundreds; students tweet, blog, and text with abandon; they read “e-books;” they watch what media they want, when they want, and where they want; they have “smart” phones that easily outperform the sophisticated computers of only a few years ago; and, the boundaries between private lives and public personas have become obscured. But what are the consequences of such dramatic changes in information volume, access, and use on the nature of teaching and learning? That is the fundamental question explored in this presentation. Specifically, Patricia Alexander shares findings from a series of studies that have examined the way in which students' conceive of and engage with online sources in the completion of academic tasks. Implications of this program for educational research and instructional practice, as well as for the assessment of students' information management and knowledge building in various academic domains are discussed.

Dienstag,  
04.03.2014,  
09:00 - 10:00,  
Casino-Gebäude  
Raum 823

Chair(s): Beatrice Rammstedt (GESIS-Leibniz-Institut für Sozialwissenschaften)

Diskutant/in: Jürgen Baumert (IPN, Max-Planck-Institut für Bildungsforschung)

Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) untersucht mit dem Programme for the International Assessment of Adult Competencies (PIAAC) in mehr als 20 Ländern Schlüsselkompetenzen der erwachsenen Bevölkerung für die erfolgreiche Teilnahme am gesellschaftlichen Leben. Als zentral hierfür werden die Lesekompetenz, die alltagsmathematische Kompetenz und – insbesondere vor dem Hintergrund der zunehmenden Technologisierung moderner Industriegesellschaften – das technologiebasierte Problemlösen betrachtet. Diese bilden die Grundlage für die Entwicklung weiterer spezifischer Kompetenzen, wie z.B. spezieller beruflicher Kompetenzen. Der Fokus von PIAAC liegt auf der Bevölkerung im erwerbsfähigen Alter (16 bis 65 Jahre).

PIAAC liefert Informationen darüber, inwieweit sich die erwachsene Bevölkerung in den jeweiligen Ländern in den untersuchten Grundkompetenzen unterscheidet, welche Faktoren mit dem Erwerb und dem Erhalt dieser Kompetenzen in Zusammenhang stehen und welche die Auswirkungen dieser Kompetenzen auf die gesellschaftliche und insbesondere wirtschaftliche Teilhabe haben. Der internationale Vergleich ermöglicht, Stärken und Verbesserungsbedarfe in den einzelnen Ländern zu identifizieren und zu überprüfen, welche Rolle beispielsweise Bildungsinstitutionen, Weiterbildungsaktivitäten und das Lernen am Arbeitsplatz für die Ausbildung und den Erhalt der Grundkompetenzen spielen.

In dem Symposium werden zentrale Ergebnisse aus deutscher Sicht präsentiert und mit Befunden verwandter Studien kontrastiert.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
Casino-Gebäude  
Raum 823



Natascha Massing (GESIS – Leibniz-Institut für Sozialwissenschaften), Débora Maehler (GESIS – Leibniz-Institut für Sozialwissenschaften), Christof Wolf (GESIS – Leibniz-Institut für Sozialwissenschaften), Ursula M. Staudinger (GESIS – Leibniz-Institut für Sozialwissenschaften)

### ***Bildung, soziale Herkunft und grundlegende Kompetenzen***

Mit dem Programme for the International Assessment of Adult Competencies (PIAAC) wurden erstmals grundlegende Kompetenzen von Erwachsenen in Deutschland im internationalen Vergleich erhoben. Die Vermittlung dieser Kompetenzen wird als eine der zentralen Aufgaben des Bildungssystems angesehen und bestehende Studien zeigen, dass Bildung ein Schlüsselfaktor für Kompetenzentwicklung ist (u.a. Kirsch, Jungeblut, Jenkins & Kolstad, 2002; OECD & Statistics Canada, 2000). Aus diesem Grund ist ein Zusammenhang zwischen Bildungsabschlüssen und grundlegenden Kompetenzen wie sie in PIAAC erhoben wurden zu erwarten. Das erste Ziel dieser Präsentation ist, diesen Zusammenhang anhand der Daten von PIAAC zu betrachten. Dabei geht es um die Unterschiede zwischen den verschiedenen Bildungsabschlüssen in Bezug auf erreichte Grundkompetenzen, unter Kontrolle von anderen Faktoren, wie zum Beispiel der Muttersprache. Es zeigt sich, dass jeder zusätzliche Bildungsabschluss mit höheren Kompetenzen einhergeht.

Im zweiten Teil geht es um den Zusammenhang zwischen grundlegenden Kompetenzen von Erwachsenen und sozialer Herkunft. Viele Studien in Deutschland stellen fest, dass ein starker Einfluss der sozialen Herkunft auf Bildungsentscheidungen und -abschlüsse sowie auf Grundkompetenzen besteht (u.a. Bos, Tarelli, et al., 2012; Ehmke & Jude, 2010). Auch die sogenannte Bildungsexpansion in den 1960ern und 1970ern, die sich insgesamt positiv auf Bildungschancen auswirkte, hat die soziale Ungleichheitsstruktur nur unwesentlich verändert, da alle sozialen Schichten von den Verbesserungen der Bildungsexpansion profitierten (vgl. Henz & Maas, 1995; Müller & Haun, 1994). Wir analysieren die Frage, ob sich der Effekt der sozialen Herkunft auch bei Erwachsenen in den PIAAC Daten wiederfindet. Als Indikator für die soziale Herkunft nutzen wir die in der Studie vorhandenen Informationen zur Bildung der Eltern. Es wird untersucht, wie die Bildung der Eltern mit Grundkompetenzen einhergeht. Dabei analysieren wir außerdem, ob Geschlechtereffekte erkennbar sind und ob dieser Zusammenhang über die eigene Bildung vermittelt wird. Durch die PIAAC-Daten wird deutlich, dass die Bildung der Eltern sich auch noch im Erwachsenenalter auf Grundkompetenzen auswirkt.

## ***Skill-Mismatch-Maße in PIAAC***

Über- und Unterqualifikation ist ein viel beachtetes Phänomen in der Arbeitsmarktforschung (vgl. Groot & v.d. Brink, 2000; Leuven & Oosterbeek, 2011). Die Befunde belegen, dass Arbeitnehmer mit höherem Bildungsabschluss als für ihren Arbeitsplatz gefordert, weniger verdienen als adäquat beschäftigte Arbeiter mit gleichem Bildungsstand. Gleichmaßen erhalten unterqualifizierte Arbeitnehmer gegenüber adäquat Beschäftigten mit gleichem Bildungsabschluss ein höheres Gehalt. Die Mechanismen, die diese Gehaltskürzungen bzw. -erhöhungen verursachen, sind jedoch noch wenig erforscht. Skill Mismatch, also die Diskrepanz zwischen erworbenen und geforderten Kompetenzen, kann eine Ursache für die beobachteten Effekte von Über- und Unterqualifikation auf das Gehalt sein.

Die Messung von Skill Mismatch ist jedoch schwierig. Insbesondere sind objektive Daten zu Fertigkeiten auf Individualebene häufig nicht verfügbar (Allen & van der Velden, 2001). Mit dem Programme of the International Assessment of Adult Competencies (PIAAC) werden solche Daten bereitgestellt. Ziel des vorliegenden Beitrags ist es, ein objektives Maß für Skill Mismatch zu entwickeln. Es werden dazu verschiedene Ansätze diskutiert und objektive Maße mit dem subjektiven Maß, das die OECD mittels in PIAAC erhobenen Hintergrunddaten berechnet hat, verglichen.

Eine wichtige methodische Einschränkung bei der Messung von Skill Mismatch mit PIAAC ist die Notwendigkeit, einer Person einen individuellen Kompetenzwert zuzuordnen, um sie als well-matched bzw. mismatched zu definieren. Von einer solchen Zuordnung eines individuellen Kompetenzwertes ist jedoch abzuraten, wenn Kompetenzwerte basierend auf der Item-Response-Theorie berechnet wurden, wie es in PIAAC der Fall ist (von Davier, Gonzales & Mislevy, 2009). Diese Einschränkung wird im vorliegenden Beitrag diskutiert.

Die vorgeschlagenen Ansätze zur Berechnung von Skill Mismatch können zur Analyse von Über- und Unterqualifikation sowohl auf nationaler als auch auf internationaler Ebene genutzt werden. Internationale Untersuchungen sind deshalb interessant, um einerseits Unterschiede im Grad des Skill Mismatch über die Länder zu messen und andererseits diese Unterschiede zu Arbeitsmarktinstitutionen und anderen Erklärungsfaktoren in Beziehung zu setzen. Darüber hinaus ist die Untersuchung von Skill Mismatch dafür relevant, die Ausrichtung von formalen und nicht-formalen Bildungsprogrammen an den Anforderungen des Arbeitsmarkts zu verbessern.

Eric A. Hanushek (University of California at Stanford), Guido Schwerdt (Universität Siegen), Simon Wiederhold (ifo Munich), Ludger Woessmann (LMU and ifo Munich)

## **Returns to Skills around the World: Evidence from PIAAC**

The skills of the population are generally viewed as a key ingredient in modern knowledge-based economies. However, existing evidence on the returns to skills in the labor market is surprisingly limited, coming almost exclusively from earnings of early-career workers in the United States. As a result, any sense of how rewards to skill evolve over the work life or of how they might differ across economies is absent. New international data from the Programme for the International Assessment of Adult Competencies (PIAAC) of the OECD dramatically changes the ability to understand how economies value skills. Using those data, this paper provides new insights into the value of skills in different economic settings by developing estimates of the returns to cognitive skills across the entire labor force for 21 countries.

Obtaining estimates of the value of individual cognitive skills has had to rely on a small number of specialized datasets that contain both skill measures and labor-market outcomes. Developing estimates of the returns to skill has most generally relied on U.S. panel data that permit observing the subsequent earnings of youth who can be followed into their initial jobs (see Hanushek and Rivkin (2012) for a survey). However, estimates based on such early-career earnings are likely to suffer from downward lifecycle bias as people with higher lifetime earnings show systematically steeper earnings growth (Haider and Solon (2006)). Furthermore, individual skills may take time to be revealed (Altonji and Pierret (2001)). Finally, returns to skills may depend on a country's specific labor-market and social institutions.

The main prior source of international comparisons has come from analyses of the International Adult Literacy Survey (IALS), collected in the mid-1990s (e.g., Leuven, Oosterbeek, and Ophem (2004); Hanushek and Zhang (2009); Barone and van de Werfhorst (2011); and the review in Hanushek and Woessmann (2011)). However, the returns from two decades ago may no longer be good indicators of the situation in economies that have undergone substantial technological change. Moreover, while informative, the IALS analyses suffer from a number of limitations: a restricted number of 14 countries, relatively small sample sizes, and skill tests that assess very basic competencies.

The new PIAAC survey provides a number of advantages over existing datasets. First, while again observing people across the entire work life, it covers more countries. Second, it provides substantially larger sample sizes per country, allowing richer analyses in subgroups. Third, PIAAC substantially extends the depth and range of measured skills including an attempt to assess technology-oriented problem solving skills.

This work makes three contributions to the literature. First, we document the pattern of returns to cognitive skills, focusing on variations in overall earnings impacts across 21 countries. Second, we provide evidence on the sensitivity of these returns to varying samples and specifications. Third, we describe how these returns differ according to features of country economies including union density and various labor and product market regulations.

Our results confirm that estimates based on early-career earnings underestimate the lifetime returns to skills, in our analyses on average by about one quarter. On average, a one-standard-deviation increase in numeracy skills is associated with an 18 percent wage increase among prime-age workers. But this masks considerable heterogeneity across countries: Eight countries, including all four participating Nordic countries, return less than 15 percent, while six are above 21 percent with the largest return being 26 percent in the United States. Estimates are highly robust to different earnings measures, the specific measure of skills, additional controls, and various subgroups. Intriguingly, the returns to skills are systematically lower in countries with higher union density, stricter employment protection, and larger public-sector shares.



Chair(s): Malte Schwinger (Universität Marburg), Désirée Laubenstein (Universität Koblenz-Landau)

Diskutant/in: Michael Grosche (Universität Potsdam)

Gemäß der UN-Behindertenrechtskonvention sollen SchülerInnen mit sonderpädagogischem Förderbedarf künftig inklusiv beschult werden. Zwar gibt es sowohl in der Gesellschaft als auch seitens der Lehrkräfte und der Wissenschaft große Zustimmung zum Konzept der Inklusion, gleichzeitig werden jedoch häufiger Probleme bei der Umsetzung inklusiven Unterrichts diskutiert. Hierzu zählen beispielsweise eine Überforderung inkludierter SchülerInnen, eine Verlangsamung des Unterrichtstempos für nicht beeinträchtigte SchülerInnen sowie die fehlende Ausbildung von Lehrkräften im Umgang mit lernschwachen Kindern. Angesichts der Tragweite der durch Inklusion anstehenden Veränderungen im Bildungssystem scheint es dringend geboten, die Bedingungen gelingender Inklusion sowohl unter Einbezug mehrerer Disziplinen (Erziehungswissenschaft, Psychologie, Sonderpädagogik) als auch unter Berücksichtigung verschiedener Perspektiven (Kinder, Eltern, Lehrkräfte, Schulen) weitergehend zu erforschen (Bless & Mohr, 2007).

In diesem Symposium soll sowohl der Disziplinen- als auch der Perspektivenvielfalt in insgesamt vier Originalbeiträgen sowie einem Diskussionsbeitrag Rechnung getragen werden. Im ersten Beitrag von Jana Vogel et al. (Universität Potsdam, Grundschulpädagogik) wird aus Perspektive der Kinder und anhand von Daten des Projekts „Inklusive Grundschule“ der Frage nachgegangen, inwiefern eine inklusive Beschulung von Kindern mit sonderpädagogischem Förderbedarf zu eher niedrigen Selbstkonzepten in Mathematik und Deutsch beiträgt und in welchem Maße das Selbstkonzept durch die Schulleistung und das Sozialklima in der Klasse vorhergesagt werden kann. Yotyodying et al. (Universität Bielefeld, Pädagogische Psychologie) thematisieren im zweiten Beitrag sowohl die Kinder- als auch die Elternperspektive. Anhand der Daten des BiLieF-Projekts untersuchen sie, ob sich inklusiv beschulte Kinder in Selbstwertgefühl und schulischem Wohlbefinden von exklusiv beschulten unterscheiden, ob Eltern exklusiv beschulter Kinder aufgrund von Stigmatisierung ein stärker kontrollierendes Erziehungsverhalten zeigen und ob dies in einer erhöhten Bedeutsamkeit elterlichen Erziehungsverhaltens für Selbstwert und Wohlbefinden bei exklusiv beschulten Kindern resultiert. Der dritte Beitrag von Kessels et al. (Universität Berlin, Pädagogische Psychologie) benennt positive affektive Einstellungen der Lehrkräfte als entscheidenden Faktor für erfolgreiche Inklusion. Um Tendenzen sozialer Erwünschtheit und andere Verzerrungen bei der Einstellungserfassung zu vermeiden, werden in dieser Studie die impliziten Einstellungen von Lehramtsstudierenden zur Inklusion erhoben. Désirée Laubenstein (Universität Koblenz-Landau, Sonderpädagogik) analysiert im vierten und letzten Beitrag die Herausforderungen an Schulen und Schulverwaltungen, die eine gelingende Integration lernbeeinträchtigter Kinder in den Regelunterricht mit sich bringen. Auf Basis qualitativer und quantitativer Daten aus dem Projekt GeSchwind geht sie der Frage nach, wie der gemeinsame Unterricht an Schwerpunktschulen weiterentwickelt werden kann und muss, um den Anforderungen einer inklusiven Beschulung genügen zu können.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
IG-Farben-Haus  
Raum 311

Jana Vogel (Universität Potsdam), Stefanie Bosse (Universität Potsdam), Christian Jäntsch (Universität Potsdam), Thorsten Luka (Universität Potsdam), Nadine Spörer (Universität Potsdam)

## ***Selbstkonzepte im inklusiven Grundschulunterricht***

Unter dem Begriff Selbstkonzept wird die Gesamtheit der kognitiven Repräsentationen eigener Fähigkeiten zusammengefasst. Die Annahme bereichsspezifischer Selbstkonzepte im Gegensatz zu einem globalen Konstrukt ist heute allgemein akzeptiert (Möller & Trautwein, 2009). Neben dem akademischen Selbstkonzept stellt das soziale Selbstkonzept, als Teilbereich des nicht-akademischen Selbstkonzeptes, eine wesentliche Komponente dar. Die Ausprägung und Entwicklung des Selbstkonzeptes unterliegt dabei zahlreichen Determinanten. Eine zentrale Herausforderung für die Schulentwicklung stellt derzeit das inklusive Lernen von Kindern mit und ohne sonderpädagogischen Förderbedarf dar. Verschiedene Arbeiten verweisen auf signifikant niedrigere Werte im akademischen Selbstkonzept von integrativ beschulten Kindern mit sonderpädagogischem Förderbedarf (SFB) im Vergleich zu Kindern derselben Schulform ohne SFB und Kindern einer Förderschule (Sauer, Ide & Borchert, 2007) und räumen der Separierung in Sonderschulen eine förderliche Wirkung auf das akademische Selbstkonzept ein (z. B. Kornmann, 2005; Marsh, 2005). Effekte einer inklusiven Beschulung stehen bislang noch aus. Die Wechselwirkungen zwischen akademischem Selbstkonzept und Schulleistung sind mittlerweile unbestritten (Marsh & Craven, 2005, 2006). Im schulischen Kontext gilt jedoch neben einer Reihe weiterer Faktoren das Sozialklima im Klassenraum als eine der entscheidenden Einflussgrößen auf die Entwicklung des Selbstkonzeptes (Grewe, 2003). Bislang wurde diese Annahme allerdings kaum empirisch geprüft. Aktuelle Studien (z. B. Wagreich, 2010) verweisen auf einen geringen Zusammenhang zwischen den Konstrukten.

Der vorliegende Beitrag greift diese Desiderata auf und nimmt das akademische und nicht-akademische Selbstkonzept von N = 1465 inklusiv unterrichteten Schülerinnen und Schülern aus 72 Klassen der 2. und 3. Jahrgangsstufe in den Blick. Dabei werden folgende Fragestellungen untersucht: 1) Lassen sich bedeutsame Unterschiede in der Ausprägung bereichsspezifischer Selbstkonzepte zwischen Kindern mit und ohne SFB identifizieren? 2) Inwieweit können die Selbstkonzepte durch die Schulleistung und das Sozialklima prognostiziert werden?

Grundlage der Studie bilden die Längsschnittuntersuchungen zu Schüler-, Lehrer- und Unterrichtsmerkmalen im Rahmen des Pilotprojekts „Inklusive Grundschule“ des Landes Brandenburg, das seit dem Schuljahr 2012/13 von der Universität Potsdam wissenschaftlich begleitet wird. Berichtet werden die Querschnittsergebnisse des Schülerfragebogens zum ersten Messzeitpunkt. Neben den Selbstkonzepten Lesen und Mathematik (basierend auf Rauer & Schuck, 2003,  $\alpha = .76$  und  $\alpha = .76$ ) wird das soziale Selbstkonzept (ELEMENT,  $\alpha = .71$ ) der Schülerinnen und Schüler betrachtet. Die Schulleistung wurde anhand standardisierter Testverfahren in Deutsch (ELFE 1-6, Lenhard & Schneider, 2006; WLLP-R, Schneider, Blanke, Faust, & Küspert, 2011) und Mathematik (HRT 1-4, Haffner, Baro, Parzer & Resch, 2005) erhoben. Aussagen über den Einfluss des Sozialklimas werden anhand der Skalen Soziale Integration ( $\alpha = .78$ ), Klassenklima ( $\alpha = .74$ ) sowie Gefühl des Angenommenseins ( $\alpha = .77$ , sämtlich basierend auf Rauer & Schuck, 2003) getätigt.

Die Daten befinden sich momentan in der Auswertung und sollen anhand von Mehrebenenregressionsmodellen analysiert werden. Erste Analysen zeigen für Schülerinnen und Schüler mit und ohne SFB signifikante Unterschiede in der Ausprägung der Selbstkonzepte Lesen und Mathematik zugunsten der Kinder ohne Förderbedarf. Ferner konnten neben der fachspezifischen Leistung Indikatoren des Sozialklimas als relevante Prädiktoren der akademischen Selbstkonzepte identifiziert werden.

Sittipan Yotyodying (Universität Bielefeld), Malte Schwinger (Universität Marburg), Birgit Lütje-Klose (Universität Bielefeld), Elke Wild (Universität Bielefeld)

## **Zum differentiellen Einfluss elterlichen Erziehungsverhaltens auf Selbstwertgefühl und schulisches Wohlbefinden bei Kindern in inklusiven und exklusiven Fördersettings**

Theoretischer Hintergrund: Die sonderpädagogische Förderung von Kindern mit Lernbeeinträchtigungen verfolgt unterschiedliche Ziele. Neben guten schulischen Leistungen besteht ein weiteres wichtiges Ziel in der Entwicklung eines positiven Selbstwertgefühls und der Sicherstellung ausreichenden schulischen Wohlbefindens (Bless & Mohr, 2007). Im Zuge der Umsetzung der UN-Behindertenrechtskonvention wird derzeit diskutiert, inwiefern eine inklusive verglichen mit einer exklusiven Beschulung von Kindern mit sonderpädagogischem Förderbedarf im Bereich Lernen (SFB-L) eher zu einer Entwicklung von positivem Selbstwertgefühl und schulischem Wohlbefinden beitragen kann. Zusätzlich zu direkten Effekten inklusiver vs. exklusiver Fördersettings (z.B. durch Bezugsgruppeneffekte, die das Selbstkonzept beeinflussen; Marsh, 2005) ist davon auszugehen, dass kontextuelle Risiko- bzw. Resilienzfaktoren etwaige Schulformeffekte moderieren. Solche Kontextfaktoren können sowohl für den schulischen Bereich (z.B. Klassenklima, Lehrerexpertise) als auch für das familiäre Umfeld (z.B. sozialer Status, Bildungsaspirationen) angenommen werden.

Fragestellung: Bezüglich des familiären Umfelds ist der positive Einfluss autonomieförderlichen, nicht kontrollierenden Erziehungsverhaltens vielfach nachgewiesen worden (Grolnick, Friendly, & Bellas, 2009). Allerdings existieren bislang kaum Studien, die den Einfluss elterlichen Erziehungsverhaltens auf Selbstwertgefühl und schulisches Wohlbefinden von GrundschülerInnen mit SFB-L untersucht haben. Ebenso ist unklar, ob diesbzgl. Effekte unterschiedlich für inklusiv vs. exklusiv beschulte Kinder ausfallen. Die vorliegende Studie greift diese beiden Fragestellungen auf, wobei wir davon ausgehen, dass Eltern von exklusiv beschulten Kindern sich häufiger stigmatisiert fühlen und daher ein stärker kontrollierendes Erziehungsverhalten zeigen. Dies sollte wiederum in einer erhöhten Bedeutsamkeit elterlichen Erziehungsverhaltens für Selbstwert und Wohlbefinden exklusiv vs. inklusiv beschulter Kinder resultieren.

Methode: Die Daten basieren auf der ersten Erhebungswelle der Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements (BiLieF), bei der zwischen Oktober 2012 und März 2013 im gesamten Bundesland Nordrhein-Westfalen N = 414 Schüler/-innen mit SFB-L (Alter: M = 9.12 Jahre, SD = 0.84) und ihre Eltern getrennt voneinander befragt wurden. Von den untersuchten Kindern besuchten n = 177 eine Förderschule mit Förderschwerpunkt Lernen (FS; exklusiv), n = 187 eine Integrationsklasse bzw. gemeinsamen Unterricht (GU; inklusiv) und n = 50 eine Grundschule mit Unterstützung durch ein Kompetenzzentrum für sonderpädagogische Förderung (KsF; inklusiv). Aus Kindersicht wurden Indikatoren des allgemeinen Selbstwerts (Höhe, Stabilität, Kontingenz, des schulischen und allgemeinen Wohlbefindens (Positiver und negativer Affekt, Zufriedenheit) sowie zur wahrgenommenen elterlichen Autonomieunterstützung und Kontrolle erhoben. Aus Elternsicht wurde ebenfalls die Autonomieunterstützung und Kontrolle sowie Stigmatisierung und Scham erfasst. Alle Skalen erwiesen sich als zufriedenstellend reliabel (Cronbachs alphas zwischen 0.76 und 0.92).

Ergebnisse: T-Tests zur Prüfung von Mittelwertsunterschieden ergaben signifikant höhere Werte in Scham und Stigmatisierung für Eltern exklusiv beschulter Kinder. Aus Sicht der Kinder zeigte sich eine signifikant höhere Selbstwertkontingenz und eine höhere wahrgenommene Kontrolle seitens der Eltern bei SchülerInnen aus exklusiven Fördersettings. In Strukturgleichungsmodellen erwies sich elterliche Scham als positiver Prädiktor elterlicher Kontrolle. Wahrgenommene elterliche Kontrolle erwies sich als negativer und wahrgenommene elterliche Autonomieunterstützung als positiver Prädiktor von Selbstwert und Wohlbefinden. Mehrgruppenanalysen zeigten jedoch, dass diese Zusammenhänge in der Gruppe der exklusiv beschulten Kinder signifikant höher ausfielen als in der Gruppe der inklusiv unterrichteten SchülerInnen. Die Ergebnisse legen nahe, dass elterliches Erziehungsverhalten bei Kindern aus exklusiven Fördersettings stärker kontrollierend ausgerichtet ist und verglichen mit inklusiv beschulten Kindern einen stärkeren Einfluss auf kindliches Selbstwertgefühl und Wohlbefinden ausübt.

## **Implizite Einstellungen von Lehramtsstudierenden zur Inklusion**

### Theoretischer Hintergrund:

Zur Umsetzung der UN-Behindertenrechtskonvention werden in Deutschland inklusive Schulen angestrebt. Als entscheidende Faktoren für eine erfolgreiche Implementierung der Inklusion gelten die Einstellungen der beteiligten Lehrkräfte (Avramidis & Norwich, 2002). Deshalb ist es ein wichtiges Forschungsthema, ob diese anstehende Umorientierung des deutschen Schulsystems durch positive Einstellungen von (angehenden) Lehrkräften begleitet wird oder aber mit starken Vorbehalten zu rechnen ist. Vorliegende Studien fokussieren hierbei vor allem die kognitiven Einstellungen, indem Lehrkräfte zum angenommenen Nutzen inklusiver Schulen, v.a. für die Kinder mit sonderpädagogischem Förderbedarf, befragt wurden (z.B. Gebhardt et al., 2011; Kunz, Luder, & Moretti, 2010). Vergleichsweise weniger Befunde liegen zu den affektiven Einstellungen von Lehrpersonen zur Inklusion vor. Zusammengefasst zeigen die meisten Studien, dass (explizit geäußerte) Einstellungen zur Inklusion insgesamt eher positiv sind, vor allem hinsichtlich kognitiver Einstellungskomponenten. Gleichzeitig sind die Erwartungen an Stress und Belastung bei der Antizipation eigenen Unterrichtens in inklusiven Settings durchaus hoch (z.B. Amrhein, 2011; Forlin, Hattie, & Douglas, 1996), was in der Regel mit negativen Emotionen einhergeht (vgl. Cohen, Tyrrell, & Smith, 1993). Auch zeigen mehrere Studien, dass Einstellungen gegenüber Menschen mit Behinderung, sofern diese mit non-reaktiven Verfahren gemessen werden und entsprechend weniger von sozialer Erwünschtheit beeinflusst sind, eher negativ sind. Diese Befunde erscheinen für die Einstellung zur Inklusion nicht unerheblich, weil beim inklusiven Unterrichten alle Lehrkräfte auch mit Kindern mit besonderem Förderbedarf interagieren werden. Auf das Konzept der Inklusion wurden bisher noch keine non-reaktiven Verfahren zur Erfassung impliziter Einstellungen angewendet, obwohl dies gerade bei der Bewertung eines Konzeptes, das als „politically correct idea“ (Avramidis & Norwich, 2002, S. 143) gilt, sinnvoll erscheint.

### Fragestellung:

Unsere Studie zielt darauf ab, die impliziten Einstellungen von Lehramtsstudierenden zum Konzept der Inklusion zu erfassen. Um zu prüfen, ob die Einstellung zum Konzept Inklusion in Abhängigkeit davon variiert, ob dabei auch die Interaktion mit Schülerinnen und Schülern mit sonderpädagogischen Förderbedarf antizipiert wird, wurde eine experimentelle Manipulation durchgeführt, in der zwei verschiedene Versuchsbedingungen realisiert wurden: Der Hälfte der Stichprobe wurde das Konzept Inklusion ohne Bezug zu geistiger Behinderung und der anderen Hälfte mit Bezug zu geistiger Behinderung vorgegeben wird (durch (Nicht-) Verwendung des Wortes „Down-Syndrom“ bei den das Konzept „Inklusion“ beschreibenden Stimuluswörtern). Gemäß vorliegender Literatur wurden als weitere mögliche Einflussfaktoren der Studiengang (Sonderpädagogik vs. allgemeine Lehrämter), die Häufigkeit des Kontaktes zu Menschen mit Behinderung und, als Proxy für die Ausbildung für inklusiven Unterricht, die Häufigkeit, mit der die Studierenden bereits mit dem Thema Inklusion zu tun hatten, einbezogen.

### Methode:

Als Messinstrument wurde der Implizite Assoziationentest (IAT) verwendet. Als Gegenkategorie zur „Inklusion“ wurde das Konzept „mehrgliedriges Schulsystem“ vorgegeben. Ein Fragebogen erfasste Art des studierten Lehramtes, Häufigkeit des Kontaktes mit Menschen mit Behinderung sowie die Häufigkeit des Kontaktes mit dem Thema Inklusion (z. B. im Studium).

### Ergebnisse:

Lehramtsstudierende zeigen im Mittel positive implizite Einstellungen zum Konzept Inklusion (im Vergleich zur Mehrgliedrigkeit), und zwar sowohl in der Bedingung, in der das Konzept Inklusion ohne einen Verweis auf eine geistige Behinderung verwendet wurde, als auch in der Bedingung, in dem auf diese verwiesen wurde. Die impliziten Einstellungen waren umso positiver, je häufiger die Studierenden Kontakt zu behinderten Menschen hatten und je häufiger sie mit dem Thema Inklusion konfrontiert gewesen waren. Bei simultaner Betrachtung und unter Kontrolle des studierten Lehramtes war ausschließlich die Häufigkeit der Beschäftigung mit dem Thema Inklusion prädiktiv für die Einstellung.

Die Ergebnisse werden in Bezug auf die Rolle der Universitäten bei der Umsetzung der UN-Behindertenrechtskonvention diskutiert.

## ***Gelingensbedingungen des gemeinsamen Unterrichts an Schwerpunktschulen in Rheinland-Pfalz (GeSchwind)***

Bildungspolitischer Hintergrund: Rheinland-Pfalz bemüht sich bereits seit Beginn der 1980er Jahre um die Integration von Schüler/innen mit Behinderung und/oder sonderpädagogischem Förderbedarf in Einzelmaßnahmen und Schulversuchen mit Integrations- und Kooperationsklassen (vgl. Bach/Pfiffmann, 1990, 1992; Krawitz/Theis-Scholz, 1999). Seit dem Schuljahr 2001/2002 gibt es im Bundesland Schwerpunktschulen (SPS) mit Integrationsklassen. Dabei handelt es sich nicht um eine neue Schulform, sondern um ein landesweites Schulkonzept, das es Kindern und Jugendlichen mit sonderpädagogischen Förderbedarf ermöglichen soll, möglichst wohnortnah eine allgemeine Schule der Primarstufe (Grundschule) oder Sekundarstufe I zu besuchen (vgl. Ministerium für Bildung, Wissenschaft, Weiterbildung und Kultur 2013). Jede einzelne allgemeine Schule, die zur SPS ernannt wurde, ist aufgefordert, ihr schuleigenes Konzept für den gemeinsamen Unterricht mit Kindern und Jugendlichen mit und ohne sonderpädagogischen Förderbedarf zu entwickeln. Nach nunmehr 12 Jahren ist eine fast flächendeckende Versorgung an SPS in RLP erreicht, unklar blieb jedoch, wie es den SPS mit ihrem erweiterten pädagogischen Auftrag der Inklusion geht, wie dieser aussieht und sich in der pädagogischen Praxis umsetzt, da bis dahin noch keine wissenschaftliche Evaluation erfolgt war, die erst im Dezember 2011 vom Projekt ‚GeSchwind‘ unter der Leitung von Jun.-Prof. Dr. D. Laubenstein, Prof. Dr. C. Lindmeier und mit Unterstützung der FÖL K. Seutter-Guthöhrlein und S. Belting aufgenommen werden konnte.

Forschungsziel und -frage

Zielsetzung des Forschungsprojektes ist eine Bestandsaufnahme der pädagogischen Arbeit unter der Fragestellung, wie – aus der Sicht der Akteure – der gemeinsame Unterricht an SPS weiterentwickelt werden kann und muss, um den Anforderungen einer inklusiven Beschulung genügen zu können.

Forschungsphasen und -methoden

I. Erhebungsphase: Um der Subjektperspektive der beteiligten Akteure gerecht zu werden erfolgte in der ersten Erhebungsphase ein qualitativer Forschungseinstieg durch Einzel-/Gruppeninterviews mit 16 Expert/innen (vgl. Glaser/Laudel, 2009). Diese ermöglichten eine kritische Auseinandersetzung mit gegenwärtigen und zukünftigen Fragen der Schulentwicklung und den darin Agierenden aus Sicht der Schulverwaltung und Beratung. Die Auswertung erfolgte mittels der Qualitativen Inhaltsanalyse nach Mayring (2010) und des Softwareprogramms MAXQDA zur computerunterstützten Datenauswertung (Kuckartz/Dresing/Rädiker/Stefer, 2008; Kuckartz/Rädiker, 2010).

II. Erhebungsphase: Die Aspekte der II. Erhebungsphase fokussierten sich auf eine quantitative Ermittlung von Grunddaten und auf eine Flächenerhebung aller Schwerpunktschulen durch Onlinebefragung (deskriptiv-statistisch) zu Themen wie Vorbereitung zur Schwerpunktschule, Teamarbeit, Unterrichtsentwicklung, Grenzen und Chancen der gemeinsamen Unterrichtung. Zum Zeitpunkt der Erhebung konnten 210 Schwerpunktschulen (Ausschöpfungsquote 70,5%) online befragt werden. Die Befragungs-Software ‚Unipark‘ diente im Rahmen des Reportings zur Erstellung von Grafiken und Splittings für eine vergleichende Analyse.

III. Erhebungsphase – Qualitative Gruppendiskussionen: Die Perspektive der Lehrkräfte wird seit Mai 2013 in Gruppendiskussionen an 28 SPS ermittelt. Die repräsentative Fallauswahl dient der Darstellung der Heterogenität des Forschungsfeldes. Die Diskussionen ermöglichen dem Forschungsteam den Zugang zu „kollektiven Orientierungen, Wissensbestände und Werthaltungen ...“ (Pryborski; Wohlrab-Sahr, 2009, 106). Die Auswertung erfolgt nach der dokumentarischen Methode (vgl. Bohnsack, 2006).

Erste Ergebnistendenzen: Der zirkuläre Auswertungsprozess ist bislang für keine Phase aufgrund des rekursiven Forschungsdesigns völlig abgeschlossen. Tendenzen weisen jedoch bereits jetzt darauf hin, dass sich die Perspektiven der im Feld Agierenden bzgl. möglicher Problemkonstellationen im Kontext der inklusiven Schulentwicklung in RLP unterscheiden, so dass Gestaltungsmöglichkeiten der Inklusion als schulisches Qualitätsziel unterschiedlich bewertet werde. Weiterhin zeigt sich, dass die durch unterschiedliche Datensplittings erhaltenen Ergebnisse anders geartete Problemkonstellationen (beispielsweise im Bereich Teamarbeit, Unterrichtsentwicklung oder Akzeptanz von Schüler/innen mit unterschiedlichem sonderpädagogischen Förderbedarf) in den Fokus der Betrachtung rücken, die sich auf die Gelingensbedingungen des gemeinsamen Unterrichts an Schwerpunktschulen in Rheinland-Pfalz auswirken und bei der Überwindung derselben zukünftig berücksichtigt werden müssen.



## **Übergänge in den Beruf: Ökonomische, arbeitsmarkttheoretische, soziologische und sozialisationstheoretische Perspektiven**

Chair(s): Susan Seeber (Georg-August-Universität Göttingen), Reinhold Nickolaus (Universität Stuttgart)

Diskutant/in: Stephan Schumann (Universität Konstanz)

Trotz rückläufiger Ausbildungsnachfrage deuten die anhaltenden Strukturverschiebungen zwischen den Sektoren des Berufsbildungssystems auf nach wie vor gravierende Passungsprobleme zwischen Nachfragern nach Ausbildungsplätzen und den Anbietern hin. So mündeten 2011 einerseits noch rund 30% der Neuzugänge in die berufliche Bildung in das Übergangssystem ein; andererseits blieb eine nicht unbedeutende Zahl an Ausbildungsstellen unbesetzt. Hier kann eine fehlende regionale Mobilität ebenso Ursache sein wie eine unzureichende Passung von Anforderungen an den Arbeitsplätzen und Berufsvorstellungen sowie Qualifikationen der Bewerber. Analysen von Paneldaten zeigen, dass vor allem vom Konjunkturverlauf – bei Kontrolle sonstiger Merkmale – ein entscheidender Einfluss auf die Übergangsprozesse ausgeht, denn es haben insbesondere Absolventen in strukturschwachen Regionen erhebliche Probleme im Ausbildungszugang (Deutsches Jugendinstitut, 2013).

Übergänge aus der allgemeinbildenden Schule in Ausbildung sind aus der Perspektive der Individuen in gesellschaftliche und institutionelle Strukturen eingebettete Statuswechsel, die sozial normiert und mit einem individuell zu bewältigenden Wechsel von Identitätssegmenten verbunden sind (Kutscha, 1991, S. 117). Für die Erklärung von Bildungsentscheidungen von Individuen werden humankapitaltheoretische Überlegungen und mikrosoziologische Ansätze (z. B. Boudon 1974) sowie Werterwartungstheorien (Esser, 1999) herangezogen, die für Übergangsprozesse im allgemeinbildenden Schulsystem gut, für die berufliche Bildung jedoch unzureichend belegt sind.

Allerdings hängt der Erfolg der Ausbildung nicht nur von einer qualitativ guten Berufswahl bzw. einer erfolgreicher Selektionsstrategie ab, sondern auch von einer guten Vorbereitung der Auszubildenden in der abgebenden Schule und von einer guten Einführung in die berufliche Grundbildung (Dietrich, 2008). Im Unterschied zur Selektion der Ausbildungsbetriebe (vgl. Connelly et al., 2012 zu signaltheoretischen Ansätzen) wurden die Bedingungen der beruflichen Sozialisation bei Eintritt in die berufliche Grundbildung bisher kaum untersucht.

Aufgrund der hohen Differenziertheit des beruflichen Bildungssystems zeigen sich erwartungskonform zudem innerhalb der Angebotstypen berufsvorbereitender Maßnahmen einer Region erhebliche Differenzen in den Eingangsvoraussetzungen nach Berufsfeld.

Ferner sind bisher die Aussagemöglichkeiten zu den Effekten des Übergangssystems stark restringiert. Es besteht nach wie vor ein erheblicher Forschungsbedarf zu den Mechanismen der Ausbildungseinmündung und des Bildungsverlaufs. Insbesondere gilt es die komplexen Wechselwirkungen von Marktkräften und institutionellen Gegebenheiten auf der einen Seite als auch individuellen Ressourcen und beruflichen Entscheidungsprozessen von Jugendlichen auf der anderen Seite näher zu untersuchen, um Erklärungsmodelle zu generieren und bildungspolitischen wie auch pädagogischen Handlungsbedarf aufzuzeigen.

Im Symposium sollen Probleme des Übergangs in eine Berufsausbildung aus interdisziplinärer Perspektive diskutiert werden, indem arbeitsmarkttheoretische, bildungsökonomische und sozialisationstheoretische sowie pädagogische Perspektiven auf diese Prozesse eingenommen werden.

Boudon, Raymond (1974): Education, Opportunity, and Social Inequality. New York: Wiley.

Connelly, B. L., Certo, S. T. & Ireland, R. D./ Reutzell, C. (2011). Signaling Theory: A Review and Assessment. Journal of Management 37(1), S. 39-65.

Deutsches Jugendinstitut (2013). Kinder- und Jugendhilfe in neuer Verantwortung. Materialien zum 14. Kinder- und Jugendbericht. Hans Dietrich: Übergang aus Ausbildung in Beschäftigung für die Jahre 1977 bis 2003 unter besonderer Berücksichtigung von Konjunktur und Demografie. München: Deutsches Jugendinstitut. [http://www.dji.de/14\\_kjb/14-KJB-Expertise-Dietrich.pdf](http://www.dji.de/14_kjb/14-KJB-Expertise-Dietrich.pdf) [aufgerufen am 27.10.2013].

Dietrich, H. (2012). Youth Unemployment in Europe. Theoretical Considerations and Empirical Findings. Berlin (Friedrich Ebert Stiftung). <http://www.fes.de/cgi-bin/gbv.cgi?id=09227&ty=pdf>;

[aufgerufen am 27.10.2013] Esser, Hartmut (1999): Soziologie: Spezielle Grundlagen. Band 1: Situationslogik und Handeln. Frankfurt am Main: Campus.

Kutscha, Günther (1991): Übergangsforschung: zu einem neuen Forschungsbereich. In: Beck, Klaus & Kell, Adolf (Hrsg.), Bilanz der Berufsbildungsforschung. Stand und Zukunftsperspektiven.

Weinheim: Deutscher Studien Verlag, 113-155.

Reineke, U. (1986): Segmentationstheorien des Arbeitsmarktes und Beschäftigungsniveau Eine Auseinandersetzung mit der theoretischen und empirischen Relevanz der Entkopplungshypothese.

Frankfurt/M., Bern, New York: Peter Lang.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G168

## **Qualifikationsentwicklung und Übergänge: Die letzten zwei Jahre in deutschen Hauptschulen**

### Theoretischer Hintergrund

Die bildungsökonomische Perspektive (Franz, 2013, Kapitel 3) legt nahe, dass die Entwicklung des Humankapitals den Arbeitsmarkterfolg bestimmt. Neben den kognitiven Fähigkeiten wird in der neueren ökonomischen Literatur auch die Entwicklung der nicht-kognitiven Fähigkeiten eine wichtige Rolle zugeschrieben (Cunha & Heckman, 2007). Besondere Maßnahmen der Berufsorientierung in der Hauptschule stellen auf die Ausbildung kognitiver und nicht-kognitiver Fähigkeiten ab, die den Übergang in eine Berufsausbildung befördern (Solga et al., 2012).

### Fragestellung

Die Stärkung der Ausbildungsfähigkeit von schwächeren Schüler(innen) und die Optimierung des Übergangs von Schule in Beruf bleibt auch angesichts sinkender Schülerzahlen eine große gesellschaftliche Herausforderung. Unsere Längsschnittstudie Fitzenberger und Lickleder (2013) führt wiederholte Befragungen für zwei Jahrgänge in den Klassen 8 und 9 der Freiburger Hauptschulen (heute: Werkrealschulen) während der Jahre 2008 bis 2010 durch. Ziel der Untersuchung ist die Beschreibung der Entwicklung der allgemeinen kognitiven Fähigkeiten und des Standes der Berufsorientierung. Weiterhin konnte der Übergang nach Abschluss der 9. Klasse beobachtet werden. Es sollen die empirischen Zusammenhänge zwischen den Schülercharakteristika sowie den Startbedingungen am Ende der Klasse 7 beschrieben und darüber hinaus die Entwicklung wichtiger Indikatoren der Entwicklung der Schüler(innen) im Hinblick auf schulische Leistungen und die Berufsorientierung sowie des Übergangs nach Abschluss der Klasse 9 aufgezeigt werden.

### Methode

Basis der Studie sind Befragungen in den vier Schulhalbjahren der Klassen 8 und 9 in allen Freiburger Hauptschulen für die zwei Jahrgänge, die im Jahr 2007 oder im Jahr 2008 mit der Klasse begonnen haben. Es werden sowohl die Schüler(innen) als auch Lehrkräfte und Eltern befragt. Der Schwerpunkt der Befragungen liegt auf der Erfassung der Maßnahmen der Berufsorientierung und der Entwicklung der Berufsorientierung. Die Erhebungen umfassen zudem die Notenentwicklung in den Fächern Deutsch und Mathematik ab dem Jahreszeugnis in Klasse 7, die Teilnahme am Zusatzunterricht als Vorbereitung für weiterführende Schulen und den Übergang nach Abschluss der Klasse 9. Es werden ökonometrische multivariate Analysen durchgeführt.

### Ergebnisse

Die Studie zeigt, dass nur ein kleiner Teil der Schüler(innen) nach Abschluss der Klasse 9 direkt eine berufliche Ausbildung startet. Ein Hauptgrund hierfür ist, dass die Schüler(innen) mit besseren schulischen Leistungen in Klasse 8 und 9 am Zusatzunterricht teilnehmen, um sich damit die Möglichkeit des Besuchs einer weiterführenden Schule zu eröffnen – mit der Aussicht auf eine Verbesserung ihrer zukünftigen Chancen am Arbeitsmarkt. Leistungsschwächere Schüler(innen) gehen eher in das Übergangssystem, als dass ihnen der direkte Übergang nach Klasse 9 in eine Berufsausbildung gelingt. Einen Wechsel in das Übergangssystem beobachten wir vor allem bei männlichen Schülern mit Migrationshintergrund. Die Noten am Ende von Klasse 7 sind mitbestimmend für den Zugang zum Zusatzunterricht während Klasse 8 und 9 und erweisen sich darüber hinaus als starker Prädiktoren für den Übergang nach Klasse 9.

### Diskussion

Der direkte Übergang in eine berufliche Ausbildung nach Klasse 9 ist für Hauptschüler(innen) in Freiburg in den Jahren 2009 und 2010 die große Ausnahme. Die leistungsstärkeren Schüler(innen) scheinen von vornherein darauf abzustellen, nach Klasse 9 eine weiterführende allgemeine oder berufliche Schule zu besuchen, um ihre langfristigen Arbeitsmarktchancen zu erhöhen. Den leistungsschwächeren Schüler(innen) gelingt meist nicht der direkte Übergang in eine berufliche Ausbildung, trotz intensiver Maßnahmen der Berufsorientierung, einer Verbesserung der schulischen Leistungen und einer merklichen Verbesserung der Berufsorientierung bei vielen Schüler(innen).

Cunha, F. und J.J. Heckman (2007). The technology of skill formation. *American Economic Review*, 97(2), 31–47.

Fitzenberger, B. und S. Lickleder (2013). Skill Formation and Transitions: The Last Two Years in a German Lower Track Secondary School. Unveröffentlichte Arbeit, Albert-Ludwigs Universität Freiburg.

Franz, W. (2013). *Arbeitsmarktökonomik*. Springer Verlag, 8. Auflage, Heidelberg.

Solga, H., Baas, M. und B. Kohlrausch (2012). Mangelnde Ausbildungsreife – Hemmnis bei der Lehrstellensuche von Jugendlichen mit Hauptschulabschluss? WZBrief Bildung 19/2012, Berlin.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G168

## **Betriebliche Strategien zur Einführung von Berufslernenden**

Ein entscheidendes Erfolgskriterium für die berufliche Sozialisation bilden soziale Verhaltensweisen und Aufgabenerfüllung von Jugendlichen im Lehrbetrieb. Studien zeigten, dass für die erfolgreiche Ausbildung im Lehrbetrieb nicht die Leistungen der Schüler/innen in einzelnen Fächern der abgebenden Volksschule im Zentrum stehen, sondern ihre Fähigkeit, mit Berufsbildenden und Betriebsmitarbeitenden zu kooperieren, sich im Betrieb zu integrieren sowie ihre Fähigkeit, die geforderten beruflichen Anforderungen zu erfüllen (Neuenschwander, 2013/im Druck). Ausgehend von der arbeitspsychologischen Forschung zum Stellenwechsel (Kammeyer-Mueller & Wanberg, 2003) wurde ein Konzept entwickelt, wonach die Zufriedenheit mit der Berufsausbildung von der Bewältigung der beruflichen Aufgaben und der sozialen Integration in den Lehrbetrieb abhängt (vgl. Neuenschwander & Gerber, 2013/im Druck). Es stellt sich die Frage, mit welchen Strategien der aufnehmende Lehrbetrieb die soziale Integration, die Bewältigung von beruflichen Aufgaben und die berufliche Zufriedenheit sechs Monate nach Beginn der beruflichen Grundbildung beeinflusst?

Methode: Zur Überprüfung der Frage bzw. Hypothesen wurde eine Stichprobe mit 550 Jugendlichen verwendet, die im Juni 2011, am Ende ihres letzten Schuljahrs, einen Eingangsfragebogen online ausgefüllt hatten (51% Frauen, Durchschnittsalter 16 Jahre, 9% Ausländeranteil). Zwischen 494 und 432 dieser Jugendlichen füllten zwischen August und Dezember im ersten Lehrjahr jeden Monat einen kurzen Online-Fragebogen aus. 82% der Jugendlichen (N=451), die den Eingangsfragebogen ausgefüllt hatten, füllten zusätzlich den online Fragebogen in der Schlusserhebung aus. Wir fanden keine systematischen Rücklaufverzerrungen zwischen den einzelnen Erhebungszeitpunkten der Längsschnittstichprobe.

Begleitend zur Schlusserhebung wurden die Berufsbildner/innen der involvierten Jugendlichen gebeten, einen Fragebogen auszufüllen. Von den 350 angefragten Berufsbildner/innen füllten 244 den Fragebogen aus (43% Frauen mit durchschnittlich 16 Jahren Berufserfahrung).

Es wurden weitgehend standardisierte Fragebogen mit bewährten Items eingesetzt, die faktorenanalytisch zu reliablen Faktoren gruppiert werden konnten. Ergebnisse: Die Korrelationen zwischen (a) den strukturierten Sozialisationsstrategien, (b) dem Feedback und (c) der Beziehung zwischen Berufsbildner/in und Berufslernenden aus der Perspektive der Berufsbildner/innen und der Lernenden sind erstaunlich gering. Aufgrund dieser drei Konzepte, erfasst aus beiden Perspektiven, wurden mit einer hierarchischen Clusteranalyse die Jugendlichen gruppiert. Die zwei-Cluster-Lösung zeigte überzeugende Gütekriterien. Die soziale Integration in den Lehrbetrieb, die wahrgenommene Bewältigung der beruflichen Aufgaben sowie die Zufriedenheit mit dem Lehrbetrieb nahmen kontinuierlich und signifikant vom 9. Schuljahr über die Zwischenerhebungen bis zur Schlusserhebung sechs Monate nach Lehreintritt ab. Es wurde aber kein Einbruch der Werte gefunden (Praxisschock). Die Veränderung dieser Variablen wurde mit Varianzanalysen mit Messwiederholung mit dem Faktor betriebliche Sozialisationstaktik (Cluster) zu erklären versucht. Zweifaktorielle Varianzanalysen mit Messwiederholung mit den Faktoren betriebliche Sozialisationstaktik (Cluster) sowie dem medianhalbierten Faktor Selbstregulation im schulischen Unterricht aus der Schüler-Eingangsbefragung ergab signifikante dreifach-Interaktionen: Bei Jugendlichen mit hohen selbstregulierenden Fähigkeiten, die mit positiven Sozialisationstaktiken im Betrieb konfrontiert waren, entwickelte sich die soziale Integration in den Lehrbetrieb sowie die Zufriedenheit mit dem Lehrbetrieb positiver als in den anderen Gruppen. Bei der Bewältigung der beruflichen Aufgaben wurde kein signifikanter Interaktionseffekt gefunden.

Diskussion: Die Ergebnisse zeigen, dass sich betriebliche Sozialisationstaktiken bei Jugendlichen mit guten Selbstregulationsfähigkeiten, die neu in die Berufslehre treten, positiv auswirken. Die berufliche Sozialisation vollzieht sich nur dann positiv, wenn sowohl der Betrieb als auch der Jugendliche günstige Voraussetzungen schaffen.

Jones, G. R. (1986). Socialization tactics, self-efficacy, and newcomers' adjustment to organizations. *Academy of Management Journal*, 29, 262-279.

Neuenschwander, M. P. (2013/im Druck). Selektionsentscheidungen beim Übergang in die Sekundarstufe I und in den Arbeitsmarkt im Vergleich. In M. P. Neuenschwander (Ed.), *Selektion in Schule und Arbeitsmarkt*. Zürich: Rüegger.

Neuenschwander, M. P. & Gerber, M. (2013/im Druck). Schulische Vorbereitung auf die berufliche Sozialisation im Lehrbetrieb. *Unterrichtswissenschaft*.

Kammeyer-Mueller, J. D. & Wanberg, C. R. (2003). Unwrapping the organizational entry process: Disentangling multiple antecedents and their pathways to adjustment. *Journal of Applied Psychology*, 88(5), 779-794.

## **Ge- und Misslingensbedingungen des Übergangs von der Schule in die postobligatorische Ausbildung – Ergebnisse der Schweizer Jugendlängsschnittstudie TREE**

Der Übergang von der Schule in die postobligatorische Ausbildung ist eine entscheidende Schnittstelle im Bildungsverlauf junger Erwachsener. Heutzutage ist diese kritische Lebensphase länger, anforderungsreicher, unübersichtlicher und risikoreicher geworden (Bergman et al., 2011, 2012). Wie vollziehen junge Menschen in der Schweiz diesen Übergang? Was sind Ge- und Misslingensbedingungen des Übergangs? Welche Rolle spielen dabei individuelle, institutionelle und sozialräumlich-strukturelle Merkmale?

TREE („Transitionen von der Erstausbildung ins Erwerbsleben“) ist in der Schweiz die erste nationale Längsschnittuntersuchung zum Übergang von der Schule ins Erwerbs- und Erwachsenenleben. Im Zentrum der Untersuchung stehen die Ausbildungs- und Erwerbsverläufe Schweizer Jugendlicher nach dem Ende der Pflichtschulzeit (Jahrgangsstufe 9). Die TREE-Stichprobe umfasst mehr als 6.000 Jugendliche, die an PISA 2000 teilnahmen und im selben Jahr aus der obligatorischen Schulpflicht entlassen wurden. Über nunmehr neun Messzeitpunkte (2001-2007 jährlich, 2010, 2014) wurden die Jugendlichen anhand von standardisierten Erhebungsinstrumenten schriftlich und telefonisch zu ihren Ausbildungs- und Erwerbsverläufen sowie weiteren damit in Zusammenhang stehenden Kontextmerkmalen befragt. Die Stichprobe ist national und sprachregional repräsentativ für rund 80.000 Schulabgänger aus dem Jahr 2000. Dem Stichprobenschwund wird durch eine entsprechende Längsschnittgewichtung Rechnung getragen (Sacchi, 2011). Die Daten erlauben Analysen aus erziehungswissenschaftlicher, soziologischer, bildungsökonomischer, psychologischer und interdisziplinärer Perspektive.

Die Ergebnisse zeigen zunächst deskriptiv, dass der Eintritt in eine weiterführende Allgemein- oder Berufsausbildung der Sekundarstufe II (1. Schwelle) für viele Jugendliche ein Nadelöhr darstellt und für ein Viertel der beobachteten Kohorte mit Wartezeiten in Zwischenlösungen und Brückenangeboten verbunden ist (Keller et al., 2010). Die postobligatorischen Ausbildungsverläufe der Jugendlichen sind häufig alles andere als gradlinig, sondern stark verzweigt. Nur gut die Hälfte der Kohorte weist eine „Standard-Normalbiographie“ auf: Ihnen gelingt es, direkt in eine weiterführende Berufs- oder Allgemeinbildung der Sekundarstufe II einzusteigen, diese bruchlos und in der Regelzeit zu durchlaufen. Multivariat erweisen sich in binär-logistischen Regressionsanalysen (Hupka-Brunner et al., 2011) das Geschlecht (zugunsten der Männer), eine hohe PISA-Lesekompetenz, der Besuch von Schultypen mit erweiterten Anforderungen, gute Schulnoten und eine hohe soziale Herkunft (HISEI) als bedeutsame Prädiktoren für einen gelingenden Einstieg in eine nachobligatorische Ausbildung. Außerdem haben strukturelle Merkmale des Bildungssystems einen eigenständigen Rationierungseffekt auf den Übergang in die Sekundarstufe II: Jugendliche aus Kantonen mit einem hohen Anteil an Schulen mit Grundanforderungen oder aus städtischen Gebieten haben ceteris paribus niedrigere Eintrittschancen in eine zertifizierende postobligatorische Ausbildung. Regionale Disparitäten deuten darüber hinaus auf unterschiedliche Zugangschancen für Jugendliche aus den drei verschiedenen Sprachregionen in der Schweiz hin.

Bergman, M. M., Hupka-Brunner, S., Keller, A., Meyer, T. & Stalder, B. E. (Hrsg.). (2011). Transitionen im Jugendalter. Ergebnisse der Schweizer Längsschnittstudie TREE. Zürich: Seismo.

Bergman, M. M., Hupka-Brunner, S., Meyer, T. & Samuel, R. (Hrsg.). (2012). Bildung – Arbeit – Erwachsenwerden. Ein interdisziplinärer Blick auf die Transition im Jugend- und jungen Erwachsenenalter. Wiesbaden: Springer VS.

Hupka-Brunner, S., Meyer, T., Stalder, B. E. & Keller, A. (2011). PISA-Kompetenzen und Übergangswegen: Ergebnisse der Schweizer TREE-Studie. In: E. M. Krekel & T. Lex (Hrsg.), Neue Jugend, neue Ausbildung? Beiträge aus der Jugend- und Bildungsforschung (S. 173-188). Bonn: BIBB.

Keller, A., Hupka-Brunner, S. & Meyer, T. (2010). Nachobligatorische Ausbildungsverläufe in der Schweiz: Die ersten sieben Jahre. Ergebnisübersicht des Jugendlängsschnitts TREE, Update 2010. Basel: TREE.

Sacchi, S. (2011). TREE-Längsschnittgewichtung: Konstruktion und Anwendung. Dokumentation der Erhebungswellen 2000 bis 2010. Basel & Zürich: TREE & cue sozialforschung.

Susan Seeber (Georg-August-Universität Göttingen), Anne Windaus (Universität Stuttgart), Sabine Gauch (Universität Stuttgart), Janna Kosanke (Georg-August-Universität Göttingen), Constanze Gruschinski (Georg-August-Universität Göttingen), Reinhold Nickolaus (Universität Stuttgart)

## **Regionaler ökonomischer Kontext und Komposition von Klassen im Übergangssystem**

Der Anteil an Jugendlichen im Übergangssystem ist in den letzten Jahren deutlich gesunken, beläuft sich jedoch immer noch auf knapp 267.000. Gleichzeitig ist die Zahl unbesetzter Ausbildungsstellen 2012 auf mehr als 30.000 angestiegen bei einem gleichzeitigen Anstieg unversorgter Bewerber (BMBF, 2013). Diese Situation verweist auf deutliche Passungsprobleme zwischen Ausbildungsplatznachfragern und Ausbildungsanbieter, die sowohl mit regionalen als auch berufsbezogenen Merkmalen konfundiert sind. Wie Analyse zeigen, ist der demografische lediglich ein moderierender Einflussfaktor, der im Vergleich zur regionalen konjunkturellen Lage und zu individuellen Eingangsvoraussetzungen der Jugendlichen ein geringeres Gewicht besitzt (Deutsches Jugendinstitut, 2013). In den letzten Jahren hat sich durch den Rückbau des Übergangssystems auch die Zusammensetzung der Jugendlichen im Übergangssystem verändert. Insbesondere reduzierte sich der Anteil an Jugendlichen mit günstigeren Eingangsvoraussetzungen, jedoch besteht nach wie vor ein substantieller Anteil schwer vermittelbarer Jugendlicher, der trotz günstigerer Übergangsbedingungen keinen Ausbildungsplatz findet (BMBF, 2013). Ob diese Annahme zutrifft und vor allem welche Auswirkungen solche Schwankungen am Ausbildungsstellenmarkt tatsächlich auf die Eingangsmerkmale der Jugendlichen haben, ist weitgehend ungeklärt. Zudem dürften die Eingangsvoraussetzungen nicht nur von der jeweiligen regionalen Konjunkturlage und Arbeitsmarktstruktur abhängig sein, sondern auch von Stadt-Land-Disparitäten, Mobilitätsbereitschaften der Jugendlichen und der dokumentierten Leistungsvarianz zwischen den Bundesländern (vgl. z. B. Pant et al., 2013).

Forschungsfragen: Das Hauptziel des Beitrags ist es, Bedingungen und Prozesse bei der Einmündung in die berufliche Bildung unter Berücksichtigung regionaler Kontextbedingungen aufzuzeigen. Dabei sollen makroanalytische Einflussgrößen wie die regionale demografische und wirtschaftliche Situation mit mikroanalytischen Analysen zum Übergang in eine Berufsausbildung aufeinander bezogen werden. In diesem Zusammenhang wird auf arbeitsmarkt- und signaltheoretische Ansätze (Conelly et al. 2012) sowie auf mikrosoziologische Theorien (Boudon, 1974) zurückgegriffen. Geklärt wird insbesondere, welche Zusammenhänge sich zwischen individuellen kognitiven, motivationalen sowie sozialen und mobilitätsbezogenen Merkmalen der Jugendlichen und regionalen Faktoren auf die Ausbildungschancen von Jugendlichen zeigen.

Methodisches Vorgehen: Für die Analysen steht ein Datensatz des Übergangssystems „Individuelle Bildungsverläufe im Übergangssystem: zur Wechselwirkung von individuellen und sozialen Merkmalen und institutionellen Bedingungen“ aus Baden-Württemberg (N= 1.116) und Niedersachsen (N=2.671) zur Verfügung. Bei Eintritt in die verschiedenen Maßnahmen des Übergangssystems wurden die Jugendlichen retrospektiv zur Schulbiografie und zu ihren Strategien der Berufsorientierung und Ausbildungsplatzsuche befragt. Darüber hinaus wurden die Kompetenzen in Mathematik, Lesen und allgemeine kognitive Fähigkeiten (CFT 20) sowie motivationale Dispositionen und selbstwirksamkeitsbezogene Fähigkeiten erhoben. In Abhängigkeit vom Berufsbereich wurde zudem berufsrelevantes Vorwissen in den Berufsfeldern Wirtschaft und Verwaltung, Hauswirtschaft und in gewerblich-technischen Fachrichtungen erfasst. Kurz vor Ablauf der berufsvorbereitenden Maßnahme erfolgte eine erneute Befragung und Testung der Jugendlichen.

Befunde: Dazu wurden die Schülerdaten auf der Ebene der BA-Bezirke der beiden Bundesländer aggregiert und es wurden regionale Wirtschaftsdaten (Arbeitslosigkeit, Angebots-Nachfrage-Relation, BIP) hinzugespielt. Die Analysen sind zwar noch nicht vollständig abgeschlossen, es zeigt sich jedoch bereits nach ersten Auswertungen, dass ein Creaming-Effekt zuungunsten der Jugendlichen in Übergangsmaßnahmen in ökonomisch benachteiligten Regionen vorliegt. Dies führt letztlich zu deutlichen Unterschieden der Zusammensetzung der Lerngruppen mit Blick auf kognitive Eingangsmerkmale der in die gleichen Angebotstypen einmündenden Jugendlichen.

Boudon, Raymond (1974): Education, Opportunity, and Social Inequality. New York: Wiley.

Connelly, B. L., Certo, S. T. & Ireland, R. D./ Reutzell, C. (2011). Signaling Theory: A Review and Assessment. Journal of Management 37(1), S. 39-65.

Deutsches Jugendinstitut (2013). Kinder- und Jugendhilfe in neuer Verantwortung. Materialien zum 14. Kinder- und Jugendbericht. Hans Dietrich: Übergang aus Ausbildung in Beschäftigung für die Jahre 1977 bis 2003 unter besonderer Berücksichtigung von Konjunktur und Demografie. München: Deutsches Jugendinstitut. [http://www.dji.de/14\\_kjb/14-KJB-Expertise-Dietrich.pdf](http://www.dji.de/14_kjb/14-KJB-Expertise-Dietrich.pdf) [aufgerufen am 27.10.2013].

Dietrich, H. (2012). Youth Unemployment in Europe. Theoretical Considerations and Empirical Findings. Berlin (Friedrich Ebert Stiftung). <http://www.fes.de/cgi-bin/gbv.cgi?id=09227&ty=pdf>;

[aufgerufen am 27.10.2013] Esser, Hartmut (1999): Soziologie: Spezielle Grundlagen. Band 1: Situationslogik und Handeln. Frankfurt am Main: Campus.



Chair(s): Klaus-Peter Wild (Universität Regensburg), Michaela Gläser-Zikuda (Friedrich-Schiller-Universität Jena)

Diskutant/in: Matthias Nückles (Albert-Ludwigs-Universität Freiburg)

In den vergangenen Jahren ist es Dank eines zunehmenden Problembewusstseins und verschiedener Fördermöglichkeiten im nationalen und internationalen Kontext zu einer verbesserten Forschungslage für den Hochschulbereich gekommen. Dies betrifft zugleich psychologische, erziehungswissenschaftliche, didaktische und soziologische Aspekte des Lehrens und Lernens an Hochschulen.

Diese Entwicklung war aus verschiedenen Gründen längst überfällig: Zum einen hat sich im vergangenen Jahrzehnt eine erhebliche Veränderung der Lehrorganisation und der damit verbundenen Qualitätssicherungsprozesse an unseren Hochschulen ergeben (Stichwort: Bologna-Prozess) ohne dass dies von einer angemessenen Forschungstätigkeit begleitet werden konnte. Zum anderen hat sich gezeigt, dass die direkte Rückbeziehung der hochschulbezogenen Forschung auf die deutlich breitere schulbezogene Forschung nicht nur zahlreiche Synergien schafft, sondern zugleich auch problematische Fehlkonzepte implizieren kann.

Im Bereich der Hochschulforschung sind nicht nur spezifische Fragestellungen etabliert, für die im schulischen Bereich kein direktes Pendant besteht (z.B. Struktur und Qualität von Lehrevaluationen, Prozesse der Studienfachwahlentscheidungen; eignungsdiagnostische Fragestellungen im Zuge der Studiengangszulassung; Fragen der Leistungserfassung im tertiären Bildungsbereich). Wir finden hier zugleich auch Fragestellungen, die im Kontext der spezifischen „Lernumgebung Hochschule“ und der in diesem Bereich spezifisch hohen Lernkompetenzen der Studierenden eine andere Qualität erhalten, die nicht nur empirisch, sondern auch theoretisch einer kontext- und domänenspezifischen Anpassung bedürfen.

Ziel des Symposiums ist es daher, aus Sicht der pädagogischen und psychologischen Forschung einen Beitrag zu einer theoretischen und domänenspezifischen Einordnung der Hochschulforschung zu leisten.

Zudem decken aus forschungsmethodischer Sicht die in diesem Symposium vorgestellten Studien diverse qualitative und quantitative Verfahren in der Hochschulforschung ab.

Aus inhaltlich-thematischer Sicht bündelt das Symposium Arbeiten, die sich einerseits auf die Lehrenden und hochschuldidaktischen Aspekte beziehen, und andererseits die Studierenden, deren Lernregulation und/oder deren Leistungsentwicklung fokussieren:

(a) Hagenauer, Volet & Gläser-Zikuda befassen sich auf der Basis einer qualitativen Studie mit der Frage, welche Bedingungen sich in der universitären Lehre als emotionsrelevant für Hochschullehrende hinsichtlich der Unterstützung von Lernprozessen der Studierenden, aber auch der eigenen Selbstregulation erweisen.

(b) Bei Seifried, Eckert & Spinath steht die Frage im Vordergrund, in welcher Weise sich die Wahrnehmung von optimalen und freiwillig angenommenen Studienaufgaben durch verschiedene Personenvariablen vorhersagen lässt.

Zwei Beiträge des Symposiums befassen sich – wenn auch auf sehr unterschiedliche Weise – mit einer diagnostischen Fragestellung:

(c) Seibert & Watermann berichten über die Entwicklung eines Fragebogenverfahrens zur Erfassung impliziter Schreibüberzeugungen im Kontext des wissenschaftlichen Schreibens.

(d) Wild befasst sich mit dem Problem der Modellierung und der Prognose „natürlicher“ Studienleistungsmaße in universitären Studiengängen.

Dienstag,

04.03.2014,

10:15 - 12:00,

PEG-Gebäude

Raum 1.G191

## **Emotionen und Emotionsausdruck von Hochschuldozierenden in der Lehrerausbildung**

Während LehrerInnenemotionen und deren Regulation im Schulbereich zunehmendes Forschungsinteresse wecken, sind sie im Hochschulsektor nahezu unerforscht. In einer aktuellen Studie konnten Postareff und Lindblom-Ylänne (2011) allerdings aufzeigen, dass Lehren an der Hochschule ebenso eine Tätigkeit darstellt, die von zahlreichen Emotionen begleitet wird. Trigwell's (2012) Ergebnisse belegen des Weiteren, dass positive Emotionen von Dozierenden mit einem studierendenzentrierten Lehrstil einhergehen. Somit ist davon auszugehen, dass Emotionen von Hochschullehrenden sowie deren Kommunikation (= Emotionsexpression) zur Qualität universitärer Lehr- und Lernprozesse beitragen. Basierend auf einem kognitiven Ansatz der Emotionsentstehung (Scherer et al., 2001) kombiniert mit einer sozialpsychologischen Perspektive (Parkinson, 1996), wird in diesem Beitrag zum einen der Frage nachgegangen, welche Bedingungen sich in der universitären Lehre als emotionsrelevant für Hochschullehrende erweisen. Zum anderen wird die Frage aufgeworfen, wie Lehrende diese Emotionen nach außen kommunizieren und welche Funktionen damit einhergehen. Emotionsausdruck wird hierbei als ein Teilbereich der allgemeinen Emotionsregulation verstanden (Gross, 2002). Sind Universitätsdozierende in der Lage, ihre Emotionen kompetent zu regulieren, so kann Emotionsausdruck gezielt eingesetzt werden, um spezifische Ziele zu erreichen (Tamir, 2011). Die Angemessenheit und in Folge auch die Wirkung des gezeigten Emotionsausdrucks hängt jedoch vom jeweiligen Kontext ab, d.h. es ist zu erwarten, dass der Emotionsausdruck und das Verständnis darüber, was „passend“ oder „unpassend“ wahrgenommen wird, interindividuell (z.B. kulturell, institutionell und disziplinär) und auch intraindividuell variiert. Zur Überprüfung der genannten Fragestellungen wurde ein explorativer, auf Leitfadeninterviews basierender Zugang gewählt. Um der Kontextspezifität von Emotionen gerecht zu werden, wird in dieser Untersuchung auf Hochschullehrende in der Lehrerbildung fokussiert, da ihnen eine duale Rolle zukommt (= Lernen ermöglichen/Wissen vermitteln und Lehren modellieren; Lunenberg et al., 2007). Die Einflüsse der Rahmenbedingungen des jeweiligen Hochschulsystems und der Kultur auf Emotionen und Emotionsausdruck werden in dieser Untersuchung ebenso berücksichtigt (Markus & Kitayama, 1991), indem Dozierende mit vergleichbarer Berufserfahrung aus zwei Ländern (Deutschland – Australien) verglichen werden. Im Hinblick auf die situative Abhängigkeit wird der Fokus auf das Lehren in Seminarsettings mit Studierenden im ersten Studienjahr gelegt, wobei typische Lehr- und Interaktionssituationen fokussiert werden. Während sich die Datengewinnung und -analyse des deutschen Dozierendensamples zum Zeitpunkt der Beitragseinreichung noch in der Ausführungsphase befindet, können die Methode und Ergebnisse der australischen Stichprobe bereits näher ausgeführt werden: Es wurden insgesamt 15 australische Lehrende (differenziert nach Fach und Geschlecht; mehr als 3 Jahre Lehrpraxis) in der universitären Lehrerausbildung zu zwei Zeitpunkten befragt. Das erste Interview fokussierte emotionales Erleben und die Emotionsregulation von Hochschullehrenden allgemein, im zweiten Interview sollten Dozierende über spezifische Erlebnisse des aktuellen Studienjahres reflektieren. Die Interviews wurden wortwörtlich transkribiert und entsprechend der Qualitativen Inhaltsanalyse mittels Zusammenfassung und Strukturierung ausgewertet. Die Interviewdaten belegen eine Fülle an positiven und negativen Emotionen beim Lehren und in der Interaktion mit Studierenden. Unterschiedlichste Faktoren tragen zu diesem Emotionserleben bei, vor allem die Erwartungserfüllung bzw. –nichterfüllung bezogen auf das Studierendenverhalten (z.B. Erledigung von Aufgaben → Freude; Zufriedenheit). Im Hinblick auf den Emotionsausdruck stimmten alle Lehrpersonen der Aussage zu, dass positive Emotionen kommuniziert werden sollten, wobei der tatsächlich realisierte Emotionsausdruck von der jeweiligen Persönlichkeit und dem professionellen Selbstbild der Lehrkraft abhängt. Negativer Emotionsausdruck muss dahingegen sehr viel vorsichtiger erfolgen und dem Balanceakt „Authentizität – Angemessenheit“ standhalten. Gelingt ein entsprechender Emotionsausdruck, so können aus Sicht der befragten Lehrpersonen unterschiedlichste Ziele erreicht werden (z.B. der Aufbau und der Erhalt einer positiven Lehrenden-Studierenden-Beziehung) (Hagenauer & Volet, 2013b). Diese Ergebnisse der australischen Interviewstudie werden mit der deutschen Studie nach Abschluss der Datenanalyse kontrastiert, womit die Beantwortung der Fragestellung um eine hochschulsystem- und kulturvergleichende Perspektive erweitert wird.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, 39 (3), 281-291.

Hagenauer, G. & Volet, S. E. (2013b). „I don't hide my feelings, even though I try to": insight into teacher educator emotion display. *Australian Educational Researcher*. Doi: 10.1007/s13384-013-0129-5.

Lunenberg, M., Korthagen, F. & Swennen, A. (2007). The teacher educator as role model. *Teaching and Teacher Education*, 23, 586-601.

Markus, H. R. & Kitayama, S. (1991). Culture and self: implications for cognition, emotion, and motivation. *Psychological Review*, 98 (2), 224-253.

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology*, 87 (4), 663-683.

Postareff, L. & Lindblom-Ylänne, S. (2011). Emotions and confidence within teaching in higher education. *Studies in Higher Education*, 36 (7), 799-813.

Scherer, K. S.; Schorr, A. & Johnstone, T. (2001). *Appraisal processes in emotion research. Theory, methods, research*. Oxford: University Press.

Tamir, M. (2011). The maturing field of emotion regulation. *Emotion Review*, 3 (1), 3-7.

Trigwell, K. (2012). Relations between teachers' emotions in teaching and their approaches to teaching in higher education. *Instructional Science*, 40 (3), 607-621.



Eva Seifried (Ruprecht-Karls-Universität Heidelberg), Christine Eckert (Ruprecht-Karls-Universität Heidelberg), Birigit Spinath (Ruprecht-Karls-Universität Heidelberg)

## **Optionale Lernangebote zum semesterbegleitenden Lernen: Wer nutzt sie und in welchem Zusammenhang steht die Nutzung mit der Entwicklung von Motivation und Lernerfolg?**

Theoretischer Hintergrund: Häufig kann in großen universitären Veranstaltungen nur eine minimale Lehrende-Lernende-Interaktion stattfinden. Daher ist es eine Herausforderung, auch in solchen Lehrveranstaltungen Leistungen mit angemessenem Aufwand zu kreditieren und gleichzeitig die Studierenden zu kontinuierlichem Lernen zu motivieren. Vorherige Studien haben gezeigt, dass eine semesterbegleitende Auseinandersetzung mit den Inhalten zu höherem Lernerfolg führen kann (vgl. Spinath & Seifried, 2012). Ein weiterer Faktor, der sich positiv auf Motivation und Leistung auswirken kann (vgl. Katz & Assor, 2007; Patall, Cooper, & Robinson, 2008), ist die Entscheidungsfreiheit im Sinne der Self-Determination Theory (Ryan & Deci, 2000). Im Zuge der Bologna-Reform und der damit einhergehenden Modularisierung und Vorstrukturierung der Studieninhalte scheinen sich Bachelor-Studierende in ihren Entscheidungsmöglichkeiten eingeschränkt zu fühlen, was wiederum bedeutsam für die geringere Studienzufriedenheit der Bachelor- gegenüber den Diplom-Studierenden sein könnte (Sieverding, Schmidt, Obergfell, & Scheiter, 2013). In einer Pflicht-Vorlesung für Bachelor-Studierende („Einführung in die Pädagogische Psychologie“) wurde daher ein neues Lehr-Lern-Format getestet, um sowohl semesterbegleitendes Lernen zu fördern als auch Wahlfreiheiten zu lassen und den Arbeitsaufwand für die Lehrenden in einem akzeptablen Umfang zu halten. Die Studierenden können sich nach jeder Sitzung durch das Verfassen von sog. Statements mit den Vorlesungsinhalten vertieft auseinandersetzen. Zwei Statements gegen Ende des Semesters sind notenrelevant. Hierzu gibt es zwei Blöcke à zwei Sitzungen, zu denen die Studierenden jeweils ein benotungsrelevantes Statement einreichen müssen. Zuvor können die Studierenden das Verfassen von Statements üben, indem sie sog. optionale Statements verfassen: Reichen die Studierenden solche ein, erhalten sie eine Rückmeldung durch Tutoren/Tutorinnen (Punktbewertung und kurzes inhaltliches Feedback). Viele Studierende nehmen jedoch nur die verpflichtenden Einreichungen vor.

Fragestellung: Daraus leiten sich zwei Fragen ab: 1) Unterscheiden sich Studierende, die optionale Statements einreichen, schon zu Beginn des Semesters in motivationalen sowie subjektiven und objektiven leistungsbezogenen Variablen von Studierenden, die keine optionalen Statements einreichen? 2) Unterscheiden sich Studierende, die optionale Statements einreichen, in ihrer motivationalen und leistungsbezogenen Entwicklung von Studierenden, die keine optionalen Statements einreichen?

Methode: Basierend auf dem dargestellten Veranstaltungskonzept wurden die Studierenden des Wintersemesters 2011/2012 (N = 108) nachträglich in zwei Gruppen unterteilt: 1) Studierende, die mindestens ein optionales Statement einreichten (n = 33), und 2) Studierende, die keine optionalen Statements einreichten (n = 75). Die motivationalen Variablen (Werte, Erwartungen) und das subjektive Wissen wurden mittels standardisierter Fragebögen zu zwei Messzeitpunkten (erste und letzte Vorlesungssitzung) erfasst. Darüber hinaus wurden drei objektive Leistungsmaße erhoben: Leistung im Vorwissenstest (erreichte Punktzahl), Leistung in der Abschlussklausur (erreichte Punktzahl) und Note in den benotungsrelevanten Statements. Ferner wurde die Abiturnote als Kontrollvariable erfasst. Die Daten wurden mithilfe varianzanalytischer Verfahren ausgewertet.

Ergebnisse: In Bezug auf die Frage, ob sich die Studierendengruppen in ihren motivationalen sowie leistungsbezogenen Voraussetzungen (subjektiv, objektiv) unterscheiden, ergaben sich außer in der Abiturdurchschnittsnote keine signifikanten Unterschiede. In Bezug auf die Frage, wie sich Motivation und Leistung für die Studierendengruppen über das Semester entwickeln, ergaben sich auf motivationaler Ebene keine signifikanten Entwicklungsunterschiede. Auf Leistungsebene zeigte sich eine signifikante Interaktion zwischen der Gruppenzugehörigkeit und der Zeit: Die Studierenden unterschieden sich zu Beginn des Semesters nicht in ihrem objektiven Wissen; allerdings verzeichneten die Studierenden, die optionale Statements eingereicht hatten, einen signifikant höheren Wissenszuwachs als Studierende, die keine optionalen Statements eingereicht hatten. Zudem unterschieden sich die beiden Studierendengruppen in der Benotung der Statements: Die Studierenden, die optionale Statements eingereicht hatten, erzielten signifikant bessere Noten als Studierende, die keine optionalen Statements eingereicht hatten. Diese Ergebnisse blieben auch nach Kontrolle der Abiturnote signifikant. Das oben beschriebene Gesamtbild konnte auch im darauf folgenden Wintersemester repliziert werden. Da diese Untersuchung kein experimentelles Design aufweist, können keine kausalen Schlussfolgerungen gezogen werden. Dennoch zeigen die Ergebnisse, dass kontinuierliches Lernen und Entscheidungsfreiheit auch oder gerade im Hochschulkontext wichtige Faktoren für Leistung darstellen.

Katz, I. & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review*, 19, 429-442. doi: 10.1007/s10648-006-9027-y

Patall, E. A., Cooper H. & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin*, 134, 270-300. doi: 10.1037/0033-2909.134.2.270

Ryan, R. M. & Deci, E. L. (2000). Self-Determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78. doi: 10.1037/110003-066X.55.1.68

Sieverding, M., Schmidt, L. I., Obergfell, J. & Scheiter, F. (2013). Stress und Studienzufriedenheit bei Bachelor- und Diplom-Psychologiestudierenden im Vergleich. Eine Erklärung unter Anwendung des Demand-Control-Modells. *Psychologische Rundschau*, 64, 94-100. doi: 10.1026/0033-3042/a000152

Spinath, B. & Seifried, E. (2012). Forschendes Lehren: Kontinuierliche Verbesserung einer Vorlesung. In M. Krämer, S. Dutke & J. Barenberg (Hrsg.), *Psychologiedidaktik und Evaluation IX* (S. 171-180). Aachen: Shaker.

## ***Erfassung impliziter Schreibüberzeugungen im Kontext des wissenschaftlichen Schreibens***

Die Fähigkeit, wissenschaftliche Texte zu verfassen, zählt zu den zentralen fachübergreifenden Zielen des Hochschulstudiums. Neben inhaltsbezogenem Wissen und dem Wissen darüber, welche Funktionen unterschiedliche Textsorten und -abschnitte erfüllen und welche Regeln es beim Verfassen wissenschaftlicher Texte es zu beachten gilt (Genrewissen), können implizite Schreibüberzeugungen den Schreibprozess beeinflussen. Unter impliziten Schreibüberzeugungen werden komplexe kognitive Prozesse verstanden, die in den Prozess des Schreibens implizit integriert sind und handlungsleitend für die Textproduktion sein können (White & Bruning, 2005). Im Vortrag wird über die Entwicklung und Validierung eines Fragebogens zur Erfassung impliziter Schreibüberzeugungen berichtet, der auf dem Writing Beliefs Inventory (WBI) von White und Bruning (2005) basiert, übersetzt und für den Kontext des wissenschaftlichen Schreibens adaptiert wurde.

Die theoretische Grundlage des WBI liefert das Schreibmodell von Bereiter und Scardamalia (1987a,b). Die unterschiedlichen Schreibperspektiven, die Bereiter und Scardamalia als Knowledge-Telling und Knowledge-Transforming bezeichnen, sind kongruent zu den von White und Bruning (2005) erfassten Schreibkonzepten des Transactional und Transmissional. Bei der Transactional-Perspektive geht es in erster Linie um den Informationstransfer, der sich auf die Darstellung von Sachverhalten begrenzt. Oftmals sind die Texte dadurch gekennzeichnet, dass in ihnen keine eigenständige Argumentationslinie erkennbar ist und die Perspektive des Lesers kaum berücksichtigt wird. Textproduzenten, bei denen die Transactional-Perspektive überwiegt, setzen während des Schreibens eigene Ideen um. Sie sehen den wissenschaftlichen Schreibprozess als Möglichkeit, Texte eigenständig, kritisch und reflektiert zu verfassen. In den verfassten Texten sind zumeist eine klare Argumentationslinie und ein kohärenter Sinnzusammenhang zu erkennen. Während also die Transactional-Perspektive durch geringe (meta-)kognitive Aktivitäten während der Textproduktion gekennzeichnet ist, wird die Transactional-Perspektive durch ein problembewusstes und reflektiertes Verfassen von Texten charakterisiert.

Der Fragebogen wurde in seiner deutschen Übersetzung und Erweiterung dem Kontext des wissenschaftlichen Schreibens angepasst und weist zufriedenstellende psychometrische Kennwerte auf. In einer ersten Studie bei Studierenden der Bildungswissenschaften (N = 175) wurde die faktorielle Struktur ermittelt, eine zweite Erhebung (N = 310) überprüfte die Struktur mittels konfirmatorischer Faktorenanalysen. Die Dimensionen des Originalinventars konnten repliziert werden, wobei die Dimensionen Transactional und Transmissional eine interessante Binnenstruktur aufwiesen: So zeigten sich bei Transactional die Subdimensionen Wissensübermittlung und Orientierung an Quellen, bei Transactional die Subdimensionen Ideengenerierung und Kritisches Prüfen. Korrelative Analysen im Rahmen der konvergenten Validitätsprüfung zeigten u.a., dass die Transactional-Skala mit erhöhten subjektiven Vorstellungen zu Wissen und zum Wissenserwerb (Epistemologische Überzeugungen) sowie erhöhten selbstregulativen Wirksamkeiten beim Schreiben einhergeht. Moderate bis hohe Korrelationen mit anderen relevanten Instrumenten (sprachliches Selbstkonzept, Persönlichkeit, Selbstwirksamkeit) stützten die konvergente Validität. Es wird ein Ausblick auf weitere zurzeit laufende Arbeiten gegeben, in denen die Beziehung zwischen impliziten Schreibüberzeugungen und den Schreibproben von Studierenden analysiert werden. Das Inventar stellt derzeit das einzige deutschsprachige Instrument zur Erfassung impliziter Schreibüberzeugungen dar.

White, M.J. & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology*, 30 (2), 166-189.

Bereiter, C. & Scardamalia, M. (1987a). *The psychology of written composition*. Hillsdale: Erlbaum.

Bereiter, C. & Scardamalia, M. (1987b). Knowledge telling and knowledge transforming in written composition. In: S. Rosenberg (Ed.) *Advances in applied psycholinguistics*, Vol. 2: Reading, writing and language learning, 142-176. Cambridge: Cambridge University Press



## **Intelligenz und Pädagogik? Zur Modellierung und Prognose von Studienleistungen in der Erziehungswissenschaft**

Problemstellung: Mit der Reform der Hochschulstudiengänge wurde nicht selten zugleich eine Ausweitung selektiver Verfahren zur Regelung des Hochschulzugangs vorgenommen. Die konkrete Ausgestaltung geschah und geschieht dabei häufig ohne direkten Bezug auf die nationalen oder internationalen Erfahrungen und wissenschaftlichen Befunde (vgl. z.B. Kuncel, Wee, Serafin, & Hezlett, 2010; Soares 2012; Formazin, Schröders, Köller, Wilhelm, & Westmeyer, 2011). Nicht nur von Laien wird dabei oft die prognostische Validität von Testverfahren oder von Selbstauskünften überschätzt, die üblicherweise erst im Anschluss an vorlaufende Selbstselektionsvorgänge positioniert sind. Legitimität und Fairness insbesondere von Ausschlussentscheidungen erfordern jedoch prinzipiell präzise unterstützende Informationen zur Effektivität und Effizienz des Auswahlverfahrens. Der aktuelle Stand der Theoriebildung und der empirischen Bildungsforschung konzentrierte sich bislang jedoch oft auf wenige spezifische Studienfächer (Trost, 2005) oder war auf die Frage der prognostischen Validität von hochaggregierten Schulabschlussnoten beschränkt (vgl. Hell, Trapmann, & Schuler, 2008; Hell, Trapmann, Weigand, & Schuler, 2007). Im vorliegenden Beitrag soll diese Problemstellung der Prognose in einer besonderen – zweistufigen - Weise aufgegriffen werden.

(1) Dabei wird zunächst geprüft, inwieweit sich die dokumentierten Studienleistungen in einem differenzierten Messmodell der Studienleistungen eines Studiengangs modellieren lassen. Eine solche Vorgehensweise knüpft daran an, dass mit der Einführung der modularisierten Studiengängen auch neue Prüfungsordnungen etabliert wurden, die in komplexer Weise eine Vielzahl an Teilleistungen zu aggregierten Werten (Module, Abschlussnoten) verrechnen. Auch wenn dabei nicht auf die übliche Terminologie der pädagogisch-psychologischen Diagnostik zurückgegriffen wird, so wird hiervon doch eine objektive, reliable und valide Bewertung der im Studium erworbenen Kompetenzen erwartet. Als Ausgangspunkt für die Modellierung von Studienleistungen dient daher die Zuordnung zwischen Lehrveranstaltungen und Modulen, so wie sie in einer Prüfungsordnung definiert ist. Die inhaltliche Interpretation eines Moduls im Sinne eines „latenten Faktors“ hängt dabei natürlich auch von der intendierten Modulkonzeption ab: Nicht selten finden hier thematische Bündelungen, aber auch Bündelungen im Sinne von „theoretisch-fachwissenschaftlichen“ vs. „forschungsmethodischen Kompetenzen“ ihren Ausdruck. Auch Bündelungen mit einer Differenzierung zwischen „berufsnahem Handlungswissen“ vs. „fachwissenschaftlichem Wissen“ sind hier denkbar. Im Anschluss an diese Analysen liegen Studienleistungen auf unterschiedlich aggregierten Ebenen vor: (a) Lehrveranstaltung, (b) Modul/Faktor, (c) Gesamtwert der Studienleistungen.

(2) Mit der Etablierung empirisch begründeter Messmodelle auf der Kriteriumsseite wird die Bestimmung differentieller Prädiktormodelle für jeden der ermittelten Faktoren ermöglicht. Gerade im Hinblick auf die üblicherweise „schwachen“ Prädiktoren (u.a. Studienmotivation) sollten sich hier neue Einsichten bezüglich der Einflussfaktoren auf die Entwicklung von Studienleistungen ergeben.

Methode: Auf der Basis der Prüfungsleistungen von sechs Kohorten (N=798) werden für Prognosezeiträume von einem bis sechs Jahren thematische und sequentielle Abhängigkeiten der Studienleistungen modelliert. Mit zunehmender Studiendauer reduziert sich die Anzahl der Studierenden um jene Personen, die ihr Studium in diesem Fach abbrechen. Voranalysen prüfen daher (a) in welcher Weise eine Selektion durch Studienabbruch erfolgt und (b) in welchem Umfang Kohorteneffekte zu erwarten sind. Zur Prognose der Studienleistungen werden neben der Abiturnote und sozio-demografischen Daten auch Kennwerte zur Intelligenz (PSB), der Konzentrationsfähigkeit (d2), der Studienmotivation und des Lernverhaltens (LIST) herangezogen.

Ergebnisse: Die Ergebnisse der Studie werden im Hinblick auf die gefundene psycho-diagnostische Qualität und die Gerechtigkeit von Studienabschlussnoten diskutiert. Die vorgelegten Befunde werfen kritische Fragen bezüglich der implizierten Messmodelle in modularisierten Studiengängen auf. Sie unterstützen aber gleichzeitig – in Grenzen und in Übereinstimmung mit verwandten Analysen (vgl. Formazin, 2011) - die Berechtigung globaler Studienleistungsindizes, so wie sie in der bisherigen Literatur zur Kennzeichnung von aggregierten Studienleistungen und als Kriterien für Studienleistungsprognosen herangezogen wurden.

Formazin, M. (2011). Schlussfolgerndes Denken und fachspezifisches Vorwissen als Prädiktoren der Studienleistung im Fach Psychologie. Humboldt Universität Berlin, Mathematisch Naturwissenschaftliche Fakultät II, 2011.

Formazin, M., Schröders, U., Köller, O., Wilhelm, O., & Westmeyer, H. (2011). Studierendenauswahl im Fach Psychologie. Testentwicklung und Validitätsbefunde. Psychologische Rundschau, 62(4), 221–236.

Hell, B., Trapmann, S., & Schuler, H. (2008). Synopse der Hohenheimer Metaanalysen zur Prognostizierbarkeit des Studienerfolgs und Implikationen für die Auswahl- und Beratungspraxis. In H. Schuler & B. Hell (Eds.), Studierendenauswahl und Studienentscheidung (pp. 43–54). Göttingen: Hogrefe.

Hell, B., Trapmann, S., Weigand, S., & Schuler, H. (2007). Die Validität von Auswahlgesprächen im Rahmen der Hochschulzulassung - eine Metaanalyse. Psychologische Rundschau, 58, 93–102.

Kuncel, N. R., Wee, S., Serafin, L., & Hezlett, S. A. (2010). The Validity of the Graduate Record Examination for Master's and Doctoral Programs: A Meta-analytic Investigation. Educational and Psychological Measurement, 70(2), 340–352.

Soares, J. A. (2012). The Future of College Admissions: Discussion. Educational Psychologist, 47(1), 66–70.

Trost, G. (2005): Studierendenauswahl durch die Hochschulen: Welche Verfahren kommen prinzipiell in Betracht, welche nicht? Psychologische Rundschau, 56 (2), 138–140.





Chair(s): Eveline Wuttke (Goethe Universität Frankfurt)

Diskutant/in: Cornelia Gräsel (School of Education, Bergische Universität Wuppertal)

Nicht zuletzt die enttäuschenden Ergebnisse deutscher Schülerinnen und Schüler in internationalen Vergleichsstudien wie PISA, TIMSS, IGLU und DESI haben hierzulande Fragen nach der Qualität der Lehrer/-innen-Bildung im Allgemeinen und den Möglichkeiten des Kompetenzerwerbs in der Lehrer/-innen-Bildung im Besonderen aufgeworfen. Zahlreiche Studien aus der Lehr-Lern-Forschung belegen die entscheidende Bedeutung von Lehrerkompetenz für die Unterrichtsqualität, und zwar sowohl im Hinblick auf die Gestaltung des Unterrichts (Prozess) als auch auf die Schülerleistungen bzw. das Schülerverhalten (Output). Mit Blick auf die Zielvariable „leistungsfähige und erfolgreiche Schüler/-innen“ wäre somit eine qualitativ hochwertige Lehrer/-innen-Bildung – folgt man der plausiblen, wenn auch empirisch nicht hinreichend abgesicherten Wirkungskette von Lehrer/-innen-Kompetenz, Lehrer/-innen-Handeln und Schüler/-innen-Kompetenz – der erste Schritt zur Zielerreichung.

In der Diskussion um notwendige Facetten professioneller Kompetenz bzw. Lehrerexpertise liegen verschiedene Vorschläge zu ihrer Strukturierung vor. Die gängige Unterscheidung von Shulman in allgemeines pädagogisches Wissen, Fachwissen und fachdidaktisches Wissen wurde vielfach aufgegriffen und weiterentwickelt, z.B. um Wertvorstellungen bzw. Überzeugungen, die die Wirklichkeit vorstrukturieren und so zur Orientierungs- und Verhaltenssicherheit beitragen. Die professionelle Kompetenz von Lehrkräften ist also ein komplexes und facettenreiches Konstrukt.

Die Beiträge des Symposiums beleuchten verschiedene Facetten professioneller Kompetenz und Expertise (subjektive Überzeugungen, professionelle Wahrnehmung, Kompetenzen im Bereich Klassenmanagement), ihre Entwicklung und Förderung sowie Auswirkungen auf Schülerinnen und Schüler oder Klassen (Klassenzielstruktur). Sie tun dies aus verschiedenen Perspektiven, die den verschiedenen Disziplinenclustern geschuldet sind, denen die Referenten angehören (Didaktik der Naturwissenschaften und des Sachunterrichts, Psychologie, Erziehungswissenschaft). Die Beiträge beziehen sich auf (Teil) Projekte innerhalb des Forschungsprogramms „Entwicklung von Professionalität des Pädagogischen Personals in Bildungseinrichtungen“ (ProPäda) des Bundesministeriums für Bildung und Forschung (BMBF) und vereinen meist schon innerhalb eines Projektes einen Disziplinen übergreifenden Blick auf die untersuchten Phänomene. Sie spannen den Bogen von pädagogischem Personal in Grundschulen über Mathematiklehrkräften der Sekundarstufe I bis hin zu fächerübergreifenden Fragestellungen zu Lehramtsstudierenden im Allgemeinen.

Zu den in den unterschiedlichen Projekten angewendeten und in den Beiträgen berichteten Erhebungsmethoden zählen neben (einmaligen) Befragungen (Schülerwahrnehmung der Klassenzielstruktur sowie des Unterrichtshandelns von Lehrkräften), die längsschnittliche Erfassung von Lehrüberzeugungen (Fragebogen) und der professionellen Wahrnehmung von Lehrkräften (online administrierter Videotest) sowie ein längsschnittlich angelegtes Wartekontrollgruppendesign zu Effekten von Lehrerfortbildung und Kindertrainings.

Ziel des Symposiums ist es, diese unterschiedlichen Perspektiven – aber auch Gemeinsamkeiten – herauszustellen und zu diskutieren.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G161

## **Motivationsförderliches Unterrichtshandeln von Lehrkräften aus der Sicht von Schüler(inne)n: Welche Unterrichtsmerkmale kennzeichnen eine Lernzielstruktur in Schulklassen?**

Der Zielorientierungstheorie wird eine hohe Relevanz für die Erklärung von motivationalen Bedingungen des Lern- und Leistungsverhaltens von Schüler(inne)n beigemessen (vgl. Meece, Anderman, & Anderman, 2006). Eine Vielzahl an Studien zeigt, dass eine Lernzielorientierung eine Reihe von Vorteilen gegenüber einer Leistungszielorientierung aufweist. Ob ein(e) Schüler(in) eine Lernzielorientierung oder eine Leistungszielorientierung verfolgt, hängt dabei nicht nur von individuellen Antezedenzen ab, sondern wird auch von kontextuellen Bedingungen beeinflusst. Inwieweit der Klassenkontext das Verfolgen von Lern- bzw. Leistungszielen vorgibt, kann theoretisch durch die wahrgenommene Klassenzielstruktur abgebildet werden (Ames, 1992). Zu wenig empirisch erforscht ist bislang jedoch der Einfluss, den Lehrkräfte durch ihr Unterrichtshandeln auf die Klassenzielstruktur nehmen. Eine der seltenen Studien ist die Untersuchung von Turner et al. (2011), die das instruktionale Handeln von Mathematiklehrkräften im Zusammenhang mit der Motivation der Schüler(innen) mittels Unterrichtsbeobachtungen und qualitativen Interviews erforschten. Befunde aus Studien mit quantitativer Methodik fehlen dagegen weitgehend. Der vorliegende Beitrag geht daher der Frage nach, welches konkrete Unterrichtshandeln von Lehrkräften mit einer Lern- bzw. Leistungszielstruktur in Schulklassen im Zusammenhang steht. Zur Beantwortung der Fragestellung wurde ein Datensatz mit 1452 Gymnasiast(inn)en der 6. Jahrgangsstufe aus 70 Klassen analysiert, die zu ihrem Mathematikunterricht befragt wurden. Einbezogen wurden Schüler(innen)wahrnehmungen der Klassenzielstruktur sowie des Unterrichtshandelns der Lehrkräfte gemäß der sechs Dimensionen des TARGET-Modells nach Ames (1992) (Task, Authority, Recognition, Grouping, Evaluating, Time). Um Unterschiede im konzertierten Unterrichtshandeln zwischen Lehrkräften abzubilden, die unterschiedliche Klassenzielstrukturen realisieren, wurden auf Klassenebene Gruppen von Schulklassen mit unterschiedlichen Zielstrukturen gebildet, die anschließend varianzanalytisch hinsichtlich der Unterrichtsmerkmale untersucht wurden.

Die Ergebnisse zeigen mittlere bis große Unterschiede zwischen den gebildeten Gruppen von Lehrkräften hinsichtlich eines Großteils der untersuchten Unterrichtsmerkmale. Theoriekonform wurde beispielsweise das Unterrichtshandeln der Lehrkräfte in Schulklassen mit einer stark ausgeprägten Lernzielstruktur als strukturierter, klarer, interessanter und kognitiv aktivierender wahrgenommen als jenes der Lehrkräfte in Schulklassen mit einer schwach ausgeprägten Lernzielstruktur. Weiterhin beurteilten Schüler(innen) in Klassen mit einer stark ausgeprägten Leistungszielstruktur das instruktionale Handeln der Lehrkraft als weniger autonomieunterstützend und emotional wertschätzend als Schüler(innen) in Klassen mit einer schwach ausgeprägten Leistungszielstruktur. Zusammengefasst geben die Ergebnisse Hinweise darauf, welches spezifische Muster des Unterrichtshandelns von Lehrkräften mit einer motivationsförderlichen Klassenzielstruktur im Zusammenhang steht. Die pädagogischpsychologische Relevanz der Befunde wird abschließend diskutiert.

## **Tragen schulpraktische Studien zur Veränderung subjektiver Lehrtheorien bei?**

Subjektive Überzeugungen sind eine bedeutsame Facette professioneller Kompetenz, denn sie können die Wahrnehmung und Deutung von Unterrichtssituationen ebenso wie die an Schüler und Unterricht gerichteten Zielvorgaben und Erwartungen beeinflussen. Die wissenschaftlich fundierte und reflektierte Vorstellung von Lehren und Lernen gilt somit als notwendige Voraussetzung für professionelles Handeln (z.B. Dann, 1994; Baumert & Kunter, 2006). Die empirische Befundlage fällt je nach Überzeugungstyp unterschiedlich aus, insbesondere zu subjektiven Theorien über Lehren bzw. Unterrichten ist sie jedoch nur schwach und inkonsistent. So existieren beispielsweise kaum echte Längsschnittstudien zur Entwicklung von Lehr-Überzeugungen im Verlauf der Lehrerbildung und zu den Prädiktoren dieser Entwicklung. Allgemein ist davon auszugehen, dass sich Überzeugungen durch Erfahrungen in der eigenen Bildungsbiographie entwickeln (z.B. Woolfolk Hoy, Davis, & Pape, 2006). Vermutlich dürften also auch praktische Erfahrungen zur Entwicklung subjektiver Lehrüberzeugungen beitragen. In diesem Fall dürften schulpraktische Studien im Lehramtsstudium eine relevante Lerngelegenheit darstellen, anhand derer eigene subjektive Theorien zum Lehrer reflektiert und ggf. revidiert werden können. Plausibel ist aber auch die Annahme, dass Überzeugungen zum Lehren eine individuell spezifische „Brille“ darstellen, unter der Ziele, Aktivitäten und Erfahrungen in Schulpraktika betrachtet und interpretiert werden.

Im vorliegenden Beitrag soll deshalb den Fragen nachgegangen werden, ob 1.) Erfahrungen im Praktikum die Veränderung subjektiver Lehrüberzeugungen vorhersagen und ob 2.) die subjektiven Theorien umgekehrt auch Eingang in die Erreichung der Praktikumsziele finden. Den Analysen liegt eine fünf Erhebungswellen umfassende Längsschnittstudie mit 1100 Lehramtsstudierenden aus 10 Hochschulen zugrunde, die den Zeitraum vom ersten bis zum dritten Studiensemester abdeckt. Darin wurden subjektive Lehrtheorien über zwei Fragebogenskalen erfasst, die transmissive und konstruktive Lehrüberzeugungen messen (Bauer & Drechsel, 2010). Die Erfahrungen im Praktikum werden durch verschiedene Merkmale operationalisiert. Dazu gehören zwei Skalen auf Basis der Aufgaben bzw. Ziele von Praktika ("Erfahrungen im schulischen Kontext", "Reflexion Berufswahlentscheidung"), sowie Angaben der Studierenden zur Betreuungsqualität und den Arbeitsaufgaben.

Die Fragestellungen werden aktuell mittels Cross-Lagged-Struktur-gleichungs-modellen unter Berücksichtigung von Eingangs- und Persönlichkeitsmerkmalen der Studierenden geprüft. Dabei werden getrennte Modelle für transmissive und konstruktive Überzeugungen berechnet. Erste Analysen zeigen zunächst eine hohe zeitliche Stabilität der subjektiven Lehrüberzeugungen, das trifft sowohl auf transmissive, als auch auf konstruktive Überzeugungen zu. Bezüglich Fragestellung 1 weisen die Ergebnisse darauf hin, dass sich subjektive Theorien zumindest partiell durch die Erfahrungen im vorausgehenden Schulpraktikum vorhersagen lassen. Dies trifft jedoch nicht auf früh im Studium lokalisierte, orientierende Praktika zu. Bezüglich Fragestellung 2 sagen umgekehrt auch die subjektiven (transmissiven) Überzeugungen die Erfahrungen im Praktikum teilweise vorher. Bezüglich beider Fragestellungen sind die Effektstärken insgesamt klein.

Die Ergebnisse bestärken die Vermutung einer wechselseitigen Beziehung zwischen (schul-)praktischen Erfahrungen und Überzeugungen zum Lehren. Sie weisen darauf hin, dass Praktika eine relevante Lerngelegenheit zur Veränderung subjektiver Lehrtheorien sein können, auch wenn aufgrund der kurzzeitigen Intervention nur geringe Effekte zu erzielen sind. Weitere Analysen müssen darüber Auskunft geben, welche Merkmale von Praktika und ihrer Begleitung hier tatsächlich wirksam sind.

Bauer, J. & Drechsel, B. (2010, März). Subjektive Begriffe vom „Lehren“ im Lehramtsstudium: Erste Befunde mit einem neuen Instrument. Vortrag auf dem 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE), Mainz.

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469-520.

Dann, H.-D. (1994). Pädagogisches Verstehen. In Reusser, K. (Hrsg.). Verstehen: psychologischer Prozeß und didaktische Analyse. Bern, 163-182

Woolfolk Hoy, A., Davis, H., & Pape, S. (2006). Teachers' knowledge, beliefs, and thinking. In P. A. Alexander & P. H. Winne (Eds.), Handbook of educational psychology (2nd ed., pp. 715-737).

Mahwah, NJ: Lawrence Erlbaum.

Maria Todorova (Universität Münster), Cornelia Sunder (Universität Münster), Mirjam Steffensky (IPN),  
Kornelia Möller (Universität Münster)

## **Verbessert das Masterstudium die professionelle Wahrnehmung von naturwissenschaftlichem Grundschulunterricht?**

Die professionelle Wahrnehmung gilt als wichtige Komponente der Lehrerexpertise (Sherin, 2007). Sie umfasst das Erkennen und Interpretieren lernrelevanter Situationen im Unterricht und ist Voraussetzung für das adaptive Handeln der Lehrkraft (Bromme, 2008). Für den Mathematikunterricht konnte gezeigt werden, dass die professionelle Wahrnehmung der Lehrkraft positiv mit dem fachbezogenen Professionswissen sowie mit der Leistung der Schüler/innen zusammenhängt (Kersting et al., 2012). Empirische Studien zeigen auch, dass erfahrene Lehrkräfte Unterricht professioneller wahrnehmen als Studierende (Seidel & Prenzel, 2007; van Es & Sherin, 2002). Speziell im Kontext des naturwissenschaftlichen Sachunterrichts wurden signifikante Unterschiede in der professionellen Wahrnehmung zwischen Lehrkräften bzw. Masterstudierenden und Bachelorstudierenden aufgezeigt (Wolters, Meschede, Steffensky, & Möller, eing.). Diese Ergebnisse deuten an, dass sich die Fähigkeit, Unterricht professionell wahrzunehmen, im Laufe der Lehramtsausbildung sowie infolge von Unterrichtspraxis verbessert.

Im Rahmen des BMBF geförderten ViU-Projekts wird in einer Längsschnittstudie der Frage nachgegangen, wie sich die professionelle Wahrnehmung angehender Grundschullehrkräfte im Laufe ihrer Ausbildung entwickelt. In diesem Vortrag liegt der Fokus auf der professionellen Wahrnehmung von Sachunterrichtsstudierenden bezüglich Maßnahmen der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht (diese umfasst Maßnahmen der kognitiven Aktivierung und der inhaltlichen Strukturierung) sowie der Entwicklung dieser Fähigkeit im einjährigen Masterstudium. Der Einfluss vorausgehender formaler Lerngelegenheiten im Bachelorstudium wird ebenfalls berücksichtigt, indem die professionelle Wahrnehmung von Studierenden aus BA-Studiengängen mit einem naturwissenschaftlichem Schwerpunkt (NaWi, N = 55) und Studierenden aus BA-Studiengängen ohne einen naturwissenschaftlichen Schwerpunkt (kein NaWi, N = 29) gegenübergestellt wird. Die professionelle Wahrnehmung wurde mit Hilfe eines im VIU-Projekt entwickelten, online administrierten Videotest erfasst, der aus sechs Videoszenen und insgesamt 71 Ratingitems besteht.

Die Ergebnisse einer Varianzanalyse mit den Innersubjekt-Faktoren Zeit (Anfang Masterstudium vs. Ende Masterstudium) und Maßnahmen der Lernunterstützung (kognitive Aktivierung vs. inhaltliche Strukturierung) sowie mit dem Schwerpunkt im Bachelorstudium als Zwischensubjekt-Faktor (NaWi vs. kein NaWi) zeigen, dass sich die professionelle Wahrnehmung bezüglich Maßnahmen der inhaltlichen Strukturierung im Laufe des Masterstudiums verbessert und bezüglich Maßnahmen der kognitiven Aktivierung unverändert bleibt (Zeit\*Art der Lernunterstützung:  $F(1, 82) = 4.40, p < .05, \eta^2 = .05$ ). Diese Entwicklung ist für alle Studierenden unabhängig vom Studienschwerpunkt zu beobachten. Allerdings verfügt die Gruppe der Studierenden mit naturwissenschaftlichem Studienschwerpunkt insgesamt über eine bessere professionelle Wahrnehmung als die Gruppe ohne einen solchen Studienschwerpunkt ( $F(1, 82) = 12.97, p = .001, \eta^2 = .14$ ). Es zeigt sich weiterhin, dass Studierende Maßnahmen der inhaltlichen Strukturierung insgesamt besser wahrnehmen können als Maßnahmen der kognitiven Aktivierung ( $F(1, 82) = 61.88, p < .001, \eta^2 = .43$ ).

In diesem Vortrag werden die Anlage der Längsschnittstudie sowie die Ergebnisse vorgestellt und diskutiert.

- Bromme, R. & Haag, L. (2008). Forschung zur Lehrerpersönlichkeit. In W. Helsper & J. Böhme (Hrsg.), Handbuch der Schulforschung (S. 803-819). Wiesbaden: VS Verlag Sozialwissenschaften.
- Kersting, N.B., Givvin, K.B., Thompson, B.J., Santagata, R. & Stigler, J.W. (2012). Measuring Usable Knowledge: Teachers' Analyses of Mathematics Classroom Videos Predict Teaching Quality and Student Learning. *American Educational Research Journal*, 49, 568-589.
- Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen bei Lehrpersonen mit Hilfe von Videosequenzen. *Zeitschrift für Erziehungswissenschaft, Sonderheft*, 8, 201-216.
- Sherin, M.G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (S. 383-395). Hillsdale, NJ: Erlbaum.
- van Es, E.A. & Sherin, M.G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10, 571-596.
- Wolters, M., Meschede, N., Steffensky, M. & Möller, K. (eing.). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht - Theoretische Beschreibung und empirische Erfassung.

Charlotte Hanisch (Fachhochschule Düsseldorf), Ilka Eichelberger (Klinik für Psychiatrie und Psychotherapie des Kindes- und Jugendalters der Universität zu Köln), Stefanie Richard (Fachhochschule Düsseldorf), Silke Schiffhauer (Fachhochschule Düsseldorf), Lisa Greimel (Klinik für Psychiatrie und Psychotherapie des Kindes- und Jugendalters der Universität zu Köln), Manfred Döpfner (Klinik für Psychiatrie und Psychotherapie des Kindes- und Jugendalters der Universität zu Köln)

## **Interventionen bei expansivem Problemverhalten in der offenen Ganztagsgrundschule**

Expansives Problemverhalten umfasst aggressives, oppositionelles, unruhiges und aufmerksamkeitsbeeinträchtigt Verhalten. Es ist häufig über den Entwicklungsverlauf sehr stabil (Bongers et al., 2004) und birgt ein hohes Risiko für negative psychosoziale Folgen (Fontaine et al., 2008). Impulsive Verhaltenstendenzen, Regelverstöße und Aufmerksamkeitsprobleme beeinträchtigen die schulische Entwicklung (Dodge, Coie et al., 2006; Murray, Rabiner & Hardy, 2011) und stellen für Lehrer gerade an allgemeinen Grundschulen eine große Herausforderung dar (DuPaul & Evans, 2008). In internationalen Arbeiten gelten effektives Klassenmanagement, positive Lehrer-Kind Beziehungen, Strategien der Verhaltensmodifikation und des Kontingenzmanagements und kognitive Methoden zur Verbesserung der Selbststeuerung als evidenzbasierte Maßnahmen zur Prävention von expansivem Problemverhalten im Kontext Schule (Maggin et al., 2010; DuPaul et al., 2011; Miranda et al., 2009). In Deutschland sind diese Strategien für Kinder mit expansivem Problemverhalten bisher lediglich für den therapeutischen Kontext (Döpfner et al., 2004) und den Bereich Kindergarten (Hanisch et al., 2010; Hautmann et al., 2011; Hanisch et al., im Druck) evaluiert worden. Der Studie liegt die Fragestellung zugrunde, inwieweit sich die oben genannten Strategien in Qualifizierungsmaßnahmen an pädagogische Mitarbeiter und Lehrer offener Ganztagsgrundschule so vermitteln lassen, dass darüber Regelverhalten und Aufmerksamkeit bei Kindern mit expansivem Problemverhalten verbessert werden. In einer ersten Studie (Offene Ganztagsgrundschulen, OGS) wurden pädagogische Mitarbeiter aus der Nachmittagsbetreuung offener Ganztagsgrundschulen fortgebildet. Ziel war zum einen, Strategien der Verhaltensmodifikation an das gesamte Team zu vermitteln und so das Regelverhalten der Kinder generell zu verbessern. Zum anderen wurden einzelne Pädagogen darin geschult, Aufmerksamkeitstrainings für Kinder mit Aufmerksamkeitsproblemen (n= 116) anzubieten. Diese Aufmerksamkeitstrainings zielten darauf ab, über kognitive Methoden Arbeitsorganisation und Selbststeuerungsfähigkeiten beeinträchtigter Kinder zu steigern und so das Arbeitsverhalten v.a. in der Lernzeit zu verbessern. In einer zweiten Studie (Lehrercoaching) wird nun eine ähnliche Fortbildung mit Lehrerkollegien von Grundschulen durchgeführt. Im Anschluss an die Fortbildung werden einzelnen Lehrern sechs Einzelcoaching Termine angeboten, in denen die Inhalte aus Fortbildung und Kindertraining auf ein einzelnes Zielkind mit expansivem Problemverhalten übertragen werden sollen. Effekte der Fortbildungen, des Kindertrainings und der Lehrercoachings werden in längsschnittlich angelegten Eigenwartekontrollgruppendesigns anhand von Fragebogeneinschätzungen, computergestützten Aufmerksamkeits-tests, klinischen Interviews und Verhaltensbeobachtungen bzw. Unterrichtsbesuchen erhoben.

Ergebnisse: (1) OGS- Studie: Aufmerksamkeitsprobleme ( $p \leq .000$ ), nicht aber oppositionelles Problemverhalten wurden durch die Kombination aus Fortbildung und Kindertraining reduziert. Wenngleich sich auf Gruppenebene uneinheitliche Ergebnisse auf den verschiedenen Erfolgsmaßen und insgesamt geringere Effektstärken als erwartet zeigten ( $d = 0.18$  bis  $d = 0.42$ ), verbesserten sich 19% der Kinder mit klinisch relevanten Aufmerksamkeitsproblemen in den subklinischen bzw. unauffälligen Bereich. Im Einzelfall scheinen die Interventionen somit klinisch signifikante Effekte zu erzielen. Forschungsmethodische Probleme ergaben sich durch die hohe Variabilität in der Umsetzung des offenen Ganztages zwischen den Schulen und durch eine hohe Fluktuation der pädagogischen Mitarbeiter. Ausbildungshintergrund, Berufserfahrung der pädagogischen Mitarbeiter und Faktoren der Schulorganisation scheinen die Effekte von Fortbildung und Kindertraining zu moderieren. (2) Lehrercoaching- Studie: Aktuell läuft eine erste Untersuchungswelle an acht Grundschulen in Nordrheinwestfalen. Acht Lehrerkollegien wurden ganztägig fortgebildet, und 25 Lehrer bzw. Teams aus Lehrern und Sozialpädagogen nehmen an den fallbezogenen, individualisierten Coachings teil. Aus der Lehrercoaching-Studie können somit erste, vorläufige Ergebnisse berichtet werden.

Bongers, I., Koot, H.M., van der Ende, J. & Verhulst, F.C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, 75, 1523-1537.

Dodge, K.A., Coie, J.D. & Lynam, D. (2006). Aggression and antisocial behavior in youth. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 719-788). New York: Wiley.

DuPaul, G. J. & Evans, S. W. (2008). School-based interventions for adolescents with attention deficit/hyperactivity disorder. *Adolescent Medicine: State of the Art Reviews*, 19 (2), 300-312.

DuPaul, G. J., Helwig, J. R. & Slay, P. M. (2011). Classroom Interventions for Attention and Hyperactivity. In M. A. Bray & T. J. Kehle (Hrsg.), *The Oxford Handbook of School Psychology* (S. 428-441). New York: Oxford University Press.

Fontaine, N., Carbonneau, R., Barker, E.D., Vitaro, F., Hébert, M., Côté, S.M., Nagin, D.S., Zoccolillo, M. & Tremblay, R.E. (2008). Girls' hyperactivity and physical aggression during childhood and adjustment problems in early adulthood: A 15-year longitudinal study. *Archives of General Psychiatry*, 65 (3), 320-328.

Hanisch C, Freund-Braier I, Hautmann C, Jänen N, Plück J, Brix G, Eichelberger I & Döpfner M (2010) Detecting effects of the indicated prevention Programme for Externalizing Problem behaviour (PEP) on child symptoms, parenting, and parental quality of life in a randomized controlled trial. *Behav Cogn Psychother*; 38 (1), 95-112.

Hanisch, C., Hautmann, C., Plück, J., Eichelberger, I. & Döpfner, M. (im Druck). The Prevention program for Externalizing Problem behavior (PEP) improves child behavior by reducing negative parenting practices in conflict situations: analysis of mediating processes in a randomized controlled trial. *Journal of Child Psychology and Psychiatry*.

Hautmann C, Stein P, Eichelberger I, Hanisch C, Plück J, Walter D & Döpfner M (2011). The Severely Impaired Do Profit Most: Differential Effectiveness of a Parent Management Training for Children with Externalizing Behavior Problems in a Natural Setting. *Journal of Child Family Studies*, 20, 424-435.

Maggin, D.M., Chafouleas, S.M., Goddard, K.M. & Johnson, A.H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology*, 49 (5), 529-554.

Miranda, A., Presentación, M. J., Garcia, R. & Siegenthaler, R. (2009). Intervention with students with ADHD: Analysis of the effects of a multi-contextualized program on academic and socioemotional adjustment. In T.E. Scruggs & M.A. Mastropieri (Hrsg.), *Policy and Practice. Advances in Learning and Behavioral Disabilities* (S. 227-263). Bingley: Emerald.

Murray, D. W., Rabiner, D. L. & Hardy, K. K. (2011). Teacher management practices for first graders with attention problems. *Journal of Attention Disorders*, 15, 638-645.





## Struktur und Differenzierung des akademischen Selbstkonzepts in der Sekundarstufe I

Chair(s): Malte Jansen (IQB), Irene Schurtz (Otto-Friedrich-Universität Bamberg)

Diskutant/in: Jörn Sparfeldt (Universität des Saarlandes, Saarbrücken)

Ziel des Symposiums „Struktur und Differenzierung des akademischen Selbstkonzepts in der Sekundarstufe I“ ist es, die Differenzierung des akademischen Selbstkonzepts auf verschiedenen Spezifitätsebenen aufzuzeigen.

In einem ersten Beitrag wird auf Konstruktebene der Frage nachgegangen, inwieweit sich in der Domäne der Naturwissenschaften Selbstkonzept und Selbstwirksamkeitserwartung abgrenzen lassen und sich differentielle Zusammenhänge mit der Testleistung in PISA 2006 auf Individual- und Klassenebene, aber auch mit personen- und unterrichtsbezogenen Variablen zeigen. Dabei werden psychologische und fachdidaktische Perspektiven auf selbstbezogene Kognitionen in den naturwissenschaftlichen Fächern integriert (Jansen, Scherer & Schroeders).

Auf der Ebene domänenspezifischer Selbstkonzepte wird in einem zweiten Beitrag die Ausdifferenzierung von Selbstkonzepten in den Schulfächern Deutsch, Englisch und Mathematik von der 5. bis zur 7. Jahrgangsstufe betrachtet (Schurtz). Über den personenzentrierten Ansatz der latenten Klassenanalyse werden hierbei individuelle Entwicklungsunterschiede bezüglich des vermuteten Prozesses der Differenzierung domänenspezifischer Selbstkonzepte typisiert aufgezeigt.

Die Beiträge 3 und 4 beschäftigen sich detailliert mit der Binnendifferenzierung innerhalb eines domänenspezifischen Selbstkonzepts, nämlich des verbalen Selbstkonzepts. Im dritten Beitrag wird zunächst die Trennbarkeit des Leseselbstkonzepts vom Selbstkonzept im Schulfach Deutsch thematisiert. Dazu wird die divergente und konvergente Validität beider Konstrukte anhand der Beziehungen zu Noten und Leistungstests dargelegt (Arens).

Der vierte Beitrag zeigt schließlich, dass auch innerhalb eines tätigkeitsspezifischen Selbstkonzeptes wie dem Leseselbstkonzept weitere Binnendifferenzierungen möglich sind. Demnach lassen sich verschiedene, textsortenspezifische Leseselbstkonzepte (expositorisches und literarisches) differenzieren (Henschel, Roick, Brunner & Stanat).

Abschließend werden die Beiträge von Jörn Sparfeldt übergreifend diskutiert und in die Selbstkonzeptforschung eingeordnet. Die skizzierten Beiträge blicken somit aus verschiedenen Perspektiven auf die Struktur und Differenzierung des akademischen Selbstkonzepts in der Sekundarstufe I. Eine Perspektivenvielfalt zeigt sich zum einen hinsichtlich der gewählten methodischen Zugänge. So werden im Symposium zur Beantwortung der Fragestellungen quer- und längsschnittliche Studien sowie variablen- und personenzentrierte Ansätze genutzt. Zum zweiten können die betrachteten Spezifizierungsebenen aus theoretischer Sicht als unterschiedliche Perspektiven aufgefasst werden, deren gemeinsame Betrachtung die Hierarchie und Ausdifferenzierung des akademischen Selbstkonzepts verdeutlicht.

- Altman, U., Bohrn, I. C., Lubrich, O., Menninghaus, W. & Jacobs, A. M. (2012). Fact vs fiction—how paratextual information shapes our reading processes. *Social Cognitive and Affective Neuroscience*, 1-8.
- Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2013). A short German version of the Self Description Questionnaire I: Theoretical and empirical comparability. *International Journal of Research & Method in Education*, 36, 415-438.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1-40.
- Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology*, 102, 964-981.
- Chapman, J. W. & Tunmer, W. E. (1999). Reading Self-Concept Scale. In R. Burden (Ed.), *Children's self-perceptions* (pp. 29-34). Windsor: NFER-Nelson.
- Christmann, U. & Schreier, M. (2003). Kognitionspsychologie der Textverarbeitung und Konsequenzen für die Bedeutungskonstitution literarischer Texte. In F. Jannidis, G. Lauer, M. Martínez & S. Winko (Hrsg.), *Regeln der Bedeutung* (S. 246-285). Tübingen: Narr.
- Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, 19(4), 499-505.
- Frederking, V., Henschel, S., Meier, C., Roick, T., Stanat, P. & Dickhäuser, O. (2012). Beyond functional aspects of reading literacy: Theoretical structure and empirical validity of literary literacy. *L1-Educational Studies in Language and Literature*, 12, 35-58.
- Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371-395.
- Hattie, J. (1992). *Self-concept*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitskonzeptes vom Kindergarten bis zur 6. Klassenstufe. In F. E. Weinert (Hrsg.), *Entwicklung im Kindesalter* (S. 115-132). Weinheim: Psychologie Verlags Union.
- Henschel, S. & Roick, T. (2013). Zusammenhang zwischen Empathie und dem Verstehen literarischer Texte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45, 103-113.
- Henschel, S., Roick, T., Brunner, M. & Stanat, P. (2013). Leseselbstkonzept und Textart: Lassen sich literarisches und faktuales Leseselbstkonzept trennen? *Zeitschrift für Pädagogische Psychologie*, 27, 181-191.
- Möller, J. & Bonerad, E.-M. (2007). Fragebogen zur habituellen Lesemotivation. *Psychologie in Erziehung und Unterricht*, 54, 259-267.
- Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79, 1129-1167.
- Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2013). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*.
- Roick, T., Stanat, P., Dickhäuser, O., Frederking, V., Meier, C. & Steinhauer, L. (2010). Strukturelle und kriteriale Validität der literarästhetischen Urteilskompetenz. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*. 56. Beiheft der Zeitschrift für Pädagogik (S. 165-174). Weinheim: Beltz.
- Scherer, R. (2013). Further evidence on the structural relationship between academic self-concept and self-efficacy: On the effects of domain specificity. *Learning and Individual Differences*, 28, 9-19.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407-441.
- Streblov, L. (2004). *Bezugsrahmen und Selbstkonzeptgenese*. Münster: Waxmann.
- Swann, W. B., Chang-Schneider, C., & Larsen McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist*, 62(2), 84.

Dienstag,

04.03.2014,

10:15 - 12:00,

PEG-Gebäude

Raum 1.G165



## ***Selbstkonzept und Selbstwirksamkeit in den Naturwissenschaften: Empirische Trennbarkeit und differenzielle Zusammenhänge mit Motiven, Unterrichtsaktivitäten und Leistungen***

Theoretischer Hintergrund: Positive Fähigkeitsselbststeinschätzungen schulischer Kompetenzen gehen mit verschiedenen Bildungszielen wie höherer Anstrengung und Motivation, höheren Bildungsaspirationen sowie einer positiven Leistungsentwicklung einher. Das akademische Selbstkonzept sowie die schulische Selbstwirksamkeitserwartung stellen dabei die am häufigsten untersuchten Konstrukte dar, die solche Fähigkeitseinschätzungen widerspiegeln. Beide Konstrukte zeigen neben konzeptionellen Gemeinsamkeiten auch Unterschiede, die sich insbesondere auf ihre zeitliche Orientierung (vergangenheits- versus zukunftsbezogen), die Art der Selbstbeurteilung (evaluativ versus deskriptiv) sowie die zur Selbstbeurteilung herangezogenen Quellen (z.B. relative Wichtigkeit sozialer Vergleiche und aufgabenbezogener Erfolgs- und Misserfolgserfahrungen) beziehen (Bong & Skaalvik, 2003; Parker et al., 2013). Erste Befunde legen nahe, dass diese konzeptionellen Unterschiede auch mit einer empirischen Trennbarkeit einhergehen (Ferla, Valcke, & Cai, 2009; Parker et al., 2013; Scherer, 2013). Für den Bereich der Naturwissenschaften stehen eine solche Prüfung sowie detaillierte Analysen zu differenziellen Zusammenhängen beider Konstrukte mit motivationalen, unterrichts- und leistungsbezogenen Faktoren jedoch noch aus.

Fragestellung: In der vorliegenden Studie betrachten wir zunächst die messtheoretische Trennbarkeit der beiden Konstrukte. Anschließend sollen differenzielle Effekte theoretisch abgeleiteter Einflussfaktoren auf die Fähigkeitsselbststeinschätzungen untersucht werden. Beispielsweise wurde angenommen, dass sich die mittlere Klassenleistung aufgrund der Wichtigkeit sozialer Referenzrahmen stärker auf das Selbstkonzept, und forschungsorientierte Unterrichtsaktivitäten aufgrund der Wichtigkeit aufgabenbezogener Erfahrungen stärker auf die Selbstwirksamkeit auswirken sollten. Des Weiteren wird der Einfluss beider Konstrukte auf zwei Ziele naturwissenschaftlichen Unterrichts, die schulische Leistung und die zukunftsorientierte Motivation einen MINT-Beruf anzustreben, bestimmt.

Methodik: Es wurden Daten der deutschen Stichprobe von PISA-2006 genutzt (N=4891 15-jährige Schülerinnen und Schüler aller Schularten). Die Selbstkonzept und Selbstwirksamkeitserwartung wurden fach- bzw. aufgabenbezogen mit Hilfe von Likert-Skalen erfasst. Im ersten Auswertungsschritt wird die empirische Trennbarkeit der Konstrukte mit konfirmatorischen Faktorenanalysen untersucht. Im zweiten Schritt werden die Einflüsse von a) Geschlecht, b) der mittleren Klassenleistung, c) der Freude an den Naturwissenschaften, d) der generellen Werteinschätzung naturwissenschaftlicher Fächer und e) forschungsorientierter Unterrichtsaktivitäten auf die Fähigkeitsselbststeinschätzungen mit Regressionsmodellen ermittelt. Im dritten Schritt werden Effekte der Fähigkeitsselbststeinschätzungen auf die naturwissenschaftliche Leistung und die zukunftsorientierte Motivation modelliert.

Ergebnisse: Es konnte gezeigt werden, dass das naturwissenschaftliche Selbstkonzept und die Selbstwirksamkeitserwartung empirisch voneinander trennbare Konstrukte sind, die zu  $p = .57$  miteinander korrelieren. Demnach zeigt ein zweifaktorielles Messmodell eine gute Passung zu den Daten. Erwartungskonform zeigten sich differenzielle Zusammenhänge zu den verschiedenen Prädiktoren: Das Selbstkonzept wurde stärker durch Geschlecht, die durchschnittliche Klassenleistung und den wahrgenommenen Wert der Naturwissenschaften vorhergesagt. Die Selbstwirksamkeitserwartung zeigte hingegen höhere Zusammenhänge mit der Durchführung naturwissenschaftlicher Experimente im Unterricht. Zudem waren differenzielle Muster bei der Prädiktion von Bildungsergebnissen durch das Selbstkonzept und die Selbstwirksamkeitserwartung vorhanden: Bei gleichzeitiger Berücksichtigung beider korrelierter Konstrukte in Regressionsmodellen zeigte nur die Selbstwirksamkeitserwartung substantielle Effekte auf die Testleistung und nur das Selbstkonzept auf Karriereaspirationen in den naturwissenschaftlichen Fächern.

Diskussion: Neben ihren theoretischen Differenzen sind das Selbstkonzept und die Selbstwirksamkeitserwartung in den naturwissenschaftlichen Fächer eindeutig empirisch trennbar. Die Befunde legen nahe, dass sich das Selbstkonzept, im Vergleich zur Selbstwirksamkeitserwartung, eher aus sozialen Vergleichen speist und auch durch fachbezogene Einstellungen stärker beeinflusst wird. Die Selbstwirksamkeitserwartung wird hingegen stärker durch praktisches Experimentieren im Unterricht beeinflusst, was die theoretische Annahme stützt, dass aufgabenbezogene Erfolgserfahrungen wichtige Quellen der Selbstwirksamkeitserwartung darstellen. Auch scheint sie eine realistischere Einschätzung der eigenen Leistung in standardisierten Leistungstests zu sein. Für die Motivation sich im späteren Berufsleben mit einem MINT-Fach beschäftigen zu wollen, spielt sie hingegen eine geringere Rolle als das Selbstkonzept. Diese Befunde stehen im Einklang mit der theoretischen Konzeption der Selbstwirksamkeitserwartung als deskriptive Bewertung konkreter, aufgabenbezogener Fähigkeiten im Vergleich zum Selbstkonzept als globale, evaluative Bewertung der eigenen fachbezogenen Begabung. Eine Konsequenz der konzeptionellen und empirischen Distinktheit von Selbstkonzept und Selbstwirksamkeitserwartung ist, dass je nach Zielsetzung und Design einer pädagogisch-psychologischen Studie der relevante und sensitive Indikator der Fähigkeitsselbststeinschätzung herangezogen werden sollte.

## ***Differenzierung der schulfachbezogenen Selbstkonzepte in Deutsch, Mathematik und Englisch - Ein personenzentrierter Ansatz***

Theoretischer Hintergrund: Das schulfachbezogene Selbstkonzept stellt eine zentrale Variable bei der Prädiktion leistungsbezogener und motivationaler Faktoren dar. Studien, die sich mit der Entwicklung des Selbstkonzepts im schulischen Kontext befassten, konnten zeigen, dass die zunächst undifferenziert hohen, mittleren schulfachbezogenen Selbstkonzepte von Grundschülerinnen und Grundschülern im Schulverlauf absinken (z.B. Helmke, 1998; Streblov, 2004). Darüber hinaus ergaben Forschungsarbeiten zur Beziehung von mathematischen und verbalen Selbstkonzepten, dass sich die schulfachbezogenen Selbstkonzepte der verbalen und mathematischen Domäne während der Schulzeit zunehmend ausdifferenzieren (Brunner et al., 2010; Möller, Pohlmann, Köller, & Marsh, 2009).

Fragestellung: Bisherige Studien zur Entwicklung und Differenzierung schulfachbezogener Selbstkonzepte verfolgten mehrheitlich variablenzentrierte Ansätze und berichteten folglich mittlere Entwicklungstrends. Dem entgegen geht der vorliegende Beitrag der Frage nach, ob sich über die Methode der latenten Klassenanalyse (LCA) Subgruppen von Schülerinnen und Schülern auffinden lassen, die sich in dem vermuteten Prozess der Selbstkonzeptdifferenzierung unterscheiden. Die Analysen beziehen sich auf die ersten drei Jahre der Sekundarstufe. Dies ist ein sensibles Zeitfenster für die Selbstkonzeptentwicklung, da die mit dem Sekundarstufenübergang verbundene Neugruppierung in einen leistungssegregierten Schulkontext mit einer Neudefinition der schulfachbezogenen Selbstkonzepte einhergehen sollte (Hattie, 1992).

Methode: Die Datenerhebung erfolgte innerhalb der längsschnittlich angelegten BiKS-8-14-Studie. Die verwendete Stichprobe umfasste N=1903 Schülerinnen und Schülern (50.2% Mädchen). Das durchschnittliche Alter am Ende der 5. Jahrgangsstufe betrug 11.4 Jahre (SD=0.5). Der Anteil von Kindern mit Migrationshintergrund lag bei 12.0%. Das Gymnasium besuchten 54.5%, 17.7% gingen auf die Realschule, 14.5% auf die Hauptschule, 11.8% auf die Gesamtschule und 1.4% der Kinder besuchten eine Förderschule. Die Erfassung der schulfachbezogenen Selbstkonzepte in Mathematik, Englisch und Deutsch erfolgte jeweils zum Ende der 5., 6. und 7. Jahrgangsstufe über drei Items je Schulfach (z.B. „Mathematik kann ich gut.“). Zur Beantwortung der Items wurden 5 Kategorien dargeboten (1 = gar nicht; 2 = wenig; 3 = mittel; 4 = ziemlich; 5 = sehr). Die Reliabilitäten waren zufriedenstellend (Mathematik: .80-.91; Deutsch: .83-.88; Englisch: .88-.91). Die LCA erfolgte mit der Software Mplus (Muthén & Muthén). Als Variablen gingen die manifesten Selbstkonzeptskalen der drei Schulfächer zu den drei Messzeitpunkten (insgesamt 9 Variablen) ein. In aufsteigender Reihenfolge wurden die Lösungen für 1-10 latente Klassen berechnet. Die Bestimmung der endgültigen Anzahl latenter Klassen erfolgte anhand inhaltlich-theoretischer Gesichtspunkte. Unterstützend wurde der BIC, adjusted BIC, AIC, die Entropie sowie der parametric-bootstrapped Likelihood-Ratio-Test als Gütekriterien herangezogen.

Ergebnisse: Die Ergebnisse der LCA ergaben bezüglich der herangezogenen Gütekriterien keine eindeutige Bevorzugung einer Lösung. Jedoch verwies die Entropie ebenso wie die Beurteilung der latenten Klassen hinsichtlich inhaltlich-theoretischer Kriterien auf die Lösung mit 9 latenten Klassen. Die latenten Klassen dieser Lösung unterscheiden sich sowohl bezüglich des Verlaufsprofils über die drei Messzeitpunkte als auch bezüglich des Niveaus der Selbstkonzeptwerte. Es ergaben sich zwei latente Klassen, welche auf unterschiedlichen Niveaustufen keine Differenzierung in den Selbstkonzepten der drei Schulfächer aufzeigen (latente Klasse 5 und 6). Die latente Klasse 2 zeigt erst zum Ende der Jahrgangsstufe 7 Tendenzen der Selbstkonzeptdifferenzierung. Die übrigen latenten Klassen weisen verschiedene Profile fachlich differenzierter Selbstkonzepte zu Beginn der Sekundarstufe auf. Von Jahrgangsstufe 6 zu 7 verlaufen die fachspezifischen Selbstkonzepte dieser latenten Klassen jedoch parallel oder bewegen sich aufeinander zu.

Diskussion: Die Untersuchung verfolgte das Ziel, individuelle Entwicklungsunterschiede der Selbstkonzepte in den Schulfächern Mathematik, Deutsch und Englisch zwischen Jahrgangsstufe 5 und 7 typisiert nachzuzeichnen. Die Analysen ergaben 9 latente Klassen, deren Profile sich vorrangig hinsichtlich der Niveaus und Positionsfolgen der Fächer unterschieden. Obwohl die Mehrheit der latenten Klassen zu Beginn der Sekundarstufe ein differenziertes Selbstkonzeptprofil aufwies, zeigte sich zum Ende der 7. Jahrgangsstufe ein leicht gegenläufiger Trend. Eine zunehmende Differenzierung konnte somit für diese Fächer und den betrachteten Zeitraum nicht beobachtet werden.

## ***Binnendifferenzierung des verbalen Selbstkonzepts: Unterschiede und Gemeinsamkeiten der Selbstkonzepte im Lesen und in Deutsch***

Theoretischer Hintergrund: Zur Erfassung des verbalen Selbstkonzepts von Schülerinnen und Schülern der mittleren Klassenstufen wurden in vorherigen Studien sowohl Skalen zum Selbstkonzept in der Muttersprache als auch Skalen zum Selbstkonzept im Lesen eingesetzt. Dies verleitet zu der Annahme, dass diese beiden Selbstkonzepte zumindest bei Schülerinnen und Schülern der mittleren Klassenstufen die gleichen Konstrukte erfassen. Diese implizite Annahme wurde allerdings empirisch noch nicht untersucht.

Fragestellung: Auf der Grundlage dreier empirischer Studien mit Schülerinnen und Schülern der fünften Klassenstufe wurde untersucht, ob Skalen zum Selbstkonzept in Deutsch und Skalen zum Leseselbstkonzept das gleiche Konstrukt messen. Dabei werden sowohl die interne Struktur beider Konstrukte als auch, im Sinne konvergenter und divergenter Validität, Zusammenhänge mit anderen Selbstkonzeptfaktoren und mit Leistungsmaßen betrachtet.

Methoden: In der ersten (N = 163) und zweiten (N = 222) Studie wird das Selbstkonzept im Lesen erfasst. In der dritten Studie (N = 436) wird das Selbstkonzept in Deutsch erfasst. In allen Studien werden zusätzlich die Selbstkonzepte für Mathematik und Schule im Allgemeinen als weitere akademische Selbstkonzeptfacetten erfasst. Zur Erfassung der akademischen Selbstkonzepte wird auf Skalen der deutschen Version des Self-Description Questionnaire I zurückgegriffen (Arens, Yeung, Craven, & Hasselhorn, 2013). Als Indikator der verbalen Leistung wird in Studie 1 ein Leseverständnistest durchgeführt. In Studie 2 und 3 werden die Noten der Schülerinnen und Schülern im Fach Deutsch erfasst. Die Auswertung erfolgt mit Hilfe von konfirmatorischen Faktorenanalysen und der Betrachtung der latenten Zusammenhänge zwischen den verschiedenen akademischen Selbstkonzeptfacetten und den Leistungsindikatoren.

Ergebnisse: In Studie 1 und 2 zeigt sich, dass das Selbstkonzept im Lesen eine separate, von den Selbstkonzepten in Mathematik und Schule getrennte, Selbstkonzeptfacette ausbildet. In Studie 3 erweist sich auch das Selbstkonzept in Deutsch als eigenständige Selbstkonzeptfacette. Ebenso belegen alle drei Studien, dass sich die Selbstkonzepte im Lesen und in Deutsch in eine Kompetenzdimension und eine Affektdimension unterteilen lassen. Diese Befunde sprechen dafür, dass die Selbstkonzepte im Lesen und in der Muttersprache eine ähnliche interne Struktur aufweisen, was für ihre konzeptuelle Ähnlichkeit spricht. Eine konzeptuelle Ähnlichkeit der Selbstkonzepte im Lesen und in der Muttersprache sollte sich aber auch in vergleichbaren Zusammenhängen zu Leistungsindikatoren als externe Validierungskriterien zeigen. In Studie 1 zeigt das Selbstkonzept im Lesen einen höheren Zusammenhang zu den Werten im Leseverständnistest als die Selbstkonzepte für Mathematik und Schule. Allerdings zeigen das Selbstkonzept im Lesen und das Selbstkonzept für Schule in Studie 2 ähnlich hohe Zusammenhänge zur Schulnote in Deutsch, während das Selbstkonzept in Deutsch in Studie 3 die relativ höchsten Zusammenhänge zur Schulnote in Deutsch aufweist.

Diskussion: Somit deuten die drei Studien auf Gemeinsamkeiten in der internen Struktur der Selbstkonzepte im Lesen und in Deutsch, aber gleichzeitig auch auf Unterschiede in ihren Zusammenhängen zu Leistungsindikatoren. Entsprechend scheinen die Selbstkonzepte im Lesen und in Deutsch nicht als austauschbare Indikatoren für das verbale Selbstkonzept von Schülerinnen und Schülern verwendet werden zu können. Vielmehr unterstützen diese Befunde die Annahmen des „specificity matching principle“ (Swann, Chang-Schneider, & Larsen McClarty, 2007), wonach Selbstkonzepte und Validierungskriterien dann die höchsten Zusammenhänge aufweisen, wenn sich beide auf den gleichen inhaltlichen Bereich beziehen. Das Selbstkonzept im Lesen scheint die Selbstwahrnehmungen im spezifischen Fähigkeitsbereich des Lesens abzubilden und daher geeignet zu sein zur Vorhersage und Erklärung von lesespezifischen Leistungen. Das Selbstkonzept im Lesen scheint hingegen zu bereichsspezifisch zur Vorhersage und Erklärung der Deutschnote, die verschiedene verbale Fähigkeiten und nicht nur Lesekompetenzen abbildet. Zur Vorhersage und Erklärung der Deutschnote scheint sich somit das Selbstkonzept in Deutsch besser zu eignen, das entsprechend als breiterer Indikator für das verbale Selbstkonzept von Schülerinnen und Schülern betrachtet werden kann.

## **Entwicklung und Validierung textartspezifischer Leseselbstkonzepte**

Theoretischer Hintergrund: Literarische und expositorische Texte unterscheiden sich in ihrer funktionalen Zielsetzung und sprachstrukturellen Gestaltung. Die Verarbeitung beider Textarten könnte deshalb mit teilweise unterschiedlichen kognitiven Prozessen einhergehen (z.B. Altman, Bohrn, Lubrich, Menninghaus, & Jacobs, 2012; Christmann & Schreier, 2003; Graesser, Singer, & Trabasso, 1994). Entsprechend zeigen empirische Studien, dass sich das Verstehen literarischer und expositorischer Texte als voneinander unterscheidbare Aspekte der Lesekompetenz betrachten lassen (Frederking et al., 2012; Roick et al., 2010). In diesem Zusammenhang ist auch anzunehmen, dass individuelle Lernermerkmale textartspezifische Verarbeitungsprozesse differenziell begünstigen (z.B. Henschel & Roick, 2013). Zu den motivationalen Merkmalen, die als Prädiktoren für das allgemeine Textverstehen identifiziert wurden, zählt unter anderem das Leseselbstkonzept (Chapman & Tunmer, 1999). Mit der derzeitigen Konzeptualisierung des Leseselbstkonzeptes ist jedoch die Einschränkung verbunden, dass es die Fähigkeitsselbsteinschätzung im Lesen unabhängig von der gelesenen Textart erfasst (vgl. Möller & Bonerad, 2007). Dass eine textartspezifische Ausdifferenzierung sinnvoll sein könnte ergibt sich daraus, dass Jugendliche am Ende der Sekundarstufe I differenzielle textartspezifische schulische und außerschulische Leseerfahrungen und Leistungsrückmeldungen im Umgang mit literarischen und expositorischen Texten erfahren. Daraus könnten sich textartspezifische motivationale Überzeugungen (z.B. Interesse, lesebezogene Selbstkonzepte) entwickeln, die sich voneinander abgrenzen lassen und differenziell mit dem literarischen und expositorischen Textverstehen zusammenhängen. Da Leseerfahrungen mit spezifischen Textsorten gemacht werden, ist anzunehmen, dass Leseselbstkonzepte durch konkrete textsortenspezifische Leseselbstkonzepte (z.B. für Gedichte, Zeitungsartikel) auf einer darunterliegenden Hierarchieebene bestimmt werden.

Fragestellung: Die vorliegende Studie verfolgt das Ziel, Selbstkonzeptskalen für das Verstehen von literarischen und expositorischen Texten zu entwickeln und die strukturelle sowie kriteriale Validität der Skalen empirisch zu prüfen.

Methode: An der Untersuchung nahmen 1004 bayerische Jugendliche aus 46 Klassen der neunten Jahrgangsstufe (46% Mädchen, Alter:  $M = 15.69$  Jahre,  $SD = 0.76$ ) von Realschulen und Gymnasien teil. Die Jugendlichen bearbeiteten einen Fragebogen, mit dem u.a. die Skalen zum Leseselbstkonzept (literarisch: 9 Items;  $\alpha = .80$ , z.B. „Es fällt mir leicht, ein Gedicht zu interpretieren.“; expositorisch:  $\alpha = .81$ , z.B. „Die Hauptaussage aus einem Zeitungskommentar kann ich schnell erfassen.“) und weitere motivationale (z.B. intrinsische Lesemotivation, Selbstkonzept Deutsch) sowie leistungsbezogene Merkmale (Schulnoten, literarisches und expositorisches Textverstehen) erfasst wurden. Die Überprüfung der strukturellen und kriterialen Validität der Leseselbstkonzeptskalen erfolgte mittels konfirmatorischer Faktorenanalysen bzw. im Rahmen von Strukturgleichungsanalysen.

Ergebnisse: Aufgrund der bisherigen Erkenntnisse zur hierarchischen Struktur von Selbstkonzepten (Shavelson, Hubner, & Stanton, 1976) ist die Annahme einer hierarchischen Leseselbstkonzeptstruktur nahe liegend, die sowohl die Textsorten- als auch die Textartspezifität berücksichtigt. Die strukturelle Validität des Leseselbstkonzeptes wurde deshalb konfirmatorisch untersucht, indem zwei hierarchische Modelle gegeneinander getestet wurden. Im ersten Modell wurde ein übergeordneter Generalfaktor für die Leseselbstkonzeptstruktur angenommen. Auf der darunterliegenden Hierarchieebene wurden die textsortenspezifischen Leseselbstkonzepte unterschieden. Das zweite Modell berücksichtigt literarisches und expositorisches Leseselbstkonzept als Faktoren zweiter Ordnung. Der Vergleich beider Modelle macht deutlich, dass das eindimensionale Generalfaktormodell eine schlechtere Anpassung an die empirischen Daten aufweist als das hierarchische zweifaktorielle Modell ( $\Delta\chi^2 = 98.26$ ,  $df = 2$ ,  $p < .01$ ). Die Korrelation zwischen den beiden Faktoren zweiter Ordnung fällt zwar substantiell, aber nicht perfekt aus ( $r = .60$ ).

Die kriteriale Validität der zweidimensionalen Leseselbstkonzeptstruktur wurde durch differenzielle Zusammenhänge mit motivationalen und leistungsbezogenen Merkmalen überprüft. So korreliert das literarische Leseselbstkonzept im Vergleich zum expositorischen enger mit dem Selbstkonzept Deutsch, der intrinsischen Lesemotivation und der Deutschnote. Darüber hinaus erklärt das literarische Leseselbstkonzept in einem Strukturgleichungsmodell unter Kontrolle des expositorischen Leseselbstkonzeptes ausschließlich Varianz im literarischen Textverstehen. Das expositorische Leseselbstkonzept ist unter Kontrolle des literarischen Leseselbstkonzeptes dagegen ausschließlich mit dem Verstehen expositorischer Texte assoziiert und trägt nicht zur Erklärung des literarischen Textverstehens bei (Henschel, Roick, Brunner, & Stanat, 2013).

Diskussion: Die Ergebnisse der Studie erweitern die Erkenntnisse über sprachliche Selbstkonzepte und stellen eine solide Grundlage für differenzierte Analysen textartspezifischer Verstehensleistungen dar.





Chair(s): Didier Vaccaro (Staatsinstitut für Schulqualität und Bildungsforschung)

Diskutant/in: Hans Anand Pant (Institut zur Qualitätsentwicklung im Bildungswesen (IQB))

Im Zusammenhang mit dem Thema Bildungsmonitoring wird immer wieder die Frage aufgeworfen, inwieweit Schulen mit unterschiedlicher Zusammensetzung von Schülerinnen und Schülern bei der Bewertung von Schülerleistungen, aber auch in Bezug auf andere Themen, wie z. B. bei der Vergabe von Mitteln, unterschiedlich behandelt werden müssen. Dabei wird insbesondere auf die Bedeutung der Verfügbarkeit bildungsrelevanter Ressourcen im Elternhaus hingewiesen. Unter dem Schlagwort „fairer Vergleich“ wurden in der Vergangenheit verschiedene Ansätze vorgeschlagen, die ermöglichen sollen, den sozioökonomischen Hintergrund einer Schülerschaft bei der Einordnung von Testergebnissen zu berücksichtigen bzw. den soziokulturellen Rahmenbedingungen einer Schule Rechnung zu tragen. In den meisten Fällen wird dabei versucht, über einen sogenannten „Sozialindex“ die Schülerschaft einer Schule so zu charakterisieren, dass Vergleichsgruppen mit ähnlicher Schülerschaft gebildet werden können. Eine besondere Herausforderung stellt die Erhebung der bildungsrelevanten Informationen über die Schülerfamilie dar, zumal viele Merkmale, wie z. B. der Bildungshintergrund der Eltern, i.d.R. in keiner Statistik zentral erfasst sind. Darüber hinaus werden in methodischer Hinsicht hohe Anforderungen an solche Indices gestellt. In dem Symposium sollen neben grundsätzlicher methodischer Kritik an eben diesen Verfahren (Mythos fairer Vergleiche) verschiedene Lösungsansätze vorgestellt werden, die auf sehr unterschiedlicher Informationsgrundlage im Wesentlichen dieselbe Zielsetzung verfolgen, nämlich Schulen „fair“ zu behandeln. Dabei wird es zum einen um die Frage gehen, wie zeitstabil ein Index ist, der u. a. über die Befragung von Schülern und Eltern ermittelt wurde bzw. welche norm- oder kriteriumsorientierten Interpretationsmöglichkeiten bei fehlender Stabilität gegeben sind (Zur Stabilität von Sozialindikatoren – Hamburger Ergebnisse und Perspektiven). Bei einem weiteren Beitrag wird es ebenfalls um zeitliche Stabilität gehen, diesmal jedoch in Bezug auf den Zusammenhang von Kontextmerkmalen und Leistung, und verschiedene Ansätze werden im Hinblick auf ihre Tauglichkeit für den Routinebetrieb und ihre Kommunizierbarkeit verglichen (Vergleich verschiedener Adjustierungsansätze). Schließlich soll die Frage beleuchtet werden, wie trotz fehlender Befragung einzelner Personen (Schüler/innen bzw. Eltern) dennoch ein schulscharfer Sozialindex erstellt werden kann, wobei die Nützlichkeit zentraler Merkmale in Abhängigkeit von ihrer Definition thematisiert wird. Gemeinsam spiegeln die Beiträge die Vielfalt der Ansätze wider, die aktuell in Deutschland zu Anwendung kommen, und stecken damit den argumentativen Raum ab, in dem sich derzeit die Diskussion um die „faire“ Behandlung von Schulen dreht.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G131



## Mythos fairer Vergleiche

Für Vergleiche zwischen den einzelnen Schulen bzw. zwischen den einzelnen Schulen und dem Gesamtmittelwert können die Ergebnisse der flächendeckenden Tests bzw. der zentralen Abschlussprüfungen potentiell genutzt werden. Da diese Ergebnisse nicht nur von den schulischen, sondern auch von außerschulischen Einflussfaktoren abhängig sind, ist es sicher einleuchtend, dass sich diese Ergebnisse nicht für eine faire Beurteilung der Effektivität von Schulen eignen. Nichtsdestotrotz haben diese Vergleiche einen evaluatorischen Charakter: Sie zeigen, wo die Handlungsfelder liegen. Es gibt bisweilen den Wunsch, die Ergebnisse zentraler Tests bzw. Prüfungen auch für die Beurteilung der Effektivität von Schulen zu verwenden. Um dies fair zu gestalten, wird versucht, die Effekte der außerschulischen Variablen aus diesen Ergebnissen heraus zu partialisieren. Die Schulen können demnach für ihre dadurch „bereinigten“ (adjusted) Ergebnisse verantwortlich gemacht werden. Zu diesem Zweck werden i. d. R. regressionsanalytische Modelle verwendet, in denen die Testergebnisse nur durch die außerschulischen Einflussfaktoren erklärt werden. Als erklärende Faktoren kommen dafür sowohl einzelne Variablen in Frage (z. B. Geschlecht) als auch Indizes (wie Sozialindex), die sowohl auf Individualebene als auch auf Schulebene zusammengesetzt werden können, und für die entsprechende Daten zur Verfügung stehen. Anhand des Varianzanteils, der durch diese Modelle erklärt wird, kreiert man Erwartungswerte für die einzelnen Schulen, und anhand des Varianzanteils, der unerklärt bleibt (Residuen), werden Abweichungen zu diesen Erwartungswerten für die einzelnen Schulen berechnet. Diese Abweichungen werden dann als Differenzen zu den vergleichbaren Schulen interpretiert. Diese Vergleiche werden gern als „fair“ bezeichnet. In dem Vortrag wird der Frage nachgegangen, ob diese Vergleiche wirklich fair sind. Es wird diskutiert, ob die Residuen nur diejenigen Varianzanteile beinhalten, die durch die Berücksichtigung der schulischen Einflussfaktoren hätten erklärt werden können, oder ob es weitere Ursachen dafür gibt. Als mögliche Ursachen werden u. a. folgende Faktoren diskutiert:

- a) ausgelassene erklärende Variablen auf Kontextebene, z. B. Effekte der früheren Bildungsinstanzen sowie Effekte der Schulaufsichtsbehörden und des Schulträgers
- b) ausgelassene erklärende Variablen auf Individualebene, z. B. Intelligenz, Legasthenie und Dyskalkulie,
- c) ausgelassene Interaktionseffekte zwischen individuellen und Schulvariablen,
- d) Messfehler, sowohl bei den Testergebnissen als auch bei den im Modell berücksichtigten Individualvariablen bzw. Sozialindizes,
- e) unzutreffende Annahmen über das Skalenniveau der abhängigen und der erklärenden Variablen,
- f) unzutreffende Annahmen über die Linearität der Zusammenhänge usw.

Es wird der Frage nachgegangen, inwieweit die Übertragung der Stichprobenergebnisse bzw. der Ergebnisse der Totalerhebungen auf die einzelnen Untersuchungseinheiten – wie die „Fairer-Vergleich-Modelle“ es tun – den probabilistischen Charakter der einschlägigen Modelle zugunsten einer deterministischen Schlussfolgerung zuwider läuft. Dürfen überhaupt die regressionsanalytischen Modelle, die für Systemanalysen entwickelt worden sind, für die Einzelfallanalysen verwendet werden oder stellt dies eine unzulässige Zweckentfremdung dar? Und es wird darüber diskutiert, ob unterschiedliche Varianten der Querschnittsanalysen (einschließlich Mehrebenenanalysen) und Längsschnittsanalysen (einschließlich Value-Added-Modelle) mit ähnlichen methodischen Problemen behaftet sind.

Zum Schluss wird die Frage gestellt, ob man sich mit erwartungsgemäß schlechten Ergebnissen einer Schule, die entsprechend schlechte Ausgangsbedingungen aufweist, zufrieden geben darf und ob diese methodisch bedingte Vorstellung der Fairness (ungeachtet aller ihrer methodischen Unzulänglichkeiten) auch inhaltlich und vor allem gegenüber den betroffenen Schülern fair ist.

Hanushek, Eric A. and Jackson, John E.: Statistical methods for social scientists. NY: Academic, 1977.

Goldstein, Harvey: Multilevel Statistical Models, 3rd edn. London: Hodder Arnold, 2003.

Reardon, Sean F. and Raudenbush, Stephen W.: Assumptions of Value-Added Models for Estimating School Effects. In: Journal of Education, Finance and Policy, 2009.

## Zur Stabilität von Sozialindikatoren – Hamburger Ergebnisse und Perspektiven

In Hamburg gibt es seit 1996 einen Sozialindex für Grundschulen und weiter-führende Schulen mit Sekundarstufe I. Dieser beschreibt, seit 2006 (Bos, Pietsch, Gröhlich, & Janke, 2006) basierend u. a. auf der Kapitaltheorie von Bourdieu (1982, 1983), die sozialen Rahmenbedingungen der pädagogischen Prozesse an Schulen. Die auf dem Index basierende Zuordnung zu sechs abgestuften Belastungsgruppen hat Auswirkungen auf verschiedenen Ebenen: Auf der einen Seite beeinflusst der Sozialindex die Ressourcenallokationen auf Systemebene (z. B. kleinere Klassenfrequenzen oder höhere Sprachfördermaßnahmen für Schulen mit niedrigeren Indizes). Auf der anderen Seite wird der Sozialindex auch in weiteren Zusammenhängen genutzt: bei der Bildung repräsentativer Stichproben im Rahmen von wissenschaftlichen Untersuchungen und Evaluationen (z. B. bei der Auswahl einer repräsentativen Kernstichprobe von Schulen pro Schuljahr für die Schulinspektion), bei der Berechnung von Vergleichswerten („fairer Vergleich“) für die schulbezogenen Ergebnismeldungen im Rahmen von standardisierten Schulleistungstests oder bei der Bildung von Vergleichsgruppen im Kontext der Bildungsberichterstattung. Hamburg reagiert damit bildungspolitisch auf den über die Jahre leicht entkoppelten, aber auch aktuell noch beschriebenen Zusammenhang zwischen der sozialen Herkunft und dem Kompetenzerwerb in Deutschland, wie auch in den PISA Ergebnissen gezeigt werden konnte (Klieme et al., 2010). Dem Sozialindex liegt ein theoretisches Modell zugrunde, das verschiedene soziokulturelle Dimensionen voneinander unterscheidet: soziales Kapital, ökonomisches Kapital, kulturelles Kapital und Migrationsmerkmale. Die drei erstgenannten Facetten orientieren sich an Bourdieus Ansatz der Kapitalarten (Bourdieu, 1982, 1983), der Unterschiede im bestehenden sozialen Status und deren Reproduktion in den Familien der Schülerinnen und Schüler, aber auch im Bildungssystem (Bourdieu & Passeron, 1971) aufzeigt. Desweiteren werden bei der Konstruktion von Sozialindizes Migrationsmerkmale der Schülerinnen und Schüler herangezogen, da in Bezug auf den Migrationshintergrund noch immer bedeutsame soziale Disparitäten nachgewiesen werden können, die unabhängig von den weiteren Dimensionen sind (Klieme et al., 2010). Um über diese verschiedenen theoretischen Facetten hinaus unterschiedliche Datenquellen zur höheren Validität des Sozialindex zu nutzen, werden für die Konstruktion des Index neben der Befragung von Schülerinnen und Schülern sowie deren Eltern zusätzlich Sozialraumdaten (z. B. die Arbeitslosenquote) integriert. Diese Sozialraumdaten liegen auf Ebene der Statistischen Gebiete vor, in denen die Schülerinnen und Schüler in Hamburg leben.

Im Laufe der Verwendung des Sozialindex wurden verschiedene methodische Verfahren genutzt: Clusteranalysen, Skalierungen auf Basis der Item Response Theory sowie konfirmatorische Faktorenanalysen. Bei der wiederholten Erhebung und Entwicklung von Sozialindikatoren in Hamburg ergeben sich die folgenden Fragestellungen: 1. Sind die Belastungswerte und Stufenbildungen durch den Sozialindex über jeweils zwei Zeitpunkte hinweg stabil (Grundschulen: 2007 und 2011, weiterführende Schulen 2005 und 2011)? 2. Falls nicht, gibt es die Möglichkeit, neben der normorientierten Interpretation der schulspezifischen Belastung auch eine kriteriumsorientierte Interpretation, die unabhängig von der Verteilung der Belastungswerte ist, vorzunehmen? Welches Verfahren der Stufeneinteilung durch eine solche Standardsetzung ist angemessen? Um diese Fragen zu überprüfen wurden u. a. Korrelationen und latente Klassenanalysen berechnet. Grundsätzlich scheint das Konstrukt sozialer Belastung relativ zeitstabil zu sein, wie die Stabilitätsanalysen zeigen. Nichtsdestotrotz ergibt sich das Problem nicht-stabiler häufigkeitsbasierter Einteilungen der Belastungsgruppen aufgrund messfehlerbedingter Unschärfen in den Schwellenbereichen der Indexstufen. Über eine Standardsetzung im Rahmen eines Experten-Workshops (Variante der Angoff-Methode) haben sich vielversprechende Hinweise, jedoch keine verwertbaren Ergebnisse ergeben (Schulte, Hartig, & Pietsch, 2013). Um die Komplexität für die Teilnehmer/-innen gering zu halten, bieten sich eher Methoden der Standardsetzung an, die mit Prototypen arbeiten, welche dann wiederum im Sinne der bookmark-Methode (Cizek & Bunch, 2007) nach Belastungsgruppen eingeschätzt werden können. Voraussetzung für diese Methoden können gruppenentdeckende Verfahren, wie die latente Klassenanalyse sein. Erste Analysen, weitergehende Überlegungen sowie Perspektiven werden aufgezeigt.

Bourdieu, P. (1982). Die feinen Unterschiede – Kritik der gesellschaftlichen Urteilskraft. Frankfurt am Main: Suhrkamp.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In: R. Kreckel (Hrsg.): Soziale Ungleichheiten, Sonderband 2 der Zeitschrift „Soziale Welt“, S. 183-198. Göttingen: Schwartz & Co.

Bourdieu, P. & Passeron, J.-C. (1971). Die Illusion der Chancengleichheit: Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs. Klett. Stuttgart.

Bos, W., Pietsch, M., Gröhlich, C. & Janke, N. (2006). Ein Belastungsindex für Schulen als Grundlage der Ressourcenzuweisung am Beispiel von KESS 4. In W. Bos et al. (Hrsg.), Jahrbuch der Schulentwicklung, Band 14, S. 149-160. Weinheim: Juventa.

Cizek, G.J. & Bunch, M.B. (2007). Standard-setting. A guide to establishing and evaluating performance standards on tests. California: Sage Publications Inc.

Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W. & Stanat, P. (Hrsg.) (2010). PISA 2009. Bilanz nach einem Jahrzehnt. Münster: Waxmann.

Schulte, K., Hartig, J. & Pietsch, M. (2013). Der Sozialindex für Hamburger Schulen – Versuch eines Standard Settings. Vortrag auf der 1. Tagung der Gesellschaft für Empirische Bildungsforschung, 11.-13. März in Kiel.



## Vergleich verschiedener Adjustierungsansätze

Die Bedeutung von Schulergebnissen aus Leistungstests hat in den letzten Jahren stetig zugenommen und dies nicht nur bei denjenigen, die sich professionell mit dem Bildungssystem beschäftigen, sondern weit darüber hinaus, wenn z.B. Eltern Leistungsdaten bei der Schulwahl für ihre Kinder heranziehen möchten. So veröffentlicht das Land Berlin schulspezifisch seit 2013 Abiturergebnisse, ab 2014 Ergebnisse aus dem Mittleren Schulabschluss (MSA). Der Aussagekraft von Schulergebnissen - sei es als veröffentlichte, sei es als interne Rückmeldungen an die Schulen - kommt damit eine zentrale Funktion für die Effektivität des Systems schulischer Leistungsmessung zu. Aus Rückmeldungen muss hervorgehen, wo eine Schule im Gesamt des Schulsystems positioniert ist und wo sich Stärken und Schwächen in den Lernständen der Schülerschaft zeigen. Ein wesentliches Problem bei der Einordnung von Schulergebnissen ist, dass Leistungswerte zu einem erheblichen Teil mit Faktoren zusammenhängen, die von Schulen und Lehrkräften nicht oder kaum beeinflusst werden können, wie etwa der sozio-ökonomische Status der Schüler/innen und ihrer Eltern (Baumert, Stanat, & Watermann, 2006; Ditton & Krüsken, 2007) oder der Migrationshintergrund (Autorengruppe Bildungsberichterstattung, 2012; Pant et al., 2013).

Konsequenterweise wird daher auch und gerade aus Schulen immer wieder die Forderung laut, dass für eine angemessene Interpretation der Leistungen in den Vergleichs- und Prüfungsarbeiten die Zusammensetzung der Schülerschaft zu berücksichtigen sei und die "rohen" Leistungsergebnisse um den Einfluss von Kontextfaktoren adjustiert werden müssten. Wenn Adjustierungen standardmäßig bei den jährlich anfallenden Vergleichs- und Prüfungsarbeiten vorgenommen werden sollen, sind nicht nur leistungsrelevante Kontextfaktoren heranzuziehen, sondern muss ein ressourcenschonendes Verfahren gefunden werden, das routinemäßig eingesetzt und zudem den Lehrkräften und einer breiteren Öffentlichkeit verständlich gemacht werden kann (Kuhl et al., 2011). In dem Beitrag soll zunächst der bis 2013 verwendete Berliner Adjustierungsansatz und seine Weiterentwicklung skizziert werden. Diese Ansätze lassen sich als Spielarten eines Adjustierungstypus auffassen, bei dem Erwartungswerte durch Vergleichsgruppen von Schulen mit ähnlicher Schülerschaft gewonnen werden (Fiege, Reuther, & Nachtigall, 2011). Beide verwenden nur zwei Kontextmerkmale, nämlich die Zuzahlungsbefreiung bei Lernmitteln als Sozialindikator und die nichtdeutsche Herkunftssprache als Indikator des Migrationshintergrunds. Abschließend werden darüber hinausgehende Ansätze dargestellt, die weitere Merkmale berücksichtigen.

Erkenntnisleitende Untersuchungsfragen sind: 1) Wie ist das Beziehungsgefüge zwischen Kontext- und Leistungsmerkmalen sowohl über die Zeit hinweg als auch in Abhängigkeit von Schulart und Region? 2) Wie gut sind die verschiedenen Ansätze im Vergleich: Hinsichtlich statistischer Kriterien, in ihrer Tauglichkeit für einen Routinebetrieb und in ihrer Kommunizierbarkeit?

Datengrundlage der Analyse sind die Mathematikergebnisse von rund 50 Realschulen und 90 Gymnasien aus dem Mittleren Schulabschluss und den Vergleichsarbeiten in der achten Jahrgangsstufe seit dem Schuljahr 2007/2008 sowie dazugehörige Verwaltungsdaten. Die Ermittlung von Vergleichswerten auf Schulebene basiert sowohl auf der Bildung von Vergleichsgruppen als auch auf Regressionsanalysen.

Erste Ergebnisse zeigen, dass Kontextmerkmale weitaus stabiler sind als Leistungswerte, dass der Zusammenhang zwischen beiden Merkmalskomplexen mit der Schulart und mit den Berliner Stadthälften Ost/West variiert und begründet nachvollziehbar den Wechsel des Adjustierungsansatzes ab 2014. Zudem werden Alternativen der weiteren Entwicklung in einem Prozess des Abwägens von Aufwand und Ertrag dargestellt.

Autorengruppe Bildungsberichterstattung (2012). Bildung in Deutschland 2012. Bielefeld: Bertelsmann

Baumert, J., Stanat, P. & Watermann, R. (2006). Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften.

Ditton, H. & Krüsken, J. (2007). Sozialräumliche Segregation und schulische Entwicklung. Diskurs Kindheits- und Jugendforschung, 1, 23-38.

Fiege, Chr., Reuther, F. & Nachtigall, Chr.: Faire Vergleiche? - Berücksichtigung von Kontextbedingungen des Lernens beim Vergleich von Testergebnissen aus deutschen Vergleichsarbeiten. Z f Bildungsforschung 1, 133-149

Kuhl, P. et al. (2011). Die Kontextuierung von Leistungswerten bei Vergleichs- und Prüfungsarbeiten: Verschiedene Wege, die Zusammensetzung der Schülerschaft in den Rückmeldungen an Schulen und die Schulinspektion zu berücksichtigen. In S. Müller, M. Pietsch & W. Bos (Hrsg.). Schulinspektion in Deutschland. Eine Zwischenbilanz in empirischer Sicht (S. 9-32). Münster: Waxmann.

Pant, H. A. et al. (Hg.): IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Münster: Waxmann.

## **Erstellung eines schulscharfen Sozialindex ohne Schüler- bzw. Elternbefragung**

Die Nutzung von Ergebnissen aus Leistungstests gehört in Deutschland mittlerweile zum gängigen Repertoire des Bildungsmonitorings (KMK, 2006). In diesem Zusammenhang wird jedoch immer wieder problematisiert, dass die soziokulturelle Zusammensetzung der Schüler einer Schule ein mehr oder weniger günstiges Lern- und Entwicklungsmilieu bedingt (Baumert et. al., 2006). Angesichts der dadurch bestimmten Entwicklungschancen für den einzelnen Schüler ergibt sich bzgl. der Arbeitsbedingungen und pädagogischen Herausforderungen, über alle Schüler hinweg, eine kumulative „Privilegierung oder Benachteiligung“ bestimmter Schulen (ebda., S. 97). Je nachdem, welche bildungsrelevanten Ressourcen in den Elternhäusern vorhanden sind, die die Schülerinnen und Schüler beim Lernen unterstützen, muss die Schule mehr oder weniger Ressourcen für pädagogische Prozesse aufwenden, um etwaige Mängel zu kompensieren (Bonsen et al., 2010). Nicht zuletzt um diesem Umstand bei der Einordnung der Testergebnisse Rechnung zu tragen haben sich in den letzten Jahren verschiedene Ansätze zur Charakterisierung der Schülerschaft über einen sogenannten „Sozialindex“ herauskristallisiert (z.B. BMBF, 2010). Grundlage für die Indexbildung sind in der Regel - in Anlehnung an die Arbeiten von Bourdieu (z. B. 1983) und Coleman (z. B. 1987) - Indikatoren, die Aufschluss über die jeweilige Verfügbarkeit bildungsrelevanter Ressourcen in der Familie geben. Üblicherweise werden hier neben der sozioökonomischen Stellung der Familie auch deren kulturelles und soziales Kapital unterschieden. Die meisten dieser Indikatoren können jedoch nur über eine direkte Erhebung (z. B. per Fragebogen) exakt bestimmt werden, was im Falle einer flächendeckenden Erhebung je Bundesland sowohl datenschutzrechtlich als auch ressourcentechnisch eine erhebliche Herausforderung darstellt. Inwieweit es möglich ist, Indices zu entwickeln, die ausschließlich auf befragungsunabhängig zugänglichen Daten basieren, bleibt daher eine zentrale Frage. Einen Ansatz liefert z. B. Schräpler (2011), welcher die entsprechenden Indikatoren über sozialgeographische Merkmale des Schulstandortes erfasst. Dieser Ansatz bringt jedoch die Schwierigkeit mit sich, dass bestimmte Schulen auch Schüler aus schulfernen Standorten rekrutieren, die bei diesem Ansatz nicht schulscharf erfasst würden.

Im folgenden Vortrag wird ein Ansatz vorgestellt, bei dem Postleitzahlen der Schülerwohnorte aus den Adressdatenbanken der Schulen dazu genutzt werden, sozialgeographische Merkmale schulscharf zu ermitteln und für die Entwicklung eines Sozialindex zu nutzen. Die Studie zeigt, dass aufgrund von sozialgeographischen Merkmalen im Zusammenspiel mit Informationen zum Migrationshintergrund der Schüler ein belastbarer Sozialindex ohne direkte Befragung erstellt werden kann. Ferner zeigt sich, dass das Merkmal Migrationshintergrund differenziert erhoben werden muss, da dieser Indikator je nach Definition für eine Indexbildung mehr oder weniger geeignet ist. Abschließend werden Möglichkeiten zur Kompensation fehlender Angaben von Schülerpostleitzahlen diskutiert.

- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In: J. Baumert, P. Stanat, R. Watermann (Hg.). *Herkunftsbedingte Disparitäten im Bildungswesen* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bundesministerium für Bildung und Forschung (BMBF)(Hg.)(2010). *Zur Konstruktion von Sozialindizes*. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung. Berlin: BMBF.
- Bonsen, M., Bos, W., Gröhlich, C. & Wendt, H. (2010). Der Index zur Erfassung der sozialen Komposition von Einzelschulen. In: Bundesministerium für Bildung und Forschung (BMBF) (Hg.), *Zur Konstruktion von Sozialindizes*. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung (S. 15-30). Berlin: BMBF.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hg.), *Soziale Ungleichheiten, Soziale Welt, Sonderband 2*, 183-198. Göttingen: Schwartz.
- Coleman, J. S. (1987). Families and schools, *Educational Researcher*, 16(6), 32-38.
- KMK (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. (1. Aufl.) Neuwied: Luchterhand.
- Schräpler, J.P. (2011). Konstruktion von SGB II – Dichten als Raumindikator und ihre Verwendung als Indikator im Rahmen der Sozialberichterstattung am Beispiel der „sozialen Belastung“ von Schulstandorten in NRW – ein Kerndichte-Ansatz. *ASTA Wirtschafts- und Sozialstatistisches Archiv*, 5 (2), 97-124.



## ***Textqualität erfassen – Computergestützte Auswertung als innovative Perspektive für die fachbezogene Bildungsforschung?***

Chair(s): Sanna Pohlmann-Rother (Otto-Friedrich-Universität Bamberg)

Obwohl die Schreibkompetenz für Berufsfähigkeit und Lebensbewältigung genauso wichtig ist wie die Lesekompetenz, wird sie in den nationalen und internationalen Leistungsvergleichen bislang selten erfasst, vermutlich des Aufwands bei der Erhebung und Auswertung wegen. In allen Klassenstufen gehört das Verfassen von Texten zu den Bildungszielen des Deutschunterrichts (Bildungsstandards für die Grundschule: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005). In den Untersuchungen des IQB sind deshalb regelmäßig große Anzahlen von Schülertexten objektiv, reliabel und valide zu beurteilen. Zwei Verfahren sichern die notwendige Beurteilerübereinstimmung und die Reliabilität der Urteile: (1.) die analytische Beurteilung auf der Grundlage von Kriteriensätzen und (2.) die Mehrfachbeurteilung nach globalem Ersteindruck, bei der mehrere Beurteiler einen Text nach ihrem ersten „holistischen“ Eindruck einstufen und aus diesen Urteilen ein Mittelwert gebildet wird (Baurmann, 2006; Grzesik & Fischer, 1984). Im Unterschied zum holistischen Urteil erlauben Kriterienkataloge differenzierte Urteile über die Qualität von Texten im Hinblick auf verschiedene Aspekte, wie z. B. die formale, sprachliche, strukturelle und inhaltliche Qualität. Bei Schreibkompetenzuntersuchungen im Rahmen der empirischen Bildungsforschung werden meist holistische und analytische Beurteilungen kombiniert, wie z.B. im Rahmen der seit 2005 durchgeführten Evaluation der Bildungsstandards (Böhme, Bremerich-Vos & Robitzsch, 2009). Insbesondere die analytische Beurteilung anhand eines differenzierten Kriterienkatalogs ist jedoch bei großen Stichproben besonders kosten- und zeitintensiv. Vor diesem Hintergrund wird in dem Vortrag der Frage nachgegangen, ob und inwieweit sich im Rahmen eines analytischen Ratings Kriterien auch computergestützt und damit effizient und weniger zeitintensiv erfassen lassen. Die verwendeten Texte stammen aus dem DFG-Projekt NaSch1 (Narrative Schreibkompetenz in Klasse 1) das auf der Schreib- und Erzählforschung und der Schreib- bzw. Deutschdidaktik sowie der empirischen Bildungsforschung aufbauend die Textqualität in Klasse 1 in Abhängigkeit von individuellen, familiären und unterrichtlichen Bedingungen untersucht. Die Datengrundlage bilden 618 Texte, die in 37 Klassen Mitte der ersten Jahrgangsstufe erhoben wurden. Die Texte wurden sowohl holistisch als auch analytisch anhand eines umfangreichen Kriterienkataloges ausgewertet. Auf Basis der interdisziplinären Kooperationsbeziehung zwischen NaSch1 und dem Ubiquitous Knowledge Processing Lab wurde anhand dieser Texte darüber hinaus eine computergestützte Erfassung der Textqualität realisiert. Im Vortrag werden die manuelle und die computergestützte Erfassung von Textqualität abgebildet und ihre Chancen und Grenzen diskutiert. Das Vorgehen bei der computergestützten Auswertung verwendet zunächst die manuell vorgenommenen Textbewertungen als Referenzdaten für eine algorithmische Modellierung der Qualitätskriterien. Je nach Typ der einzelnen Qualitätskriterien werden diese entweder extensional oder intensional im Bewertungssystem modelliert. Eine extensionale Modellierung bedeutet hierbei, dass Kriterien mittels händisch erstellter, expliziter Regeln direkt am Text gemessen werden können. Intensional modellierte Qualitätskriterien werden hingegen mit Hilfe maschinellen Lernens erfasst. Dabei wird der Text zunächst als Summe seiner qualitätsrelevanten Eigenschaften repräsentiert. Ein Lernalgorithmus nutzt dann die menschlichen Referenzbewertungen, um die für die Vorhersage eines Qualitätskriteriums relevanten Texteeigenschaften zu identifizieren. Auf diese Weise können automatische Qualitätsmaße von manuell erstellten Referenzdaten erlernt werden. Im Vortrag werden mit diesem kooperativen Vorgehen Möglichkeiten der Vernetzung von Bildungswissenschaften und Informatik veranschaulicht, um bewährte Instrumente und Methoden zu optimieren und neue Perspektiven zu eröffnen.

Baurmann, J. (2006). Schreiben - Überarbeiten - Beurteilen. Ein Arbeitsbuch zur Schreibdidaktik. Seelze: Kallmeyer.

Böhme, K., Bremerich-Vos, A. & Robitzsch, A. (2009). Aspekte der Kodierung von Schreibaufgaben. In D. Granter, O. Köller, A. Bremerich-Vos & et al. (Hrsg.), Bildungsstandards Deutsch und Mathematik (S. 290–329). Weinheim, Basel: Beltz.

Grzesik, J. & Fischer, M. (1984). Was leisten Kriterien für die Aufsatzbeurteilung? Theoretische, empirische und praktische Aspekte des Gebrauchs von Kriterien und der Mehrfachbeurteilung nach globalem Ersteindruck. Opladen: Westdt. Verl.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). Bildungsstandards im Fach Deutsch für den Primarbereich (Jahrgangsstufe 4). Beschluss vom 15.10.2004. München: Wolters Kluwer.

## Vergleich analytischer und holistischer Ratingstrategien zur Bestimmung von Textqualität

Trotz der zunehmenden Prozessorientierung in der gegenwärtigen Schreibdidaktik und -forschung ist das Interesse an der Bestimmung von Textqualität ungetrübt. Nach wie vor besteht in diesem Zusammenhang ein grundlegender Forschungsbedarf (Nutz, 2003). In der heutigen Debatte um eine objektive und gerechte Leistungsbewertung steht jedoch weniger die hinreichend belegte unterschiedliche Aufsatzbewertung (Ingenkamp, 1975) im Fokus, sondern etwa die Frage nach der Optimierung von Bewertungskriterien.

Als empirisch fundierte Instrumente zur Bestimmung von Textqualität gelten dabei die holistische Beurteilung sowie der analytische Kriterienkatalog - beide Verfahren sichern die notwendige Beurteilerübereinstimmung bzw. die Reliabilität der Urteile (Baurmann, 2006; Grzesik & Fischer, 1984).

Der Vortrag verfolgt das Ziel, diese beiden Bewertungsinstrumente miteinander in Beziehung zu setzen und auf dieser Basis einzelne Komponenten der Textqualität empirisch zu erfassen. Im DFG-Projekt NaSch1 (Narrative Schreibkompetenz in Klasse 1), das die Textqualität in Abhängigkeit von individuellen, familiären und unterrichtlichen Bedingungen untersucht, wurden 618 Texte von Erstklässlern sowohl analytisch anhand eines umfangreichen Kriterienkatalogs erfasst sowie nach ihrem „holistischen“ Eindruck beurteilt. Die kriteriale Textbewertung setzt sich dabei aus ausdifferenzierten Einzelkomponenten wie Wortschatz oder Kohärenz zusammen, während der holistische Eindruck im Rahmen einer mehrstufigen Skala ein Globalurteil bewirkt.

Die Grundlage für den Vergleich der Instrumente bildet eine Rasch-Skalierung der Textqualität anhand der kriterialen Bewertung. Damit lassen sich zum einen das Konstrukt sowie die Dimensionen von Textqualität abbilden. Zum anderen kann der Frage nachgegangen werden, inwieweit das holistische Rating mit der Gesamtskala des analytischen Ratings übereinstimmt oder ggf. nur eine Subdimension des Konstrukts Textqualität wiedergibt. Auf diese Weise lässt sich belegen, anhand welcher Kriterien ein Text ein holistisches Qualitätsurteil erhält. Daraus lassen sich auch Konsequenzen für die Aufsatzbeurteilung auf Seiten der Lehrkräfte ableiten

Das Projekt NaSch1 insgesamt sowie die Fragestellungen des geplanten Vortrags verbinden linguistische, deutschdidaktische und grundschulpädagogische Perspektiven im Rahmen der empirischen Bildungsforschung.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G135

Baurmann, J. (2006). Schreiben - Überarbeiten - Beurteilen. Ein Arbeitsbuch zur Schreibdidaktik. Seelze (Velber): Kallmeyer.

Grzesik, J. & Fischer, M. (1984). Was leisten Kriterien für die Aufsatzbeurteilung? Theoretische, empirische und praktische Aspekte des Gebrauchs von Kriterien und der Mehrfachbeurteilung nach globalem Ersteindruck. Opladen: Westdt. Verl.

Ingenkamp, K. (1975). Die diagnostische Problematik des Aufsatzes als Prüfungsinstrument und die Bemühungen zur Verbesserung der Auswertungsqualität. In O. Beck & F.-J. Payrhuber (Hrsg.), Aufsatzbeurteilung heute. Problematik, Diagnose, Therapieansätze. (S. 16–31). Freiburg im Breisgau: Herder.

Nutz, M. (2003). Beurteilung sprachlicher Leistungen. In P. Klotz (Hrsg.), Didaktik der deutschen Sprache. Band 2. Ein Handbuch (S. 924–937). Stuttgart: UTB GmbH.

## **Literarische Textschwierigkeit interpretieren? Einschätzung der Textschwierigkeit literarischer Texte durch Experten mit unterschiedlichem Ausbildungsstand bzw. unterschiedlicher Berufserfahrung**

Die Studie wird angeleitet von der Frage, wie Studierende, Referendarinnen und Referendare sowie Lehrerinnen und Lehrer des Faches Deutsch die Schwierigkeit eines literarischen Textes bzw. dessen Verstehensanforderungen im Hinblick auf Schülerinnen und Schüler einer zehnten Jahrgangsstufe/Gymnasium einschätzen. Die Studie ist vor dem Hintergrund der Entwicklung eines Modells der Textschwierigkeit literarischer Texte zu sehen (vgl. Frickel, 2010; Frickel & Filla, 2013), da es bislang an einem etablierten Konzept fehlt, das als Instrument für die Textauswahl und Itementwicklung von Lesetests sowie für die Planung des Literaturunterrichts fungieren kann (vgl. Rupp, 2007). Das vorgeschlagene Modell, das eine Übersicht über die vorstellbaren zentralen Dimensionen von Textschwierigkeit aufführt und die Text-Leser-Kontext-Relation transparent macht, versucht eine Synthese aus bisher vorliegenden Ansätzen und Konzepten (s. Frickel & Filla, 2013). In diesem Modell wird die Rolle des Experten als Einflussgröße für die Bestimmung der Textschwierigkeit exponiert und lieferte für diese Studie die Grundlage zur Entwicklung unserer Untersuchungshypothesen. Der Experte als Einflussgröße wurde in den vorliegenden Ansätzen bisher vernachlässigt bzw. weisen Richard Bamberger und Richard Vanecek in ihrem Werk Lesen-Verstehen-Lernen-Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache (1984) darauf hin, dass eine objektive Bestimmung im Hinblick auf literarische Texte mit Hilfe der Sprachschwierigkeit als Prädiktor schwer möglich sei, da diese hierzu vielmehr „interpretiert“ (Bamberger & Vanecek, 1984, S. 49) werden müssten, d.h. der Experte eine wichtige Einflussgröße darstellt. Auch Thomas Zabka (2012), der ein Konzept für die Analyse von Verstehensanforderungen vorgeschlagen hat, räumt ein, dass im Zusammenhang mit seinem Ansatz Experteneinschätzungen zukünftig untersucht werden müssten (vgl. Zabka, 2012, S. 159). Der in den Jahren 2011/12 entwickelte Studie lag die Hypothese zugrunde, dass Lehramtsstudierende, Referendare und Lehrer des Faches Deutsch die Textschwierigkeit einer Kurzgeschichte unterschiedlich einschätzen bzw. unterschiedliche Prognosen hinsichtlich des Textverständnisses der Schülerinnen und Schüler abgeben. Der Expertenfragebogen enthielt dementsprechend vier große Fragekomplexe: Fragen zum sprachlichen bzw. poetologischen Schwierigkeitsgrad des Textes, zu verständnisrelevanten Dispositionen des Lesers, zur leser- bzw. themabezogenen Motivation und zum Interesse sowie eine Auswahl der aus meiner Unterrichtserfahrung mit diesem Text hergeleiteten Deutungshypothesen. Die Teilnehmer sollten anhand überwiegend sechsstufiger Ratingskalen aus ihrer Sicht als (gegenwärtige oder zukünftige) Experten Bewertungen zum Ausmaß der Textschwierigkeit, Motivationswirkung, verstehensrelevanten Leserdispositionen etc. vornehmen. Darüber hinaus wurden u.a. Angaben zu Geschlecht, Alter, Berufserfahrung und Studienfächern bzw. der Studienrichtung innerhalb des Faches Deutsch erhoben. Die Probanden setzten sich zusammen aus Lehrern, Referendaren und Studenten verschiedener Schulen (Gymnasien/Gesamtschulen aus NRW), Studienseminare (Gymnasium, NRW) und Universitäten. Unsere Analysen waren in erster Linie auf Unterschiedshypothesen konzentriert. Hinsichtlich der Einschätzung der sprachlichen bzw. poetologischen Textschwierigkeit (d.h. z.B. Vokabular, Syntax, Metaphorik, Figurendarstellung, Handlungsaufbau usw.) zeigte sich ein homogenes Ergebnis entlang der Metadaten aller Gruppen ohne signifikante Unterschiede. Signifikante Unterschiede ergaben sich dabei insb. bei Items, die die Bewertung der Leserdispositionen und der schwierigkeitskonstituierenden Textmerkmale aus der Schülerperspektive abdecken. Die Ergebnisse bieten Ansatzpunkte für die Diskussion, insb. im Hinblick auf die Relevanz des Experten bei der Bestimmung der Textschwierigkeit und für Aspekte, die im Zusammenhang mit der DeutschlehrerInnen-Bildung zukünftig stärker reflektiert werden sollten.

Bamberger, Richard/Vanecek, Erich: Lesen – Verstehen – Lernen – Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache. Wien 1984.

Frickel, Daniela A.: Textschwierigkeit als Parameter im Prozess literarischen Textverstehens. Erörtert am Beispiel Kleiner Prosa. In: Rupp, Gerhard/Boelmann, Jan/Frickel, Daniela A. (Hg.): Aspekte literarischen Lernens. Junge Forschung in der Deutschdidaktik. Münster 2010 (= Leseforschung 2), 113-128.

Frickel, Daniela A./Filla, Manfred: Textschwierigkeit „interpretieren“ – Dimensionen und Problemgrößen bei der Einschätzung der Textschwierigkeit literarischer Texte. In: Frickel, Daniela A./Boelmann, Jan (Hg.): Literatur – Lesen – Lernen. Frankfurt/M. 2013, 105-131.

Köster, Juliane: Die Bedeutung des Vorwissens für die Lesekompetenz. In: Abraham, Ulf (u.a.) (Hg.): Deutschdidaktik und Deutschunterricht nach PISA Freiburg i.Br.: Fillibach 2003, 90-105.

Leupner, Martin/Saupe, Anja: Textverstehen im Literaturunterricht und Aufgaben. Baltmannsweiler 2008.

Rupp, Gerhard: Textseitige Einflussfaktoren auf die Leseleistung. In: Hinrichs, Nicole/Limburg, Annika (Hg.): Gedankenstriche – Reflexionen über Sprache als Resource. Für Wolfgang Boettcher zum 65. Geburtstag. Tübingen 2010, 111-128.

Spinner, Kaspar H.: Wie Fachwissen das literarische Verstehen stört und fördert. In: Irene Pieper/Dorothee Wieser (Hg.): Fachliches Wissen und Literarisches Verstehen. Frankfurt/M. (= Beiträge zur Literatur- und Mediendidaktik) 2011, 53-70.

Willenberg, Heiner: Ein handhabbares System, um Textschwierigkeiten einzuschätzen. Vorschläge für eine Textdatenbank von Sachttexten. In: Fix, Martin; Jost, Roland (Hg.): Sachttexte im Deutschunterricht. Für Karlheinz Fingerhut zum 65. Geburtstag.

Baltmannsweiler 2005, 94-106.

Winkler, Iris: «Ich finde den Text schwierig ...» Textschwierigkeit als Resultat einer Wechselbeziehung von Text- und Lesermerkmalen. In: Frickel, Daniela A./Boelmann, Jan (Hg.): Literatur – Lesen – Lernen. Frankfurt/M. 2013, 393-409.

Zabka, Thomas: Analyserituale und Lehrerüberzeugungen. Theoretische Untersuchung vermuteter Zusammenhänge. In: In: Pieper, Irene/Wieser, Dorothee (Hg.): Fachliches Wissen und Literarisches Verstehen. Frankfurt/M. 2011 (= Beiträge zur Literatur- und Mediendidaktik), 35-52.

Zabka, Thomas: Didaktische Analyse literarischer Texte. Theoretische Überlegungen zu einer Lehrerkompetenz. In: Frickel, Daniela A./Kammler, Clemens/Rupp, Gerhard (Hg.): Literaturdidaktik im Zeichen von Kompetenzorientierung und Empirie. Kompetenzen und Probleme. Freiburg i.Br. 2012, 139-162.



## **Zwei Perspektiven auf Rechtschreibung: Sprachdidaktik und Computerlinguistik entwickeln ein Instrument zur Rechtschreibdiagnostik**

Das DFG-geförderte Forschungsprojekt „Automatisierte Rechtschreibdiagnose“ verbindet die Disziplinen Sprachdidaktik und Computerlinguistik. Ziel des Projekts ist die Entwicklung einer Software zur automatischen Rechtschreibanalyse von frei verfassten Lernerschreibungen. Eine solche Software bedient zwei Desiderata:

1. Für den didaktischen Alltag ist es damit möglich, differenzierte Aussagen über Stärken und Schwächen in der Rechtschreibleistung von Schreibernern zu bekommen. Die besondere Facette der Rechtschreibung in eigens verfassten Texten bietet – im Gegensatz zu den bekannten standardisierten Tests - besondere Einblicke in die Rechtschreibkompetenz beim „integrierten Schreiben“, welches als die „authentischere“ Schreibdimension gilt (vgl. Ludwig, 1995; Fay, 2010).

2. In der empirischen Forschung zur Orthographiedidaktik ist es damit erstmals möglich, sehr große Stichprobenumfänge ökonomisch zu untersuchen. Eine qualitative Rechtschreibanalyse auf Basis eines differenzierten Kategorienrasters (Baustein zahlreicher empirischer Studien zur Rechtschreibdidaktik) ist manuell mit einem prohibitiv hohen Arbeitsaufwand verbunden. Eine automatische Analyse erleichtert und flexibilisiert das Forschen in diesem Gebiet.

Die Software, von der wir bislang einen Prototypen entwickelt haben, teilt sich in zwei Module:

Modul 1 ermittelt automatisch aus dem Geschriebenen des Schülers die orthographisch korrekte Version des Textes. Dabei werden Technologien aus der automatischen Sprachverarbeitung angewandt, indem nämlich das Geschriebene erst automatisch in eine Phonemfolge („Sprachsynthese“) und diese dann wiederum in die wahrscheinlichste korrekte Wortfolge („Spracherkennung“) übersetzt wird. Als Sprachsynthesetool benutzen wir MARY (Schröder & Trouvain, 2003); als Spracherkennungstool benutzen wir das Janus Recognition Toolkit (JRTk) (Soltau et al., 2001).

Modul 2 stellt dann automatisch die Lernerschreibung und ihre korrigierte Fassung einander gegenüber. Sowohl Fehler als auch Richtigschreibungen können dann – ebenfalls automatisch - annotiert und klassifiziert werden. Im Ergebnis erscheint ein differenziertes Analyseprofil mit normalisierten Angaben zu 36 Rechtschreibkategorien (Kategorienraster von Fay, 2010), je nach Input für bspw. einen einzelnen Text, einen Klassensatz Texte oder einer umfangreichen Datenbank mit sehr vielen Texten.

In unserem Vortrag stellen wir zum einen die Software in ihrem Vorgehen vor (vgl. u.a. auch Fay, Berkling, & Stüker, 2012) und zum anderen erste Beispiele aus der Anwendung auf einen Schülertextkorpus. Bei dem Korpus handelt es sich um 1.767 Texte (153.026 Wörter) von Schülern der Klasse 2 bis 8. Die Texte wurden in den Jahren 2011 bis 2013 in Schulen in und um Karlsruhe erhoben, wobei die Schreibaufgaben so offen gestaltet waren, dass frei konzipierte Lernertexte entstehen konnten (Zum Korpus: Berling/Fay/Ghayoomi/Hein/Stüker 2014 i.V.).

Berkling, K.; Fay, J.; Stüker, S. (2011): Speech Technology-based Framework for Quantitative Analysis of German Spelling Errors in Freely Composed Children's Texts. In Proceedings of the 2011 workshop of the ISCA (International Speech Communication Association) Special Interest Group (SIG) on Speech and Language Technology, 24.-26. August 2011, Venedig, Italien. [reviewed]

Berkling, K.; Fay, J.; Ghayoomi, M.; Hein, K.; Stüker, S. (2014, i.V.): A database of freely written texts of german school students for the purpose of automatic spelling error classification. LREC 2014.

Fay, J.; Berkling, K. (2013): Rechtschreibdiagnostik. In: Fay (Hrsg.): (Schrift-) Sprachdiagnostik heute. Schneider-Verlag: Baltmannsweiler. S. 84-108.

Fay, J.; Berkling, K.; Stüker, S. (2012): Automatische Analyse von Rechtschreibfähigkeit auf Basis von Speech-Processing-Technologien. In: Didaktik Deutsch. Heft 33. Schneider Verlag: Baltmannsweiler. S. 14-37. [reviewed]

Fay, J. (2010): Die Entwicklung der Rechtschreibkompetenz beim Textschreiben. Eine empirische Untersuchung von Klasse 1 bis 4. Peter Lang: Frankfurt/M.

Ludwig, O.: Integriertes und nicht-integriertes Schreiben. Zu einer Theorie des Schreibens: eine Skizze. In: Baurmann, Jürgen; Weingarten, Rüdiger (Hrsg.): Schreiben. Prozesse, Prozeduren und Produkte. Opladen 1995, S. 273-288.

Schröder, M.; Trouvain, J.: The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching. In: International Journal of Speech Technology, 2003, pp. 365-377.

Soltau, H.; Metzke, F.; Fügen, C.; Waibel, A.: A one pass- decoder based on polymorphic linguistic context assignment. In Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU '01), pages 214-217, Madonna di Campiglio Trento, Italy, December 2001.

Stüker, S.; Fay, J.; Berkling, K. (2011): Towards Context-dependent Phonetic Spelling Error Correction in Children's Freely Composed Text for Diagnostic and Pedagogical Purposes. In Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), 28.-31- August 2011, Florenz, Italien. [reviewed]



## Expertise in der Domäne Schreiben bei angehenden Lehrpersonen

Leistungsstudien zur Schreibkompetenz von SchülerInnen zeigen, dass ein gewichtiger Anteil der Kinder und Jugendlichen nur über geringe bzw. unzureichende Schreibkompetenzen verfügt (vgl. Neumann & Lehmann, 2008; Salahu-Din, Persky & Miller, 2008). Soll der Anteil der SchülerInnen mit unzureichenden Schreibkompetenzen verringert werden, rückt u.a. der Schreibunterricht ins Zentrum der Aufmerksamkeit. So beklagt Juzwik (2010) eine Kluft zwischen Schreibforschung und -praxis, indem Befunde zu empirisch wirksamen Verfahren v.a. von Vertretern eines «natürlichen» Schreibunterrichts abgelehnt würden. Ähnliches konstatieren Troia und Maddox (2004) bei Mittelstufen-Lehrpersonen. Ob Empfehlungen zu empirisch wirksamen Fördermassnahmen in die Schulpraxis implementierbar sind, lässt sich also nur bestimmen, wenn mehr über den Schreibunterricht in der Schulpraxis bekannt ist, als dies bisher der Fall ist (Gilbert & Graham, 2010). An dieser Stelle setzt das Projekt «Novizinnen und Novizen im Schreibunterricht» an und geht den folgenden Leitfragen A und B nach:

- A: Welches fachliche und fachdidaktische Wissen und welche Überzeugungen bringen angehende Lehrpersonen für den Schreibunterricht mit und wie verändern sich diese im Verlauf des ersten Berufsjahrs?

Um die Schreibunterrichtspraxis angemessen beschreiben zu können, werden in der Teilstudie B Unterrichtsbeobachtungen durchgeführt, da Selbstauskünfte von Lehrpersonen – wie sie in Teilstudie A erhoben werden – nur bedingt Aufschluss über die tatsächliche Praxis geben (ebd.; Kiuvara, Graham & Hawken, 2009).

- B: Auf welcher schreibdidaktischen Basis gestalten Lehrpersonen im ersten Berufsjahr ihren Schreibunterricht?

Die Ergebnisse werden auf Implikationen für die Aus- und Weiterbildung von Lehrpersonen geprüft. In der Teilstudie A wurde im Mai 2013 bei Studierenden des Ausbildungsgangs Primarstufe an der Pädagogischen Hochschule FHNW am Ende ihrer Ausbildung ein Fragebogen eingesetzt (n=99; Rücklauf 46%): Erfasst wurde die Selbstwirksamkeit in der Domäne Schreiben (Bruning, Dempsey, Kauffmann, McKim & Zumbunn, 2012) sowie ihre Selbstwirksamkeit, Schreiben zu unterrichten (hierzu wurde die Skala von Bruning et al., 2012, adaptiert, da sich die Skala von Graham, Harris, Fink & MacArthur, 2001, in einem Pretest nicht bewährte). Des Weiteren wurde fachliches Wissen in der Domäne Schreiben mit zwei offenen Aufgaben auf der Basis von Barbeiro (2010) erfasst: Zum einen sollen die ProbandInnen ihren eigenen Schreibprozess, zum anderen jenen von SchülerInnen mit eher geringen Schreibkompetenzen beschreiben. Fachdidaktisches Wissen wurde mit zwei offenen Aufgaben erhoben, bei denen die ProbandInnen einen Schülertext schriftlich beurteilen und begründen sowie einen Lehrerkommentar zu diesem Text verfassen sollten (hierzu wurde eine Aufgabe aus dem Projekt MyMoment adaptiert, vgl. Schneider, Wiesner, Lindauer & Furger, 2012). Des Weiteren wurden die Studierenden aufbauend auf Cutler & Graham (2008) und Graham, Harris, Fink-Chorzempa & MacArthur (2003) u.a. gefragt, inwiefern sie sich durch die Ausbildung an der Pädagogischen Hochschule auf den Schreibunterricht vorbereitet fühlen. Die Befragung wurde mit dem Online-Tool TOSS, das an der Pädagogischen Hochschule FHNW im Rahmen von Schreibkompetenzabklärungen bei Studierenden eingesetzt wird, durchgeführt. Um allfällige Veränderungen im Verlauf des ersten Berufsjahrs erfassen zu können, wird die Befragung im Mai/Juni 2014 wiederholt. Die Skalen zur Selbstwirksamkeit wurden faktorenanalytisch ausgewertet: In jener zum eigenen Schreiben ergaben sich analog zu Bruning et al. (2012) die Faktoren Ideenfindung ( $\alpha = .84$ ), Norm ( $\alpha = .84$ ) sowie Selbstregulation ( $\alpha = .70$ ); Ähnliches gilt für die Selbstwirksamkeit, Schreiben zu unterrichten ( $\alpha = .72-80$ ). Die offenen Aufgaben zum fachlichen Wissen werden auf der Basis von Breetvelt, van den Bergh & Rijlaarsdam (1994) sowie u.a. von Troia (2006) inhaltsanalytisch ausgewertet – daraus wird ein Summenindex Fachwissen berechnet –, der und jeweils von zwei Ratern kodiert. Die diesbezüglichen Auswertungen sind im Gange. Für die Teilstudie B wird aus A kriteriengeleitet eine Teilstichprobe gezogen (n=8). Dazu werden die Variablen Fachwissen sowie Selbstwirksamkeit, Schreiben zu unterrichten, herangezogen. Erste Ergebnisse aus der Teilstudie A können im März 2014 vorgestellt werden.

Barbeiro, L.F. (2010). What happens when I write? Pupils' writing about writing. *Reading and Writing*, 24, 813–834.

Breetvelt, I., van den Bergh, H. & Rijlaarsdam, G. (1994). Relations between Writing Processes and Text Quality: When and How? *Cognition and Instruction*, 12 (2), 103–123.

Bruning, R., Dempsey, M., Kauffmann, D.F., McKim, C. & Zumbunn, S. (2012). Examining Dimensions of Self-Efficacy for Writing. *Journal of Educational Psychology*, (Advance online publication).

Cutler, L. & Graham, S. (2008). Primary Grade Writing Instruction: A National Survey. *Journal of Educational Psychology*, 100 (4), 907–919. Verfügbar unter: <http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ823719> [Stand: 3.8.2011].

Gilbert, J. & Graham, S. (2010). Teaching Writing to Elementary Students in Grades 4–6: A National Survey. *The Elementary School Journal*, 110 (4), 494–518.

Graham, S., Harris, K.R., Fink, B. & MacArthur, C. (2001). Teacher Efficacy in Writing: A Construct Validation With Primary Grade Teachers. *Scientific Studies of Reading*, 5 (2), 177–202.

Graham, S., Harris, K.R., Fink-Chorzempa, B. & MacArthur, C. (2003). Primary Grade Teachers' Instructional Adaptations for Struggling Writers: A National Survey. *Journal of Educational Psychology*, 95 (2), 279–292.

Juzwik, M.M. (2010). Challenges and Possibilities for Professional Development in Writing Instruction. In G.A. Troia, R.K. Shankland & A. Heintz (Hrsg.), *Putting Writing Research into Practice. Applications for Teacher Professional Development* (S. 259–275). New York/London: Guilford Press.

Kiuvara, S.A., Graham, S. & Hawken, L.S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, 101 (1), 136–160.

Neumann, A. & Lehmann, R.H. (2008). Schreiben Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie* (S. 89–103). Weinheim/Basel: Beltz.

Salahu-Din, D., Persky, H. & Miller, J. (2008). The Nation's Report Card: Writing 2007 (Statistical Analysis Report No. NCES 2008468). US Department of Education. Verfügbar unter:

<http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2008468> [Stand: 28.12.2010].

Schneider, H., Wiesner, E., Lindauer, T. & Furger, J. (2012). Kinder schreiben auf einer Internetplattform. Resultate aus der Interventionsstudie myMoment2.0. *dieS-online*, (2), 1–37.

Troia, G.A. (2006). Writing Instruction for Students with Learning Disabilities. In C.A. MacArthur, S. Graham & J. Fitzgerald (Hrsg.), *Handbook of Writing Research* (S. 324–336). New York: Guilford Press.

Troia, G.A. & Maddox, M.E. (2004). Writing Instruction in Middle Schools: Special and General Education Teachers Share Their Views and Voice Their Concerns. *Exceptionality*, 12 (1), 19–37.





## **Wie wirkt additive Sprachförderung? Analyse einer 2-jährigen Sprachförderprogramms für Grundschüler nichtdeutscher Erstsprache**

Chair(s): Raphaela Schätz (LMU München)

**Problemstellung:** Kompetenzstudien belegen weiterhin den Rückstand von Schülern nichtdeutscher Erstsprache gegenüber ihren Mitschülern (Schwippert, Wendt & Tarelli, 2012). Dennoch gibt es im deutschen Sprachraum kaum Sprachförderprogramme, die adäquat und längerfristig umgesetzt werden und deren Wirkungen analysiert sind. **Theoretischer Rahmen:** Ziele der Sprachförderung in der Zweitsprache von 5-8-jährigen Kindern sind gut entwickelte Fähigkeiten in Wortschatz und Grammatik, um zielsprachliche Sätze verstehen und produzieren zu können. Das mündliche Erzählen übt bereits Merkmale der Bildungssprache und ist daher ein weiteres wichtiges Förderziel. Der Einfluss der Erzählfähigkeit auf spätere schriftsprachliche Leistungen ist auch empirisch belegt (O'Neill, Pearce & Pick, 2004). Zur Förderung des Zweitspracherwerbs sollte dem Kind ein strukturierter Sprachinput angeboten werden, der die natürlichen Entwicklungsstufen berücksichtigt (Pienemann, 1989) und eine Stufe über dem Lernniveau liegt (Vygodsky, 1986). Der Sprachinput sollte in authentischen Kommunikationssituationen dargeboten werden, um echte und damit motivierende Sprechanlässe zu schaffen. Der Lehrende sollte Lerneräußerungen durch richtiges Wiederholen modellieren und ermöglicht dadurch auch implizites Lernen (Long, 2007). Dieser Förderansatz kann vor allem in additiver Sprachförderung realisiert werden, die zusätzlich zum regulären Unterricht in Kleingruppen stattfindet. Bevor Wirkungen von Interventionen untersucht werden, ist es wichtig die Umsetzungstreue zu überprüfen, die als Maß der Übereinstimmung zwischen Konzept und tatsächlicher Umsetzung verstanden wird (Mowbray et al., 2003). Im Kontext von Sprachförderung sind hier die inhaltliche, didaktische und organisationale Ebene relevant.

**Kontext:** Untersucht wird ein 2-jähriges, additives Sprachförderprogramm für Schulanfänger nichtdeutscher Erstsprache, das seit 2011 an fünf Berliner Grundschulen pilotiert wird. Förderziele sind die Erweiterung der Wortschatz-, Grammatik- und Erzählfähigkeit. Das Programm basiert auf dem Fördermaterial „Deutsch für den Schulstart“ (Kaltenbacher, 2009), das die oben genannten theoretischen Überlegungen beinhaltet, und wurde zur Anpassung an Zielgruppe und Rahmenbedingungen inhaltlich und didaktisch ergänzt.

**Fragestellungen:** (1) Wie wird die Sprachförderung inhaltlich, didaktisch und organisational umgesetzt? (2) Wie wirkt die Sprachförderung auf den Spracherwerb der Kinder?

**Methode:** Stichprobe: Die Stichprobe umfasste sieben Fördergruppen mit jeweils fünf bis sieben Kindern (n = 39) und entsprechende Vergleichsgruppen (n = 37) aus fünf Grundschulen. Die Auswahl erfolgte anhand des Pretests, der die schwächsten Kinder der Fördergruppe und die nächstschwächsten Kinder der Vergleichsgruppe zuwies. Die Vergleichsgruppe hatte ein höheres Ausgangsniveau, aber dennoch Sprachförderbedarf.

**Instrumente und Design:** Mit den Förderkräften wurden zu zwei Zeitpunkten strukturierte Interviews geführt. Zusätzlich wurden kontinuierlich theoriebasierte Checklisten zu Inhalt, Didaktik und Durchführung der einzelnen Förderstunden eingesetzt. Ein standardisierter Sprachstandtest (Elizitationsverfahren) erfasste in einem Prä-Post-Kontrollgruppendesign die Fähigkeiten in Wortschatz, Grammatik und Erzählen zu vier Messzeitpunkten.

**Ergebnisse:** (1) Umsetzung: Nach Einschätzung der Förderkräfte ließen sich die Inhalte und didaktischen Prinzipien des Konzepts sehr gut umsetzen und waren für die Zielgruppe sehr passend. Auf organisationaler Ebene unterschied sich die Häufigkeit der durchgeführten Sprachförderung zwischen den Fördergruppen, so dass in drei Fördergruppen mindestens 80 % der geplanten Förderzeit und in einer Fördergruppe unter 50 % realisiert wurde.

(2) Wirkung: Die Fördergruppe zeigte in den Bereichen Wortschatz und Satzbau, sowie Erzählen eine stärkere Entwicklung als die Vergleichsgruppe. Die Kinder der Fördergruppe glichen sich in diesen Bereichen innerhalb von zwei Jahren an die Vergleichsgruppe an, die ein höheres Ausgangsniveau hatte und keine Sprachförderung erhielt.

Beim Artikelgebrauch (Bereich Grammatik) ließ sich kein Unterschied zwischen Förder- und Vergleichsgruppe feststellen, da sich beide Gruppen parallel entwickelten.

**Diskussion:** Die Studie zeigt, dass additive Sprachförderung, die eine entsprechende Didaktik verfolgt, passende Inhalte anbietet und adäquat umgesetzt wird, für die Schwächsten eines Jahrgangs geeignet sein kann, um den Sprachstand in den Bereichen Wortschatz und Satzbau, sowie Erzählen zu erhöhen. Damit leistet die Studie einen wichtigen Beitrag zur Diskussion um die Wirkungen von additiver Sprachförderung. Eine Herausforderung stellt jedoch die Gewährleistung der Umsetzungstreue im Schulkontext dar. Methodische Grenzen bestehen weiterhin bei Wirkungsanalysen von Interventionen, da randomisierte Untersuchungsdesigns nur schwierig zu realisieren sind.

Kaltenbacher, E. (2009). Vom Nutzen der Spracherwerbsforschung für die Sprachvermittlung. Zeitschrift für Literaturwissenschaft und Linguistik, 39 (153), 39-59.

Long, M.H. (2007). Problems in SLA. London: Lawrence Erlbaum Associates.

Mowbray, C. T., Holter, M. C., Teague, G. B. & Bybee D. (2003). Fidelity Criteria: Development, Measurement, and Validation. American Journal of Evaluation, 24 (3), 315-340.

O'Neill, D.K., Pearce, M.J. & Pick, J.L. (2004). Predictive relations between aspects of preschool children's narratives and performance on the Peabody Individualized Achievement Test – Revised: Evidence of a relation between early narrative and later mathematical ability. First Language, 24, 149-183.

Pienemann, M. (1989). Is Language Teachable? Psycholinguistic Experiments and Hypotheses. Applied Linguistics, 1 (10), 52-79.

Schwippert, K., Wendt, H., & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Hrsg.), IGLU 2011.

Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S.191-207). Münster: Waxmann.

Vygodsky, L. S. (1986). Thought and Language. Cambridge: MIT Press.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G147



## ***Vorschulische familiäre sprachbezogene Förderpraktiken und deren Effekte auf schriftsprachliche Vorläuferfertigkeiten bei Kindern mit und ohne Migrationshintergrund***

Der Schriftspracherwerb ist eine komplexe Aufgabe, die ein Kind bereits früh zu meistern beginnt. Dieser Prozess wird durch Interaktionen des Kindes mit seinen Sozialisationsumwelten, wie bspw. der Familie, unterstützt. Aus der bisherigen Forschung ist bekannt, dass die frühe familiäre Literacy-Förderung eine wichtige Rolle für die Schriftsprachentwicklung spielt (Bus, van IJendoorn & Pellegrini, 1995; de Jong & Leseman, 2001; Hood, Conlon & Andrews, 2008; Scarborough & Dobrich, 1994; Sénéchal & LeFevre, 2002). Als zentrale Aktivitäten in diesem Zusammenhang werden alle prä- und paraliterarischen Aktivitäten gesehen (Hurrelmann, Hammer & Nieß, 1993), die im Folgenden als familiäre sprachbezogene Förderpraktiken bezeichnet werden. Hierzu zählen unter anderem das Vorlesen, das Singen von Liedern, das Aufsagen von (Kinder-)Reimen und Gedichten, das Spielen von Sprachspielen und das Erzählen von Geschichten. Bisher sind nur wenige Studien bekannt, in denen die vorschulische familiäre Förderung und ihre Effekte auf schriftsprachliche Kompetenzen bei Kindern mit und ohne Migrationshintergrund untersucht wurden. Im vorliegenden Beitrag sollen daher die folgenden Forschungsfragen beantwortet werden:

- 1) Wie sieht die vorschulische familiäre Fördersituation in Familien mit und ohne Migrationshintergrund aus?
- 2) Wie wirken sich die vorschulischen familialen sprachbezogenen Förderpraktiken auf schriftsprachliche Vorläuferfertigkeiten bei Kindern mit und ohne Migrationshintergrund aus?

Untersucht wurde dies an einer Stichprobe von N = 275 Vorschulkindern, wovon 80 Kinder einen Migrationshintergrund aufweisen (d.h. mind. ein Elternteil spricht eine andere Muttersprache als Deutsch). Im letzten Kindergartenjahr durchliefen 224 Kinder ein Training zur Förderung der phonologischen Bewusstheit, während die restlichen 51 Kinder einer untrainierten Kontrollgruppe angehörten. Die Kompetenzen der Kinder in den Bereichen phonologische Bewusstheit, Wortschatz, Buchstabenkenntnisse und Grammatik wurden zu Beginn und am Ende des letzten Kindergartenjahres, d.h. direkt vor und direkt nach der Trainingsphase, erhoben. Zudem wurden die Eltern im Rahmen einer Elternbefragung um Angaben zur Häufigkeit, mit der sie elf verschiedene sprachbezogene Förderpraktiken vor der Einschulung des Kindes umgesetzt haben, gebeten. Diese Förderpraktiken umfassten unter anderem die folgenden Tätigkeiten: Bücher lesen, Geschichten erzählen, Lieder singen, mit Alphabet-Spielzeug spielen und Wortspiele spielen. Des Weiteren wurden im Elternfragebogen auch Angaben zum sozioökonomischen Status (SES), dem Buchbesitz im Haushalt und der Sprache, die für die Förderpraktiken überwiegend verwendet wurde, erhoben.

Erste Ergebnisse aus Kovarianzanalysen deuten darauf hin, dass es zwischen Kindern mit und ohne Migrationshintergrund auch unter Kontrolle des SES signifikante Unterschiede hinsichtlich der Häufigkeit bestimmter Förderpraktiken gibt. So lesen Eltern ohne Migrationshintergrund ihrem Kind signifikant häufiger vor und betrachten häufiger Bücher gemeinsam mit ihrem Kind. Dies geht auch mit dem Ergebnis einher, dass in Familien ohne Migrationshintergrund eine signifikant höhere Anzahl an Büchern und Kinderbüchern vorliegt. Eltern mit Migrationshintergrund hingegen spielen mit ihrem Kind signifikant häufiger mit Alphabet-Spielzeug. Hinsichtlich der Effekte der sprachbezogenen Förderpraktiken auf die Vorläuferfertigkeiten zu t2 scheint es laut ersten Ergebnissen aus Regressionsanalysen Unterschiede zwischen Kindern mit und ohne Migrationshintergrund zu geben. So wirken sich in beiden Gruppen unterschiedliche Förderpraktiken positiv auf die gemessenen Vorläuferfertigkeiten aus. Bei Kindern mit Migrationshintergrund zeigen die einzelnen Förderpraktiken zudem eher positive Effekte auf verschiedene Vorläuferfertigkeiten als dies bei Kindern ohne Migrationshintergrund der Fall ist. Die Ergebnisse sollen vor dem Hintergrund aktueller Forschungsergebnisse sowie der Rahmenbedingungen der eigenen Studie diskutiert werden.

Bus, A. G., van IJendoorn, M. H. & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65(1), 1-21.

de Jong, P. F. & Leseman, P. P. M. (2001). Lasting effects of home literacy on reading achievement in school. *Journal of School Psychology*, 39(5), 389-414.

Hood, M., Conlon, E. & Andrews, G. (2008). Preschool Home Literacy Practices and Children's Literacy Development: A Longitudinal Analysis. *Journal of Educational Psychology*, 100(2), 252-271.

Hurrelmann, B., Hammer, M. & Nieß, F. (1993). Lesesozialisation Band 1: Leseklima in der Familie: Eine Studie der Bertelsmann Stiftung. Gütersloh: Bertelsmann Stiftung.

Scarborough, H. S. & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, 14(3), 245-302.

Sénéchal, M. & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73(2), 445-460.

## Evaluation des Bundesprogramms „Offensive Frühe Chancen: Schwerpunkt-Kitas Sprache und Integration“

Theoretischer Hintergrund: Die Nachteile von Kindern aus bildungsfernen Familien und Kindern mit Migrationshintergrund bestehen in Deutschland bereits beim Eintritt in die Grundschule (Dubowy et al., 2008; Weinert et al., 2010) und sind vorrangig auf mangelnde Sprachkenntnisse zurückzuführen. Vor diesem Hintergrund fördert das BMFSFJ im Rahmen des Programms „Offensive Frühe Chancen: Schwerpunkt-Kitas Sprache und Integration“ (OFC) Kindertageseinrichtungen, um sie bei der Aufgabe der alltagsintegrierten sprachlichen Bildung zu unterstützen. In der internationalen Forschung zur frühpädagogischen Qualität wird davon ausgegangen, dass Aspekte der Strukturqualität und Orientierungsqualität die pädagogische Prozessqualität beeinflussen, welche Einfluss auf die kindliche Entwicklung hat (Kuger & Kluczniok, 2008; Kluczniok et al., 2011). Das Programm zielt deshalb darauf ab, die sprachpädagogische Arbeit in den Einrichtungen über Verbesserungen aller Qualitätsbereiche in Gang zu setzen. Eine entscheidende Rolle spielen hierbei auch die Orientierungen der pädagogischen Fachkräfte. Im Rahmen der OFC sollen Orientierungs- und Prozessqualität durch die Verfügbarmachung einer zusätzlichen, auf sprachpädagogische Arbeit spezialisierten Fachkraft verbessert werden.

Fragestellung: Die wissenschaftliche Evaluation untersucht die Auswirkungen des Bundesprogramms auf verschiedenen Ebenen. Effekte sind theoretisch sowohl beim Team der Einrichtung und der zusätzlichen Fachkraft als auch bei den Eltern und Kindern zu erwarten. In der aktuellen Studie wird untersucht (1) wie das Bundesprogramm in den Einrichtungen implementiert wurde und welche Umsetzungsformen sich abzeichnen, (2) welche pädagogischen Einstellungen und Überzeugungen der Fachkräfte mit den verschiedenen Umsetzungsformen der Initiative verbunden sind, (3) wie verschiedene Realisierungsformen des Bundesprogramms mit der Prozessqualität in den Kindertageseinrichtungen in Verbindung stehen und (4) ob die Prozessqualität in den Kindertageseinrichtungen mit dem Stand der kindlichen Sprachentwicklung im Zusammenhang steht.

Im Rahmen der Evaluationsstudie werden folgende Stichprobengruppen miteinander verglichen:

- zwei Gruppen von Einrichtungen, die am Programm teilnehmen und jeweils eine unterschiedliche Form zusätzlicher fachlicher Unterstützung erhalten,
- eine Gruppe von Einrichtungen, die ohne zusätzliche fachliche Unterstützung am Programm beteiligt ist sowie
- eine Vergleichsgruppe von Einrichtungen, die nicht am Bundesprogramm teilnimmt.

Insgesamt beteiligen sich 335 Einrichtungen und 1331 Kinder/Familien an der Evaluation.

Methode: Die Evaluation ist als quasi-experimentelle Längsschnittstudie mit drei Messzeitpunkten angelegt und verfolgt einen multimethodischen Ansatz. Die vorliegenden Ergebnisse basieren auf Daten des ersten Messzeitpunktes, die durch Online-Befragungen der Einrichtungsleitungen und der geförderten Zusatzkräfte, Qualitätsbeobachtungen und Interviews in den Einrichtungen sowie standardisierte Entwicklungstests zum kindlichen Sprachstand erhoben wurden. Zur Feststellung der sprachförderlichen Prozessqualität wurden Teile der KES-E (Roßbach & Tietze, in Vorbereitung), KES-R und KRIPS-R (Tietze et al., 2007a, 2007b), der Dortmunder Ratingskala zur Erfassung sprachförderrelevanter Interaktion (Do-Resi) (Fried & Briedigkeit, 2008) sowie eine deutsche Version der TILRS (Girolametto et al., 2000) eingesetzt. Als Instrumente zur Ermittlung des Sprachentwicklungsstandes fanden der Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), Subtests aus dem SETK-2 und dem SETK 3-5 (Grimm et al., 2000, 2001) sowie aus dem Test zum Satzverständnis von Kindern (Sieg Müller et al., 2011) Verwendung. Für die Beantwortung der Fragestellungen werden deskriptive Auswertungsformen und regressionsanalytische Verfahren genutzt.

Ergebnisse: Die Ergebnisse des ersten Messzeitpunktes verweisen auf eine hohe Varianz hinsichtlich der Umsetzung des Bundesprogramms in den Kindertageseinrichtungen. Seine Implementierung wurde von den Einrichtungen als positiv wahrgenommen. Es zeigt sich eine deutliche Präferenz von Tätigkeiten aus den Schwerpunktbereichen „direkte Arbeit mit Kindern“ (58,4 %) und „Arbeit mit dem Team“ (27,3 %) im Vergleich zum Bereich „Arbeit mit Familien“ (14,3 %). Bisherige Ergebnisse bestätigen den Zusammenhang zwischen pädagogischer Orientierungsqualität und Prozessqualität. Unter anderem zeigt sich ein positiver Zusammenhang (.14\*) zwischen der sprachpädagogischen Arbeit während der Freispielzeit und den KES-E Subskalen Lesen und Diversität. Aufbauend auf diesen Ergebnissen werden Daten zum kindlichen Sprachstand ausgewertet und in Verbindung zu Orientierungs- und Prozessqualität gesetzt.

Dubowy, M., Ebert, S., von Maurice, J., & Weinert, S. (2008). Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten: Ein Vergleich von Kindern mit und ohne Migrationshintergrund. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 40, 124-134.

Dunn, L. M., & Dunn, L. M. (2007). Peabody Picture Vocabulary Test. Fourth Edition. London, UK: Pearson.

Fried, L., & Briedigkeit, E. (2008). Sprachförderkompetenz – Selbst- und Teamqualifizierung für Erzieherinnen, Fachberatungen und Ausbilder. Berlin, Düsseldorf, Mannheim: Cornelsen Verlag Scriptor.

Girolametto, L., Weitzman, E., & Greenberg, J. (2000). Teacher Interaction and Language Rating Scale. Toronto, ON: The Hanen Centre.

Grimm, H., Aktas, M., & Frevert, S. (2000). SETK-2: Sprachentwicklungstest für zweijährige Kinder. Diagnose rezeptiver und produktiver Sprachverarbeitungsfähigkeiten. Göttingen: Hogrefe.

Grimm, H., Aktas, M., & Frevert, S. (2001). SETK 3-5. Sprachentwicklungstest für drei- bis fünfjährige Kinder. Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen. Göttingen: Hogrefe.

Kluczniok, K., Anders, Y., & Ebert, S. (2011). Förderereinstellungen von Erzieherinnen. Einflüsse auf die Gestaltung von Lerngelegenheiten im Kindergarten und die kindliche Entwicklung früher numerischer Kompetenzen. Frühe Bildung, 0(1), 13-21.

Kuger, S., & Kluczniok, K. (2008). Prozessqualität im Kindergarten – Konzept, Umsetzung und Befunde. Zeitschrift für Erziehungswissenschaft, Sonderheft 11, 159-178.

Sieg Müller, J., Kauschke, C., von Minnen, S., & Bittner, D. (2011). TSVK: Test zum Satzverstehen von Kindern. München: Elsevier, Urban & Fischer Verlag.

Tietze, W., Bolz, M., Grenner, K., Schlecht, D., & Wellner, B. (2007a). Krippen-Skala. Revidierte Fassung (KRIPS-R). Berlin: Cornelsen Verlag Scriptor.

Tietze, W., Schuster, K.-M., Grenner, K., & Roßbach, H.-G. (2007b). Kindergarten-Skala. Revidierte Fassung (KES-R). Berlin: Cornelsen Verlag Scriptor.

Roßbach, H.-G., & Tietze, W. (in Vorbereitung). Kindergarten-Skala Erweiterung (KES-E). Deutsche Fassung der „The Early Childhood Environment Rating Scale-Extension (ECERS-E)“ von Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2006).

Weinert, S., Ebert, S., & Dubowy, M. (2010). Kompetenzen und soziale Disparitäten im Vorschulalter. Zeitschrift für Grundschulforschung, 1, 32-45.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G147



## **Individuelle und instruktionale Rückmeldungen im Zusammenhang mit Lesekompetenzen von Kindern im Grundschulunterricht – Ergebnisse aus einer empirischen Studie**

Individuelle Rückmeldungen stellen wichtige Bedingungen für erfolgreiche Kompetenzerwerbsprozesse bei Kindern im Grundschulalter dar. Aufgrund der hohen Heterogenität an Lernausgangslagen (vgl. Bos, Bremerich-Vos, Tarelli & Valtin, 2012; Einsiedler, Martschinke & Kammermeyer, 2008) kommt der individuellen Unterstützung der Lernprozesse von Schülerinnen und Schülern im Grundschulunterricht eine hohe Bedeutung zu. Dennoch zeigen Befunde aus den Internationalen Grundschul-Lese-Untersuchungen (vgl. Lankes & Carstensen, 2007; Tarelli, Lankes, Drossel & Gegenfurtner, 2012), dass etwa die Hälfte aller Kinder im Leseunterricht in der Grundschule in Deutschland im Klassenverband ohne individuelle Unterstützung unterrichtet wird: Leseprozesse von Schülerinnen und Schülern werden durch die Grundschullehrkräfte nur wenig strukturiert, individuelle Rückmeldungen zu Leseprozessen nur selten erteilt.

Empirische Befunde geben Hinweise darauf, dass dem Unterstützungsverhalten von Lehrkräften im Sinne von Strukturierungen, Hilfestellungen und Rückmeldungen eine besondere Rolle hinsichtlich der Kompetenzerwerbsprozesse von Schülerinnen und Schülern zukommt. Dabei zeigt sich, dass durch die Qualität von Rückmeldungen oder die Präferenz für verschiedene Feedbacktypen seitens der Lehrkräfte Unterschiede in fachlichen und überfachlichen Kompetenzen von Schülerinnen und Schülern aufgeklärt werden können. So konnte in diesem Zusammenhang beispielsweise Law (2011) im Rahmen einer Querschnitts-studie mit Zweitklässlern sowie deren Lehrkräften aufzeigen, dass sich sowohl die durch die Schülerinnen und Schüler wahrgenommenen sowie die durch die Lehrkraft erteilten instruktionalen Rückmeldungen als förderlich hinsichtlich der Motivation und der Lesekompetenz der Schülerinnen und Schüler erwiesen. Weiterhin konnten in Studien vielfach förderliche Wirkungen der Betonung einer individuellen Bezugsnormorientierung durch die Lehrkraft im Unterricht hinsichtlich der Entwicklung der akademischen Selbstkonzepte und der Motivationen der Schülerinnen und Schüler herausgestellt werden (vgl. z.B. Lüdtke & Köller, 2002; Lüdtke, Köller, Marsh & Trautwein, 2005; Henderlong Corpus & Lepper, 2007). Im Detail zeigte sich in diesen Studien, dass ein effizientes Rückmeldeverhalten eine der zentralen Determinanten für Schulleistungen darstellt (vgl. Hattie, 2009; Hattie & Timperley, 2007). Weitgehend ungeklärt ist in diesem Zusammenhang allerdings, welche Rolle der Art und der Häufigkeit der Erteilung von unterschiedlichen Rückmeldungen im Leseunterricht der Grundschule durch die Lehrkraft zukommt, wie diese Rückmeldungen von den Schülerinnen und Schülern wahrgenommen werden und welche Auswirkungen diese hinsichtlich ihrer Fähigkeitsselbstkonzepte sowie ihrer Kompetenzerwerbsprozesse im Lesen haben. Diese Forschungslücke wird im Rahmen des vorliegenden Projekts anhand des Leseunterrichts an Grundschulen bearbeitet:

An der vorliegenden Studie sind insgesamt N=693 Schülerinnen und Schüler aus dem dritten und vierten Grundschuljahr sowie ihre Grundschullehrerinnen und -lehrer (N=34) beteiligt gewesen. Die Kinder wurden zum einen zu ihren lesebezogenen Selbstkonzepten und ihrer Wahrnehmung der Intensität der erteilten instruktionalen Rückmeldungen befragt, zum anderen bearbeiteten sie Aufgaben aus dem Leseverständnistest ELFE („Ein Leseverständnistest für Erst- bis Sechstklässler“; vgl. Schneider & Lenhard, 2006). Die an der Studie beteiligten Lehrkräfte wurden mittels eines Fragebogens zu der Häufigkeit des Einsatzes individuellen Feedbacks sowie instruktionalen Feedbacks im Leseunterricht der Grundschule befragt.

Die Ergebnisse verdeutlichen, dass Schülerinnen und Schüler, deren Lehrkräfte häufig individuelles Feedback erteilen, über signifikant bessere Lesefähigkeiten verfügen, als Kinder, die selten individuelle Rückmeldungen im Unterricht erhalten. Als erwartungswidrig ist dagegen der Befund einzustufen, dass Unterschiede zwischen diesen Kindergruppen in Bezug auf ihre fähigkeitsbezogenen Selbstkonzepte nicht festzustellen sind. Weiterhin zeigt sich anhand der ermittelten Befunde theoriekonform, dass Schülerinnen und Schüler, die häufig instruktionales Feedback durch die Lehrerinnen und Lehrer erhalten, eine signifikant stärkere Erteilung dieser Rückmeldungen im Leseunterricht wahrnehmen und signifikant bessere Lesefertigkeiten aufweisen, als Kinder, die eher selten instruktionale Rückmeldungen von ihren Lehrerinnen und Lehrern erhalten. Dagegen konnten – entgegen der Erwartungen – Unterschiede im Leseverständnis von Grundschulkindern und -schülern nicht durch die von den Kindern wahrgenommenen instruktionalen Rückmeldungen erklärt werden. Die vorliegenden Befunde geben Hinweise darauf, dass sowohl der fachliche als auch der überfachliche Kompetenzerwerb von Schülerinnen und Schülern durch adäquate Unterstützungsmaßnahmen der Lehrerinnen und Lehrer in gewisser Weise beeinflusst wird.

- Bos, W., Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 91-136). Münster: Waxmann.
- Einsiedler, W., Martschinke, S. & Kammermeyer, G. (2008). Die Grundschule zwischen Heterogenität und gemeinsamer Bildung. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), Das Bildungswesen in der Bundesrepublik Deutschland (S. 325-374). Reinbek: Rowohlt.
- Hattie, J. (2009). Visible Learning: A synthesis of over 800 meta-analyses relating to achievement. London: Routledge, Taylor & Francis.
- Hattie, J. & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81-112.
- Henderlong Corpus, J. & Lepper, M. R. (2007). The effects of person versus performance praise on children's motivation: Gender and age as moderating factors. Educational Psychology, 27(4), 487-508.
- Lankes, E.-M. & Carstensen, C. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 161-193). Münster: Waxmann.
- Law, Y.-K. (2011). The role of teachers' cognitive support in motivation young Hong Kong Chinese children to read and enhancing reading comprehension. Teaching and Teacher Education, 27, 73-84.
- Lüdtke, O. & Köller, O. (2002). Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht. Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 34(3), 156-166.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. Contemporary Educational Psychology, 30, 263-285.
- Schneider, W. & Lenhard, W. (2006). ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.
- Tarelli, I., Lankes, E.-M., Drossel, K. & Gegenfurtner, A. (2012). Lehr-Lernbedingungen an Grundschulen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 137-173). Münster: Waxmann.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G147





## ***Verschiedene Perspektiven auf Unterricht: Wie nehmen Lehrer/innen und Schüler/innen ihren gemeinsamen Unterricht wahr?***

**Chair(s):** Sabine Gruehn (Westfälische Wilhelms-Universität Münster)

In der empirischen Unterrichtsforschung wird die Qualität von Unterricht in der Regel anhand von Lehrerangaben, Schülerangaben oder Beobachtungsdaten erfasst. Jede dieser drei Perspektiven zeichnet sich sowohl durch ihre inhaltlichen als auch ihre methodischen Spezifikationen aus. Inhaltlich werden durch die verschiedenen Perspektiven z. T. unterschiedliche Aspekte des Unterrichts abgebildet (vgl. Clausen, 2002) und präsentieren erst zusammengenommen ein facettenreiches Bild der Unterrichtsqualität. Methodisch hat jede Perspektive Stärken und Schwächen sowie blinde Flecken, die nicht zuletzt darin begründet liegen, dass Schüler- und Beobachterangaben Fremd-, Lehrerangaben hingegen Selbstbeurteilungen darstellen. Es lassen sich dennoch differentielle Übereinstimmungen für einzelne Unterrichtsmerkmale zeigen, bspw. bestehen eher Parallelen von Lehrer- und Schülerurteilen bei Unterrichtsmerkmalen, die eine längere gemeinsame Erfahrung unterrichtlicher Interaktion und Kommunikation voraussetzen (Clausen, 2002). Für die Unterrichtsentwicklung können die verschiedenen Perspektiven dazu genutzt werden, Verbesserungsmöglichkeiten zu identifizieren. Insbesondere geben Abweichungen zwischen der Lehrer- und Schülerperspektive Hinweise darauf, inwieweit die durch methodisch-didaktische Überlegungen der Lehrkraft vorstrukturierten Unterrichtsarrangements von den Schülern wahrgenommen und entsprechend genutzt werden (können). In bisherigen Studien, die die Unterrichtsqualität mehrperspektivisch erfasst haben (vgl. Clausen, 2002; Helmke, Schneider, & Weinert, 1986; Helmke, 2012), wurden vor allem Unterrichtsbeurteilungen von Schülern der Mittelstufe (Jahrgangsstufe 6 bzw. 8) genutzt, die – eine Schwäche dieser Perspektive – einen vergleichsweise geringen Differenzierungsgrad aufweisen. Darüber hinaus wurden die Unterrichtseinschätzungen ausschließlich im Fach Mathematik erhoben. Es ist deshalb eine offene Frage, inwieweit a) Rückmeldungen von älteren Schülern, etwa aus der Oberstufe, auf differenzierteren Unterrichtseinschätzungen basieren, die ggf. engere Bezüge zu den Selbsteinschätzungen von Lehrkräften aufweisen, und b) sich diese Sachverhalte in anderen Fächern wiederfinden lassen.

Für die hier zu präsentierende Studie ergeben sich damit folgende Forschungsfragen:

- Wie nehmen Oberstufenschüler/innen und Lehrer/innen gleiche Unterrichtsmerkmale des Unterrichts wahr, wo bestehen Gemeinsamkeiten bzw. Unterschiede?
- Wie differenziert sind die Schülerurteile über den Unterricht?
- Bestehen Unterschiede in den Einschätzungen zwischen verschiedenen Fächern?

Um diese Fragen zu untersuchen, wurden in einem von der DFG geförderten Forschungsprojekt (vgl. Rahn, Gruehn, & Böttcher, 2012) im Juni/Juli 2013 an 49 Gesamtschulen und beruflichen Gymnasien in ganz NRW Schüler/innen (N=2212) der Jahrgangsstufe 11 sowie 168 ihrer Lehrkräfte zu deren gemeinsamen Unterricht befragt. Die Befragung erfolgte mittels standardisierter Fragebögen in den Fächern Deutsch, Mathematik und den als Schwerpunkt gewählten profilbildenden Fächern bei Schüler/innen an beruflichen Gymnasien. Die 39 ausgewählten beruflichen Gymnasien (N=1532) weisen zu 14 % einen erziehungswissenschaftlichen, zu 12 % einen technischen und zu 40 % einen kaufmännischen Schwerpunkt auf. Folgende Unterrichtsmerkmale wurden sowohl über die Lehrkräfte als auch die Schüler/innen mit Hilfe etablierter Skalen aus der Unterrichtsforschung erfasst: Klassenführung, Differenzierung, Elemente offenen Unterrichts sowie kooperative Unterrichtsformen. Darüber hinaus sollten die Schüler/innen ein Gesamturteil über ihre Zufriedenheit mit der Lehrkraft sowie mit dem Unterricht abgeben. Die Lehrkräfte wurden gebeten einzuschätzen, wie hoch die Zufriedenheit der Klasse mit ihrem Unterricht bzw. mit ihnen als Lehrperson ist. Die faktorenanalytische Überprüfung der Differenziertheit der Schülerurteile zeigt eine prägnantere Ausprägung der verschiedenen Unterrichtsmerkmale im Vergleich zu Studien mit jüngeren Schülern, in denen die gleichen Skalen eingesetzt wurden (z. B. Klieme, Pauli, & Reusser, 2005). Die Korrelationsanalysen weisen darauf hin, dass Unterrichtsmerkmale aus verschiedenen Perspektiven unterschiedlich wahrgenommen werden, dementsprechend zeigen die Merkmale Klassenführung und Differenzierung perspektivenspezifische Wahrnehmungstendenzen. Während die Lehrkräfte die ‚Differenzierung‘ innerhalb ihres Unterrichts höher einzuschätzen scheinen, zeigen die Schüler/innen tendenziell eine positivere Einschätzung des Klassenführungsmerkmals ‚Disziplin‘. Hinsichtlich verschiedener Unterrichtsfächer deuten sich zwischen der Wahrnehmung des Deutsch- und des Mathematikunterrichts Unterschiede in den Beurteilungen durch die Schüler/innen an. Weiterhin wird geprüft, inwieweit diese Wahrnehmungsunterschiede in Zusammenhang mit den Beurteilungen anderer Unterrichtsmerkmale stehen und ob es unter den Lehrkräften Differenzen zwischen Fächern gibt. Die Ergebnisse werden vor dem Hintergrund ihrer Bedeutung für die Unterrichtsentwicklung diskutiert.

Clausen, M. (2002). Unterrichtsqualität - eine Frage der Perspektive? Münster: Waxmann.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze: Klett-Kallmeyer.

Helmke, A., Schneider, W. & Weinert, F.-E. (1986) Quality of instruction and classroom learning outcomes: The German contribution to the IEA classroom environment study. Teaching and Teacher Education, 2(1), 1-18.

Klieme, E., Pauli, C. & Reusser, K. (2005). Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie Unterrichtsqualität, Lernverhalten und mathematisches Verständnis". Frankfurt a.M.: GPPF.

Rahn, S., Gruehn, S. & Böttcher, W. (2012). Determinanten des Schülerfeedbacks: eine fächerver-gleichende Analyse von Schüleraussagen zur Unterrichtsqualität an allgemeinbildenden und beruflichen Gymnasien . DFG-Antrag auf Sachbeihilfe. Münster.





## **Die Erfassung des Unterrichts aus Schülersicht: Erste Ergebnisse einer Systematisierung nationaler Fragebogeninstrumente**

Die Messung verschiedener Qualitätsdimensionen des Unterrichts gehört fraglos zu den Schwerpunkten der empirischen Bildungsforschung. Hierbei ist die Schülerbefragung anhand von Fragebögen eine der häufigsten Methoden zur Datengewinnung. Diese sind seit einiger Zeit fester Bestandteil empirischer Studien und auch im Bereich der Schulentwicklung sowie des Bildungsmonitorings ist der Einsatz von Schülerbefragungen unstrittig (Clausen, 2002; Helmke, 2006; Pietsch, 2009). Schülerinnen und Schüler können als Experten hinsichtlich der Wirkung pädagogischer Prozesse angesehen werden und liefern im Vergleich zu alternativen Verfahren (z.B. Selbstberichte der Lehrkräfte oder Expertenurteile) deutlich einfacher zu erhebende Daten (Clausen, 2002; Helmke, 2006). Allerdings ist der Großteil von Fragebogeninstrumenten zur Erfassung von Unterricht aus Schülersicht durch eine überaus große Heterogenität bestimmt. Bereits eine oberflächliche Betrachtung verwendeter Fragebögen zeigt neben Unterschieden in den zu erfassenden Qualitätsdimensionen vor allem Unterschiede bzgl. verschiedener Itemmerkmale. So unterscheiden sich Items hinsichtlich ihrer verhaltensbezogene Spezifität (Inferenzgrad), der Wahrnehmungsperspektive (Adressatenbezug) als auch des zu beurteilenden Unterrichtszeitraums (Beurteilerzeitraum). Dabei ist zu vermuten, dass diese von den zu erfassenden Qualitätsdimensionen zunächst unabhängigen Itemmerkmale auch einen Einfluss auf die psychometrischen Eigenschaften der verwendeten Instrumente ausüben können. Beispielsweise konnten Cohen (1981) und Feldman (1989) zeigen, dass hoch-inferente Schülereinschätzungen (d.h. Urteile, die interpretative Prozesse seitens des Beurteilers erfordern) zwar valide im Sinne signifikanter und erwartungskonformer Korrelationen mit Schulleistungen sind, allerdings auch eine deutlich niedrigere faktorielle Validität als niedrig-inferente Ratings (d.h. Urteile über ein konkret beobachtbares Unterrichtsmerkmal) aufweisen (Murray, 2007). Darüber hinaus wurden in Abhängigkeit von der gruppen- bzw. individuumsbezogenen Itemformulierung Mittelwertunterschiede und spezifische Zusammenhänge mit outcomes berichtet (McRobbie, Fisher & Wong, 1998; Baltes, Zhandova, & Parker, 2009). Ziel des vorliegenden Beitrags ist es daher, die Ergebnisse einer umfangreichen Systematisierung bekannter Fragebogeninstrumente zur Erfassung von Unterricht aus Schülersicht zu berichten. Hierzu wurden unter anderem auf Grundlage der Datenbank DaQS (Deutsches Institut für Internationale Pädagogische Forschung, 2013) Fragebögen und Skalenhandbücher einschlägiger Studien zur Unterrichtsqualität gesichtet und dokumentiert. Insgesamt konnten bislang 533 Items (120 Unterrichtsskalen) aus acht Studien entlang ihrer Itemformulierung und der statistischen Kennzahlen beschrieben werden. Deskriptive Ergebnisse wurden ggf. durch Re-analysen ergänzt. Die Ergebnisse einer ersten Untersuchung bestätigten dabei die Bedeutung des Adressatenbezugs für die Variabilität der Schülereinschätzungen auf der Klassenebene. Im Rahmen eines Mehrebenenmodells zur Vorhersage von Intraklassenkorrelationskoeffizienten (ICCs) auf Itemebene zeigten sich auch unter Berücksichtigung der zugrundeliegenden Unterrichtsdimensionen und des zu beurteilenden Unterrichtsfaches höhere ICCs für Items mit „Wir-Bezug“ als für Items mit „Ich-Bezug“. Inwieweit sich hierin auch eine höhere Validität im Hinblick auf die Erfassung der Unterrichtsqualität widerspiegelt, ist Gegenstand aktueller Analysen.

- Baltes, B. B., Zhandova, L. S. & Parker, C. P. (2009). Psychological climate: A comparison of organizational and individual level referents. *Human Relations*, 62, 669-700.
- Clausen, M. (2002). Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of findings. *Research in Higher Education*, 51, 281-309.
- Deutsches Institut für Internationale Pädagogische Forschung (2013). Datenbank zur Qualität von Schule. Zugriff am 31.10.2013 <<http://daqs.fachportal-paedagogik.de/>>.
- Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583-645.
- Helmke, A. (2006). Unterrichtsqualität: Erfassen, Bewerten, Verbessern. Seelze: Kallmeyersche Verlagsbuchhandlung.
- McRobbie, C. J., Fisher, D. L. & Wong, A. F. L. (1998). Personal and class forms of classroom environment instruments. In B. J. Fraser & K. G. Tobin (Hrsg.), *International Handbook of Science Education (Part One)* (S. 581-594). Boston: Kluwer Academic Publishers.
- Murray, H. G. (2007). Low-inference teaching behaviors and college teaching effectiveness: Recent developments and controversies. In R. P. Perry & J. C. Smart (Hrsg.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. New York: Springer.
- Pietsch, M. (2009). Die Qualität des Unterrichts an Hamburger Schulen aus Beobachterperspektive. In Institut-für-Bildungsmonitoring (Hrsg.), *Jahresbericht der Schulinspektion Hamburg 2008* (S. 40-58). Hamburg: Institut für Bildungsmonitoring.



## ***Fachübergreifende Merkmale und Muster eines kognitiv aktivierenden Unterrichts an Gymnasien – Ergebnisse aus video- und in vivo-basierten Einzelfallstudien in heterogenen und homogenen gymnasialen Klassen***

Das Konstrukt Kognitive Aktivierung ist neben den klassischen Aspekten der Unterrichtsqualität, ein zentrales Konzept in der empirischen Unterrichtsforschung. Dennoch ist es in seiner theoretischen Definition nach wie vor unscharf konturiert und bleibt aufgrund noch unzureichender empirischer Wirksamkeitsbelege ein in Teilen normativer Referenzrahmen für den zeitgemäßen Unterricht. Auch ist unklar, inwieweit die Kognitive Aktivierung als Förderung von aktiven Konstruktionsleistungen der Lernenden mit ihren unterschiedlichen Lernvoraussetzungen und in den verschiedenen Lernumwelten in gleicher Weise von Bedeutung ist (Bohl et al., 2012). Charakteristisch für einen kognitiv aktivierenden Unterricht ist die Synthese aus der Nutzung offener, problemorientierter und lebensweltnaher Aufgaben, aktiven Schülerarbeitsphasen sowie einer differenzierten Lehrerrolle (Kleickmann, 2012). Ziel ist die Förderung und der Erwerb vernetzten und transferfähigen Wissens. PISA-Befunde zum Unterricht belegen, dass derartige Unterrichtsformen häufiger an leistungsstarken Gymnasien vorgefunden werden (Baumert et al., 2004).

Ausgehend von dieser Befundlage wurde für die zu präsentierende Studie ein Forschungsansatz gewählt, der die beiden Disziplinen Allgemeine Didaktik und empirische Unterrichtsforschung miteinander verbindet. Dabei stellt das didaktische Dreieck nach Reusser (2011) einen Bezugsrahmen für die theoretische Aufarbeitung und die Operationalisierung des Konstrukts sowie die Entwicklung eines Raterinventars dar. Das Raterinventar bildete die Grundlage von explorativ angelegten Beobachtungen an Gymnasien. Ziel war es, neben fachübergreifenden Merkmalen eines kognitiv aktivierenden Unterrichts auch klassische Unterrichtsqualitätsmerkmale an Gymnasien zu erfassen. Die Stichprobe setzt sich aus einem Gymnasium für leistungsstarke und hochbegabte Schülerinnen und Schüler mit einer tendenziell eher homogenen Schülerschaft und aus zwei weiteren Gymnasien mit einer tendenziell eher heterogenen Schülerschaft zusammen. Das Forschungsdesign umfasst sowohl quantitative Erhebungen in Form von video- und in vivo-basierten Unterrichtsbeobachtungen durch externe Rater sowie Schüler- und Lehrerbefragungen, als auch qualitative Erhebungen in Form von Schüler- und Lehrerinterviews.

Vorgelegt werden die deskriptiven Befunde aus den Unterrichtsbeobachtungen (N = 38). Neben den häufig vorkommenden Merkmalen eines kognitiv aktivierenden Unterrichts an Gymnasien, werden die Unterschiede zwischen der homogenen und der heterogenen gymnasialen Schülerschaft in Bezug auf klassische Unterrichtsqualitätsmerkmale und kognitiv aktivierende Merkmale berichtet. Es zeigt sich, dass der Unterricht an dem Gymnasium mit der homogenen und leistungsstarken Schülerschaft durch einen hohen diskursiven Unterrichtsstil, ein hohes Ausmaß an Schülerkooperation und eine differenzierte Lehrerrolle geprägt ist, während an Gymnasien mit einer heterogenen Schülerschaft klassische Unterrichtsqualitätsmerkmale, wie die Klarheit und Strukturiertheit des Unterrichts, dominieren. Ein kognitiv aktivierender Unterricht äußert sich in dieser Gruppe durch eine differenzierte Aufgabenstellung, die Förderung von Lern- und Arbeitsstrategien sowie schülerseitige Rückmeldungen über Lernfortschritte oder Lernschwierigkeiten.

Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hrsg.), PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleiches. Münster: Waxmann.

Bohl, T., Kleinknecht, M., Batzel, A. & Richey, P. (2012). Aufgabenkultur in der Schule: Eine vergleichende Analyse von Aufgaben und Lehrerhandeln im Hauptschul-, Realschul- und Gymnasialunterricht. Baltmannsweiler: Schneider Verlag Hohengehren.

Kleickmann, T. (2012). Kognitiv aktivieren und inhaltlich strukturieren im naturwissenschaftlichen Sachunterricht. Handreichungen des Programms SINUS an Grundschulen. Kiel: IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität.

Reusser, K. (2011). Von der Unterrichtsforschung zur Unterrichtsentwicklung – Probleme, Strategien, Werkzeuge. In W. Einsiedler (Hrsg.), Unterrichtsentwicklung und Didaktische Entwicklungsforschung. Bad Heilbrunn: Julius Klinghardt.

**Fächerperspektiven auf gute Lehre**

Akademische Disziplinen unterscheiden sich hinsichtlich ihrer epistemologischen Charakteristika (Muis, Bendixen & Haerle, 2006) und werden entsprechend dieser Charakteristika oft in „harte“, mathematisch-naturwissenschaftliche (MINT), und „weiche“ Disziplinen, Geistes- und Sozialwissenschaften (GESO), unterschieden. MINT-Fächer wie Chemie oder Physik widmen sich häufiger mit quantitativen Methoden der Ableitung von allgemein gültigen Aussagen. Das Wissen in diesen Fachkulturen gilt als kumulativ und disparat strukturiert. GESO-Fächer wie Geschichte oder Politikwissenschaften werden als qualitativ, sich ständig wiederholend und mit Einzelheiten befassend beschrieben. (Neumann, Parry & Becher, 2002) Diese fachspezifischen Eigenschaften beeinflussen neben Forschungsmethoden und Vorstellungen über Wissen und Wissenserwerb auch Lehrensätze bzw. die Vorstellungen darüber, was gute Lehre ausmacht (Päuler & Jucks, in preparation). Die vorliegende Studie thematisiert Lehrorientierungen von Lehrenden und Studierenden unterschiedlicher Fachdisziplinen und adressiert die folgenden beiden Fragestellungen: (1) Wie unterscheiden sich Lehrvorstellungen von Lehrenden und Studierenden der Fachgruppe MINT und der Fachgruppe GESO? (2) Wie unterscheiden sich die Lehrorientierungen zwischen den Statusgruppen?

Hinsichtlich der ersten Fragestellung weisen bisherige Studien deutliche Unterschiede aus: Lehrende weicher Fächer weisen eine größere Studierendenorientierung auf als Lehrende harter Fächer, die ihrerseits stärker traditionelle lehrendenorientierte Lehrvorstellung haben (z.B. Lindblom-Ylänne, Trigwell, Nevgi & Ashwin, 2006; Lübeck, 2010; Lueddeke, 2003). In einer qualitativen Studie wurde gefunden, dass die Lehrpräferenzen von Studierenden sich in ähnlicher Weise zwischen den Fachgruppen unterscheiden (Parpala, Lindblom-Ylänne & Rytönen, 2011).

Die zweite Fragestellung ist relevant, da Kongruenz zwischen den Lehrvorstellungen Lehrender und Studierender als wesentlicher Faktor für effektive Lehr-Lern-Prozesse gilt (Vermunt & Verloop, 1999). Virtanen und Lindblom-Ylänne (2010) fanden jedoch substantielle Unterschiede zwischen den Lehrorientierungen Lehrender und Studierender der Biowissenschaften. Während für Hochschullehrende die Reflexion von Lehrvorstellungen ein wesentlicher Bestandteil ihrer Professionalisierung ist (Amundsen & Wilson, 2012), sind Studierende hier vermutlich weniger reflektiert und tendieren daher möglicherweise stärker zu traditionellen, lehrendenorientierten Vorstellungen. Zur Beantwortung der vorliegenden Fragestellungen wurde sowohl ein qualitativer als auch ein quantitativer Ansatz verfolgt. 160 Lehrende und 309 Studierende aus unterschiedlichen Fachdisziplinen verschiedener deutscher Universitäten bearbeiteten das Approaches to Teaching Inventory (ATI-R; Trigwell, Prosser & Ginns, 2005). Diese quantitativen Daten wurden hinsichtlich fach- und statusgruppen- sowie ausgewählter fachspezifischer Unterschiede analysiert. Darüber hinaus führte ein Teil der Lehrenden und Studierenden den Satzanfang „Für gute Lehre ist es aus meiner Sicht wichtig, dass...“ fort. Diese qualitativen Daten wurden vor dem Hintergrund der Lehrenden- und Studierendenorientierung inhaltsanalytisch (Mayring, 2010) ausgewertet (Päuler & Jucks, 2013). Eine multivariate Varianzanalyse (MANOVA) mit den between-subject Faktoren Fachgruppe (MINT vs. GESO) und Statusgruppe (Lehrende vs. Studierende) sowie den Skalen des ATI (Lehrenden- und Studierendenfokus) als abhängigen Variablen ergab, dass Mitglieder der MINT-Fächer und Studierende einen höheren Lehrendenfokus und geringeren Studierendenfokus zeigten als Mitglieder der GESO bzw. Lehrende. Der Statusgruppenunterschied wurde unter Berücksichtigung der Kovariaten Geschlecht und Alter jedoch nicht mehr signifikant. Eine zweite MANOVA mit einem Subset aus jeweils drei MINT- und GESO-Fächern, für die  $n > 30$  galt, ergab auch zwischen einzelnen Fächern Unterschiede: Mitglieder des Faches Geschichte wiesen im Vergleich zu Mitgliedern der Germanistik, der Mathematik, der Chemie und der Physik einen niedrigeren Lehrendenfokus auf, keine Unterschiede bestanden zu der Vorstellung von Mitgliedern der Politikwissenschaften. Innerhalb des Studierendenfokus wiesen Mitglieder aller drei Geistes- und Sozialwissenschaften signifikant höhere Werte auf als Mitglieder der drei MINT-Fächer. Die qualitativen Daten aus der Fortführung des o.g. Satzanfanges wurden ebenfalls vor dem Hintergrund der Lehrenden- und Studierendenorientierung ausgewertet (Päuler & Jucks, 2013). Studierende wiesen auch hier stärker als Lehrende eine Lehrendenorientierung auf, die Studierendenorientierung war hingegen bei Lehrenden stärker ausgeprägt. Innerhalb der Studierendenstichprobe zeigten sich darüber hinaus Unterschiede zwischen den Fachgruppen MINT (Mathematik, Informatik, Naturwissenschaften, Technik) und GESO (Geistes- und Sozialwissenschaften), die jedoch nicht einheitlich zu interpretieren waren. Die Ergebnisse werden vor dem Hintergrund hochschuldidaktischer Professionalisierung und ihrer Relevanz für die Lehrpraxis diskutiert.

- Amundsen, C. & Wilson, M. (2012). Are We Asking the Right Questions? A Conceptual Review of the Educational Development Literature in Higher Education. *Review of Educational Research*, 82(1), 90-126.
- Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction*, 7(3), 255-275.
- Lindblom-Ylänne, S., Trigwell, K., Nevgi, A. & Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education*, 31(3), 285-298.
- Lübeck, D. (2010). Wird fachspezifisch unterschiedlich gelehrt? Empirische Befunde zu hochschulischen Lehrensätzen in verschiedenen Fachdisziplinen. *Zeitschrift für Hochschulentwicklung*, 5(2), 7-24.
- Lueddeke, G. R. (2003). Professionalising teaching practice in higher education: a study of disciplinary variation and 'teaching-scholarship'. *Studies in Higher Education*, 28(2), 213-228.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse – Grundlagen und Techniken*. Weinheim: BELTZ.
- Muis, K. R., Bendixen, L. D. & Haerle, F. C. (2006). Domain-Generality and Domain-Specificity in Personal Epistemology Research: Philosophical and Empirical Reflections in the Development of a Theoretical Framework. *Educational Psychology Review*, 18(1), 3-54.
- Neumann, R., Parry, S. & Becher, T. (2002). Teaching and learning in their disciplinary context: a conceptual analysis. *Studies in Higher Education*, 27(4), 405-417.
- Parpala, A., Lindblom-Ylänne, S. & Rytönen, H. (2011). Students' conceptions of good teaching in three different disciplines. *Assessment & Evaluation in Higher Education*, 36(5), 549-563.
- Päuler, L. & Jucks, R. (2013). Direkt erfragt: die Messung von Lehrorientierungen per Fragebogen mit offenem Antwortformat. *Zeitschrift für Hochschulentwicklung*, 8(3), 95-109.
- Päuler, L. & Jucks, R. (in Vorbereitung). Perspectives on Teaching: Conceptions of Teaching and Epistemological Beliefs of University Teachers and Students in Different Disciplines.
- Vermunt, J. D. & Verloop, N. (1999). Congruence and friction between learning and teaching. *Learning and Instruction*, 9, 257-280.
- Virtanen, V. & Lindblom-Ylänne, S. (2009). University students' and teachers' conceptions of teaching and learning in the biosciences. *Instructional Science*, 38(4), 355-370.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G150



## **Typische Unterrichtsmuster - Ansätze zur Beschreibung und Erklärung der Variabilität kaufmännischen Unterrichts**

Im Zuge der wirtschaftsdidaktischen Diskurse über Handlungs- und Situationsorientierung der letzten Jahrzehnte und der damit einhergehenden Modifikation curriculärer Vorgaben und der Lehrerbildung ist davon auszugehen, dass dies gleichsam Veränderungen in der Gestaltung von Unterricht herbeiführte. Dahingehend geben die wenigen Untersuchungen der jüngeren Vergangenheit jedoch Anlass zu Skepsis. Pätzold et al. (2003) sowie Seifried, Grill und Wagner (2006) kritisieren Methodenmonotonismus oder Lehrerdominanz und resümieren, dass in der Unterrichtsrealität keine grundlegenden Veränderungen nachgewiesen werden können. Über die Wirkung weiterer Determinanten, die neben bildungsprogrammatischen Setzungen ebenso die Unterrichtsgestaltung beeinflussen, ist in unserer Domäne noch zu wenig bekannt.

Wir stellen vor diesem Hintergrund Aspekte der didaktischen Gestaltung beruflichen Lernens in den Vordergrund und untersuchen insbesondere Sozialformen und den „methodischen Grundrhythmus“ von Unterricht i. S. Meyers (2010). Dabei zielt die Beobachtungsstudie (428 Stunden, 140 Lehrkräfte, 22 Schulen) darauf, deskriptiv zu erfassen, wie Unterrichtszeit genutzt wird sowie welche Sozialformen (additional differenziert) und Unterrichtsrythmen in welcher Intensität auftreten und wie diese zusammenhängen. Darüber hinaus erfassen wir Merkmale des Curriculums (Bildungsgang etc.), des Unterrichts (Inhalt etc.), der Klasse (Anzahl etc.), der Lehrkraft (Berufserfahrung, Qualifikation etc.) und der Schule (Anzahl Gebäude etc.), um mithilfe dieser potentiellen Determinanten Erklärungsansätze für die Variabilität in der Unterrichtsgestaltung zu identifizieren.

Im Zentrum des Beitrages soll zudem der Versuch stehen, mittels clusteranalytischer Verfahren typische Unterrichtsmuster zu isolieren sowie diese genauer zu beschreiben und in Beziehung zu den genannten Merkmalen des Unterrichts zu setzen.

Im Ergebnis können wir zeigen, dass sich einerseits die zentralen Befunde vorangegangener Studien – z.B. Pätzold et al. (2003) sowie Seifried, Grill und Wagner (2006) – replizieren lassen. Andererseits können wir sowohl Einflussfaktoren auf die didaktische Ausgestaltung als auch typische Muster in den Unterrichtsstrukturen identifizieren. Die Identifikation, Beschreibung und/oder Analyse solcher Unterrichtsmuster bzw. Unterrichtsskripts bzw. Inszenierungsmuster ist seit der internationalen Vergleichsstudie TIMSS 1995 (vgl. Stigler et al., 1999; Stigler & Hiebert, 1999) deutlich stärker in den Mittelpunkt des Forschungsinteresses in den mathematisch-naturwissenschaftlichen Fachdidaktiken gerückt (vgl. u.a. Lehrke et al., 2003; Pauli & Reusser, 2003; Seidel & Prenzel, 2004; Dalehefte, 2006; Heugner, Pauli & Reusser, 2007; Hugener, 2008). Auf diese Weise wird die isolierte Analyse einzelner Unterrichtsmerkmale überwunden und die Aufmerksamkeit richtet sich eher auf einen multikriteriellen Ansatz (vgl. u.a. Lehrke et al., 2003, S. 120; Hugener, 2008, S. 83). Es kann gezeigt werden, dass unterschiedliche Unterrichtstypen in der Unterrichtsrealität bestehen, die sich auf den Sozialformeneinsatz oder die Nutzung bestimmter Phasen beziehen. So lassen sich beispielsweise lehrer- bzw. schülerzentrierte Lehr-Lern-Arrangements von einander unterscheiden und rekonstruieren. Zudem lassen sich z.B. Erarbeitungs-, Übungs- und Geschlossene Stunden identifizieren (vgl. Götzl, Jahn & Held; 2013). Vor diesem Hintergrund können vor allem wesentliche Desiderata benannt werden.

Dalehefte, I. M. (2006): Unterrichtsskripts – ein multikriterieller Ansatz. Eine Videostudie zum Zusammenspiel von Mustern unterrichtlicher Aktivität, Zielorientierung und prozessorientierter Lernbegleitung. Kiel: Dissertation.

Götzl, M., Jahn, R. W., & Held, G. (2013): Bleibt alles anders!? Sozialformen, Unterrichtsphasen und echte Lernzeit im kaufmännischen Unterricht. In: bwp@ Berufs- und Wirtschaftspädagogik – online, Ausgabe 24, S. 1-23. Online: [http://www.bwpat.de/ausgabe24/goetzl\\_etal\\_bwpat24.pdf](http://www.bwpat.de/ausgabe24/goetzl_etal_bwpat24.pdf) (25-06-2013).

Heugner, I./Pauli, C./Reusser, K. (2007): Inszenierungsmuster, kognitive Aktivierung und Leistung im Mathematikunterricht. In: Lemmermöhle, M.: Professionell Lehren. Erfolgreich Lernen. Münster: Waxmann, S. 109-121.

Hugener I. (2008). Inszenierungsmuster im Unterricht und Lernqualität. Sichtstrukturen schweizerischen und deutschen Mathematikunterrichts in ihrer Beziehung zu Schülerwahrnehmung und Lernleistung – eine Videoanalyse. Münster u.a.: Waxmann 2008.

Lehrke, M. et al. (2003). Dominierende Skripts im Physikunterricht und Unterrichtsergebnisse. In: Pitton, A. (Hrsg.): Außerschulisches Lernen in Physik und Chemie. Münster u.a.: LIT, S. 120 ff.

Meyer, H. (2010). Was ist guter Unterricht? Berlin.

Pätzold, G. et al. (2003). Lehr-Lern-Methoden in der beruflichen Bildung. Eine empirische Untersuchung in ausgewählten Berufsfeldern. Oldenburg.

Pauli, C. & Reusser, K. (2003). Unterrichtsskripts im schweizerischen und im deutschen Mathematikunterricht. In: Unterrichtswissenschaft, 31, H. 3, S. 238-272.

Seidel, T. & Prenzel, M. (2004). Muster unterrichtlicher Aktivität im Physikunterricht. In: Doll, J. (Hrsg.): Bildungsqualität von Schule. Münster u.a.: Waxmann, S. 177-194.

Seifried, J., Grill, L., & Wagner, M. (2006). Unterrichtsmethoden in der kaufmännischen Unterrichtspraxis. In: Wirtschaft und Erziehung, 58, H. 7-8, S. 236-241.

Stigler, J. W. et al. (1999). The TIMSS Videotape Classroom Study. Methods and Findings from an Exploratory Research Project on Eight-Grade Mathematics Instruction in Germany, Japan, and the United States. In: U.S. Department of Education (Hrsg.): A Research and Development Report. Washington.

Stigler, J. W. & Hiebert, J. (1999). The Teaching Gap. New York: Free press.

Chair(s): Marc Kleinknecht (TU München), Mirjam Steffensky (IPN Kiel), Eva-Maria Lankes (TU München)

Diskutant/in: Christine Pauli (Universität Freiburg/Fribourg)

Im Rahmen einer kompetenzorientierten Lehrerbildung stellt der reflexiv-wissensbasierte Umgang mit Praxissituationen eine zentrale Herausforderung dar (Baumert & Kunter, 2006). Der ersten Phase der Lehrerbildung fällt dabei die Aufgabe zu, deklarative Wissensbestände zu vermitteln und das Analysieren und Reflektieren von Lehr-Lern-Situationen, z.B. im Rahmen von Schulpraktika anzubahnen. Die zweite und dritte Phase fokussiert dagegen stärker auf die Reflexion des eigenen Handelns (Neuweg, 2002). Der Fallarbeit wird in allen Phasen der Lehrerbildung ein hohes Potenzial bei der Förderung von Kompetenzen im Bereiche der Analyse und Reflexion zugesprochen. Sie bietet die Möglichkeit, Situationen des Unterrichts- und Schulalltags aus nicht-teilnehmender Distanz systematisch zu analysieren, das heißt theoretische Konzepte für die Deutung von Einzelphänomenen zu nutzen und gleichzeitig der Theorie als auch der spezifischen Handlungslogik des Einzelfalls gerecht zu werden (Helsper, 2004). Insbesondere Videofälle eröffnen den Lernenden z.B. die Gelegenheit, alltägliche Entscheidungssituationen nachzuvollziehen, Situationen aus verschiedenen Blickwinkeln zu betrachten und den Fall in einer von allen geteilten Fachsprache zu diskutieren (Sherin, 2004). Bislang gibt es allerdings kaum Evidenz für die Wirksamkeit der (videobasierten) Fallarbeit.

Das Symposium präsentiert verschiedene theoretische Zugänge zur videobasierten Fallarbeit in den Domänen Erziehungswissenschaft, Naturwissenschafts-, Fremdsprachen- bzw. Geschichtsdidaktik und beleuchtet die Wirkungen der videobasierter Konzepte auf die Kompetenzen der Lehramtsstudierenden und Lehrkräfte. Der erste Beitrag aus dem Bereich der Erziehungswissenschaft widmet sich der Frage, wie sich die Analysefähigkeit von Klassenführung in fallbasierten Lernsettings im Bereich der Schulpädagogik fördern lässt. Es werden die Wirkungen eines problemorientierten gegenüber einem instruktionalen Text- oder Videofallseminar auf kognitive Belastung, Motivation und Emotionen sowie die Analysefähigkeit von Lehramtsstudierenden untersucht. Der zweite Beitrag thematisiert das videobasierte Lernen in der naturwissenschaftsbezogenen Grundschullehrerausbildung, wobei ein videofallbasiertes mit einem textfallbasierten Seminar und einem Seminar ohne Fallbeispiele verglichen wird. Der Fokus der Studie liegt auf der professionellen Wahrnehmung der inhaltlichen Strukturierung und der kognitiven Aktivierung. Der dritte Beitrag zielt auf die Förderung der fachbezogenen Kompetenzorientierung im Geschichtsunterricht durch ein videobasiertes Fortbildungsangebot. In der Studie wird exploriert, inwieweit sich das Wissen und Handeln der teilnehmenden Lehrkräfte nach der Fortbildung verändert hat. Der vierte Beitrag stellt den Einsatz von Videos in der Ausbildung von Lehramtsstudierenden im Fach Englisch in den Mittelpunkt. In der Studie werden die Beobachtungshilfen für strukturierte bzw. unstrukturierte Videofälle gezielt variiert und die Effekte auf die einzelnen Subprozesse der professionellen Wahrnehmung untersucht. Das Symposium schließt mit der Diskussion der Chancen und Grenzen der Fallarbeit in verschiedenen Phasen der Lehrerbildung.

Dienstag,

04.03.2014,

10:15 - 12:00,

IG-Farben-Haus

Raum 0.251

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9, 469-520.

Helsper, W. (2004). Pädagogische Professionalität als Gegenstand des erziehungswissenschaftlichen Diskurses. Einführung in den Thementeil. Zeitschrift für Pädagogik, 50, 303-308.

Neuweg, G. H. (2002). Lehrerhandeln und Lehrerbildung im Lichte des Konzepts des impliziten Wissens. Zeitschrift für Pädagogik, 48, 10-29.

Sherin, M.G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (2004). Using video in teacher education (S. 1-27). Oxford, UK: Elsevier.





## **Fallbasiertes Lernen in der Schulpädagogik am Beispiel Classroom-Management. Eine Interventionsstudie zu Effekten verschiedener fallbasierter Lernumgebungen**

Der professionelle Umgang mit unterrichtlicher Komplexität und der Transfer von wissenschaftlicher Theorie in die Unterrichtspraxis sind häufig beobachtete Herausforderungen des Lehrerberufes. Mittels Fällen kann in Hochschulseminaren die Komplexität unterrichtlichen Handelns vermittelt werden. An ihnen lassen sich Kompetenzen (z.B. zu Classroom Management) anbahnen, die später im pädagogischen Handeln adäquat angewendet werden können (Piowar, Thiel, & Ophardt, 2013). Um situiertes Wissen aus Fällen aufzubauen, fokussieren aktuelle Konzepte in der Lehrerbildung vor allem die unterrichtliche Analysekompetenz als Facette der Reflexionskompetenz. Eine theoretische Verankerung des Analyseprozesses bleibt jedoch in der Fachliteratur zum fallbasierten Lernen in der Lehrerbildung oft unscharf (Lindmeier, 2013). Vor diesem Hintergrund wurde für die vorliegende Studie Analysekompetenz von Unterricht als Fähigkeit zur Selektion und Analyse von Lehr- und Lernsituationen definiert. Dabei umfasst die Selektion das Aufgreifen relevanter, inhaltsbezogener Situationen. Die Analyse enthält in Anlehnung an Dewey (1951) Phasen, in denen die Ausgangssituation beschrieben und die wahrgenommenen und alternativen Optionen (z.B. der Lehrperson), die Ursachen und Konsequenzen dieser Optionen sowie die Entscheidung für eine Option benannt werden. Bisherige Studien haben gezeigt, dass die Analyse von Unterrichtsvideos die Analysekompetenz von Lehramtsstudierenden und unterrichtenden Lehrkräften steigert (Tripp & Rich, 2012; Seidel et al., 2007; Sherin & van Es, 2009). Fallbasiertes Lernen mit Unterrichtsvideos oder -texten ermöglicht die Analyse von Interaktionen im Klassenzimmer, Lehrerverhalten und Lernprozessen seitens der Schülerinnen und Schüler ohne Handlungsdruck (Yadev et al., 2011; Sherin, 2004). Die Arbeit mit Video- oder Textfällen erlaubt eine theoriebasierte Analyse von komplexen Situationen. Dabei dienen die Fälle nicht nur der bloßen Illustration, sondern werfen als exemplarische Schlüsselsituationen pädagogischen Handelns Probleme auf, die einen Analyseprozess initiieren. In der Interventionsstudie fokussieren wir das Thema Classroom-Management, welches ein wichtiger Bestandteil des pädagogischen Wissens ist (Shulman, 1986; Piowar, Thiel, & Ophardt, 2013). Lehrkräfte sollten wissen, wie Klassenaktivitäten organisiert und gestaltet werden, um eine höchstmögliche Lernzeit für Schülerinnen und Schüler zu garantieren. Erste Studien zeigen, dass es bereits in der ersten Phase der Lehrerbildung möglich ist, ein gutes Classroom-Management zu erlernen, ebenso wie Analysekompetenz zur Identifikation relevanter Situationen (Piowar, Thiel, & Ophardt, 2013; Sherin, 2007). In der Studie wurde untersucht, welche Effekte verschiedene fallbasierte Lernumgebungen auf Motivation, Emotionen und kognitive Belastung bei der Fallarbeit haben und in welchem Umfang diese die Analysekompetenz Lehramtsstudierender fördern. Zur Beantwortung der Forschungsfragen wurden in einer Interventionsstudie ca. 700 Lehramtsstudierende unterschiedlichen fallbasierten Arrangements in einem 2x2-Design zugeordnet. Einerseits wurden die den Seminaren zugrunde liegenden Lehr-Lern-Modelle (instruktional vs. problemorientiert) und andererseits das Fallmedium (Text vs. Video) variiert. In einer Kontrollgruppe wurde lediglich theoretische Literatur zum Thema Classroom-Management bearbeitet. Ein webbasierter Pre-Post-Test erhob die Veränderung der Analysekompetenz. Während der Fallarbeit in den Seminaren wurden mittels eines Begleitfragebogens kognitive Belastung (Sweller, 2007), Lern-/Leistungsemotionen (Eccles et al., 1983; Barrett & Russell, 1999) und Motivation bzw. Immersion (Kleinknecht & Schneider, 2013) als abhängige Variablen erfasst. Erste Ergebnisse zeigen in allen vier verschiedenen Bedingungen eine Zunahme der Analysekompetenz vor allem mit Blick auf die Einbindung theoretischer Aspekte in die Analysen. Des Weiteren kann ein Vorteil des videofallbasierten und problemorientierten Lernens für die abhängigen Variablen Immersion und Interesse festgestellt werden. Hinsichtlich der Motivation unterscheiden sich die vier Bedingungen entgegen der Vorannahmen nicht. Die kognitive Belastung nimmt in allen Seminaren nach einer Gewöhnungszeit ab, jedoch ist das Belastungsniveau in videobasierten Seminaren höher. Insgesamt deuten die Befunde darauf hin, dass fallbasierte Seminare in den vier Formen dazu geeignet sind, die Analysekompetenz Studierender zu fördern. Gleichzeitig wird fallbasiertes Lernen in allen vier Formen als motivierend und überwiegend als kognitiv angemessen wahrgenommen.

Cornelia Sunder (Universität Münster), Maria Todorova (Universität Münster), Mirjam Steffensky (IPN Kiel), Kornelia Möller (Universität Münster)

## ***Videobasiertes Lernen in der naturwissenschaftsbezogenen Grundschullehrerbildung***

Im Mittelpunkt dieses Beitrages steht die Förderung der professionellen Wahrnehmung des naturwissenschaftlichen Grundschulunterrichts. Professionelle Wahrnehmung beinhaltet das Erkennen und Interpretieren von lernrelevanten Ereignissen im Unterricht (Sherin, 2007). Sie gilt als grundlegend für das unterrichtliche Handeln und Reflektieren (Hammerness, Darling-Hammond, & Shulman, 2002). In neueren Studien zum Mathematikunterricht konnte auch ein Zusammenhang zwischen der professionellen Wahrnehmung der Lehrkräfte und den Schülerleistungen gezeigt werden (Roth et al., 2011; Kersting et al., 2012). Professionelle Wahrnehmung wird in diesem Beitrag aus einer naturwissenschaftsdidaktischen Perspektive untersucht, wobei der Fokus auf der Lernunterstützung in der Grundschule liegt; diese bezieht sich auf Maßnahmen der kognitiven Aktivierung und inhaltlichen Strukturierung.

Eine wichtige Frage in der Lehrerbildungsdiskussion ist, wie sich die professionelle Wahrnehmung bzw. vergleichbare situationsgebundene Wissensbestände und Fähigkeiten fördern lassen. Häufig werden dazu videobasierte Lehr-Lernumgebungen eingesetzt (Blomberg et al., 2013). Es fehlt aber an kontrollierten Studien zur Untersuchung der Wirksamkeit solcher Lernumgebungen, was gerade auch angesichts des (immer noch) großen Aufwandes beim Einsatz von Videos als problematisch erscheint.

Vor diesem Hintergrund ist ein Ziel des Projektes ViU (Videobasierte Unterrichtsanalyse), den Einsatz von Videos in der universitären Lehre zur Förderung der professionellen Wahrnehmung zu untersuchen. Hierzu wird ein videobasiertes und fallbezogenes Seminar (Experimentalgruppe) mit einem textbasierten und fallbezogenen Seminar (Kontrollgruppe 1) sowie mit einer Seminargruppe verglichen, die eine Veranstaltung mit entsprechenden theoretischen Inhalten, aber ohne fallbasierte Lernelemente (Kontrollgruppe 2) erhält. Zur Erfassung der professionellen Wahrnehmung werden in einem Prä-Post-Follow-Up-Design ein standardisierter Videotest sowie offene Analysen von Vignetten eingesetzt. Die Studie wird mit zwei Kohorten im Wintersemester 13/14 sowie 14/15 durchgeführt.

In einer Pilotierungsstudie wurden das Konzept und die Umsetzbarkeit der videobasierten Intervention an einer Stichprobe von 22 BA-Studierenden getestet. Im Vergleich zu einer Kontrollgruppe (N = 17) verbesserte sich die professionelle Wahrnehmung der Interventionsgruppe signifikant bezüglich Maßnahmen der inhaltlichen Strukturierung ( $F(1,36) = 4.18, p < .05$ ), allerdings nicht bezüglich Maßnahmen der kognitiven Aktivierung ( $F(1,36) = 0.52, p = .48$ ). Im Vortrag werden das Konzept der videobasierten Intervention sowie die Pilotierungsergebnisse vorgestellt und diskutiert.

## **Videobasierte Unterrichtsentwicklung im Fach Geschichte. Ergebnisse eines Interventionsprojekts zur Förderung historischer Kompetenzen**

Unterrichtsvideos gelten als anschauliche und prozessnahe Grundlage zur Reflexion und Diskussion über Fragen des Unterrichtens. Sie erfreuen sich zunehmender Beliebtheit wegen ihrer Möglichkeit, Unterrichtsprozesse in ihrer Vielschichtigkeit sowie aus räumlicher und zeitlicher Distanz zu zeigen (Brophy, 2004). Ein großes Potential wird unter anderem in der videobasierten Unterrichtsanalyse bzw. Unterrichtsreflexion als Form des situierten und fallbasierten Lernens vermutet (Borko et al., 2008; Brophy, 2004; Krammer & Reusser, 2005). Bereits vorliegende empirische Befunde belegen einen Zusammenhang zwischen videobasierter Unterrichtsreflexion und einer zunehmend differenzierten Wahrnehmung des Unterrichtsgeschehens (z.B. Sherin & van Es, 2009; Krammer et al., 2010). Hervorgehoben wird aber auch die Wichtigkeit der Orientierung an fachlichen Lernzielen und empirisch abgesicherten Merkmalen der Unterrichtsqualität bei der Auswahl von aussagekräftigen Videosequenzen und der Festlegung von Beobachtungsgesichtspunkten, die es erlauben, über disziplinäre Fragen des Unterrichtens systematisch nachzudenken (Hiebert et al., 2007; Santagata & Guarino, 2010; Reusser et al., 2007). Aus fachdidaktischer Perspektive und basierend auf den empirischen Ergebnissen der Studie „Geschichte und Politik im Unterricht“ wird für den Geschichtsunterricht auf der Sekundarstufe I in der Schweiz ein Entwicklungsbedarf in der Gestaltung prozessorientierter Lerngelegenheiten geortet. Vorliegende empirische Befunde zeigen, dass die Vermittlung von Sachwissen dominiert und auch die im Unterricht eingesetzten Lernaufgaben schwergewichtig auf die Rezeption bzw. Erschliessung von Informationen abzielen (Messner & Buff, 2007; Waldis et al., 2012). Der Förderung und Einübung prozeduraler Fähigkeiten und metakognitiven Wissens wie sie derzeit von der deutschsprachigen Geschichtsdidaktik in Kompetenzmodellen beschrieben (Gautschi, 2009; Schreiber et al., 2006) und in der englischsprachigen Literatur subsumierend unter Begriffen wie „historical reasoning activities“ oder „historical thinking concepts“ gefasst werden (Seixas & Morton, 2013; van Drie & van Boxtel, 2008) wird im Unterricht deutlich weniger Platz eingeräumt. Soll die Förderung von fachspezifischen Kompetenzen bzw. historischem Denken im Geschichtsunterricht verstärkt zum Zuge kommen, so müssen Lehrpersonen ihr fachdidaktisches Wissen über geeignete Instruktionsstrategien und Aufgabenformate sowie ihre Kenntnisse über Kognitionen der Schülerinnen und Schüler erweitern (Shulman, 1986). Dieses Ziel wurde mit der Interventionsstudie „Erweiterung professioneller Handlungskompetenzen von Geschichtslehrpersonen“ verfolgt.

Es wurde eine rund 50 Stunden beanspruchende Fortbildung für Lehrpersonen der Sekundarstufe I konzipiert, die die Einführung in die theoretische Kompetenzdebatte mit der Analyse fremder Videosequenzen und der videobasierten Reflexion eigener Videosequenzen kombinierte. Ziel war es, die Lehrpersonen für das Anliegen der Kompetenzförderung im Geschichtsunterricht zu sensibilisieren und erste Schritte der Unterrichtsentwicklung einzuleiten. Zur Evaluation des gewählten Interventionsansatzes wurde bei den im Weiterbildungskurs beteiligten Lehrpersonen (n = 8) sowie bei Kontrollgruppenlehrpersonen (n = 8) vor und nach Kursbesuch je eine Doppellektion Geschichte videografiert und ein darauf bezogenes halbstrukturiertes Interview durchgeführt. Auf der Basis dieser Daten wurde untersucht, welche Maßnahmen und Aktivitäten zur Kompetenzförderung genannt werden und ob Veränderungen in den Unterrichtskommentaren (Interviews) und im Unterrichtshandeln (Videoanalysen) zwischen den beiden Erhebungszeitpunkten sichtbar werden.

Es zeigte sich, dass die Experimentalgruppenlehrpersonen die Unterrichtsreflexion auf der Basis fremder und eigener Videoaufnahmen als äußerst produktiv empfanden. Die Kenntnisse über und die Akzeptanz des in der Weiterbildung eingeführten Kompetenzmodells fielen in der Experimentalgruppe deutlich höher aus als in der Kontrollgruppe. Trotzdem war es auch nach Kursbesuch für die Experimentalgruppenlehrpersonen keine leichte Aufgabe, theoretisch begründete Leitideen auf konkrete Situationen des eigenen didaktischen Handelns zu beziehen und die getroffenen Entscheidungen auf der vorliegenden Theoriebasis zu reflektieren. Die Videoanalysen liefern darüber hinaus keinerlei Hinweise für grundsätzliche Umstellungen der Unterrichtspraxis. Die empirischen Ergebnisse machen deutlich, dass mittels videobasierter Fallbeispielen die Unterrichtsentwicklung angeregt werden konnte und zumindest auf der Wissensebene die Lehrpersonen erreicht wurden. Zugleich zeigte sich, dass eine tiefgreifende Umstellung der Unterrichtspraxis wie sie die Kompetenzorientierung im Geschichtsunterricht mit sich ziehen würde, deutlich voraussetzungsvoller ist und weiterer Maßnahmen bedarf.



## ***Erarbeitung der Videofälle im VideoWeb. Ergebnisse einer Vorstudie mit EFL Lehramtsstudierenden***

Die Nutzung von Video in der Lehrerbildung hat relativ lange Tradition und verfolgt verschiedene Ziele (Brophy, 2004; Janík, Minaříková & Najvar, 2013 u.a.). Im Kontext „neuer“ Lehr-Lernkultur eröffnet sich die Frage, wie die Unterrichtsvideos – eingebettet in diverse instructional settings – das Lernen von LehrerInnen beeinflussen und welche Ansätze und Formate in Bezug auf welche Ziele sich als wirksam erweisen (Seidel et al., 2011; Blomberg et al., 2011). Unsere qualitative Vorstudie – die wir im Rahmen dieses Symposiums präsentieren wollen – zielt auf die oben angeführten Fragen. Die Vorstudie ist ein Teil eines breiteren Projektes GA13-21961S: Exploring professional vision and its development through video-based analysis in dem wir uns mit der Erfassung und Entwicklung der professional vision von angehenden Lehrkräften mit Englisch als Fremdsprache (EFL-Lehramtsstudierenden) beschäftigen.

Als Instrument wird eine videobasierte online Lernumgebung (IRSE VideoWeb) benutzt. IRSE VideoWeb besteht aus mehreren thematischen Modulen, die grundlegende Fragen zur Didaktik des Englischen als Fremdsprache bearbeiten (lassen). Jedes Modul besteht aus zwei (bzw. drei) strukturierten Videofällen (kurze Videosequenzen mit auf knowledge-based reasoning gezielten offenen Fragen, zusätzlichen Materialien, Experten-Kommentare u.a.) und einem Videofall ohne Strukturierung (kurze Videosequenz mit einer offenen generellen Frage). Die Fragen in den strukturierten Videofällen gehen von einem Beobachtungsrahmen aus, der die einzelnen Subprozesse des knowledge-based reasoning (Beschreiben, Interpretieren, Erklären, Prädiktion vorlegen, Bewerten, Alternative vorschlagen) betrifft. Ziel unserer Vorstudie ist es festzustellen, wie sich das Erarbeiten der (strukturierten vs. unstrukturierten) Videofälle unterscheidet, wenn ihnen der Beobachtungsrahmen explizit erläutert wurde versus wenn er nur implizit aus den Fragen bei Videofällen „zu entnehmen“ ist. Aus unserer Gesamtstichprobe der EFL-Lehramtsstudierenden wurden für diese Analyse zwei Gruppen gebildet: (A) Gruppe mit der Einführung zu den Beobachtungsrahmen (ME; 20 EFL-Lehramtsstudierenden) – die Studierenden wurden über den Beobachtungsrahmen innerhalb einer Präsenzlehrveranstaltung informiert und haben ihn auch bei der Gruppendiskussionen über Videosequenzen benutzt; (B) Gruppe ohne Einführung zu den Beobachtungsrahmen (OE; 12 EFL-Lehramtsstudierenden) – die Studierende wurden nur über das generelle Ziel informiert, keine weitere Anweisung war vorhanden.

Die Forschungsfragen unserer Studie im Bezug auf die strukturierten Videofälle sind: (1) An welche Subprozesse des knowledge-based reasoning kann in den Antworten/Kommentaren der Studierenden erschlossen werden? (2) Wie stimmen die die Antworten/Kommentaren mit den von Fragen gezielten knowledge-based reasoning Subprozessen überein? (3) Wie unterscheiden sich inhaltlich die Antworten/Kommentare bei den Studierenden in den zwei Gruppen? Die Forschungsfragen zu den unstrukturierten Videofällen sind: (4) Wie sind die Antworten/Kommentare der Studierenden (in Bezug auf den Subprozessen) strukturiert? (5) Wie unterscheiden sich inhaltlich die Antworten/Kommentare bei den Studierenden in zwei Gruppen?

Die Daten werden mit Hilfe von qualitativen Inhaltsanalysen ausgewertet, es werden diverse theorie- und datengeleitete Kategoriensysteme benutzt. Vorläufige Ergebnisse deuten daran, dass alle Subprozesse des knowledge-based reasoning bei der Erarbeitung strukturierten Videofällen in beiden Gruppen in der ähnlichen Proportion vertreten wurden. In der ME-Gruppe stimmen die Subprozesse besser mit Zielen der entsprechenden Fragen überein. Den Inhalt der Antworten betreffend, waren die Antworten innerhalb der ME Gruppe thematisch ähnlich gerichtet, während innerhalb der OE Gruppe sich die Antworten/Kommentare vielfältiger gestalteten. In den Antworten/Kommentaren bei den unstrukturierten Videofällen haben die Studierende in der OE-Gruppe die Unterrichtssituationen eher beschrieben, bewertet oder Alternativen vorgeschlagen; die ME-Gruppe hat mehr interpretiert und Prädiktionen vorgelegt. Die OE-Gruppe fokussierte in den Antworten/Kommentaren der Lehrkraft und „ihr“ Unterrichten, die ME-Gruppe thematisierte dagegen viel mehr die Schülerinnen und Schüler und ihre (vermutliche) Wahrnehmung der Unterrichtssituation.

Diese vorläufige Ergebnisse sind von begrenzter Aussagekraft, breiter gefasste Analyse werden zur Zeit realisiert. Die Ergebnisse sollten im Frühjahr 2014 vorhanden sein und präsentiert werden. Zu Beantwortung der Frage, wie die Arbeit der EFL-Lehramtsstudierenden mit Videofällen unter unterschiedlichen Bedingungen erfolgt und wie diese optimal zu unterstützen ist, besteht noch erheblicher Forschungsbedarf.

Chair(s): Jan Retelsdorf (IPN), Olaf Köller (IPN)

Diskutant/in: Oliver Wilhelm (Universität Ulm)

Theoretisch kommt schulisch erworbenen Kompetenzen eine hohe Relevanz für eine erfolgreiche berufliche Erstausbildung zu – die empirische Evidenz für diese Annahme ist jedoch weniger eindeutig. Am ehesten sind hier die von Rainer Lehmann in Hamburg durchgeführten ULME-Studien zu nennen, in denen sich schulische Kompetenzen in der Tat als signifikante Prädiktoren späterer beruflicher Kompetenzen erwiesen. Gleichzeitig ist die Frage nach der Trennbarkeit allgemeinbildender und beruflicher Kompetenzen nicht hinreichend geklärt. Vor diesem Hintergrund bündelt unser Symposium Beiträge, die sich der Modellierung und Relevanz schulischer und beruflicher Kompetenzen in Ausbildungsberufen mit hohen Anforderungen in den Bereichen Mathematik und Naturwissenschaften beschäftigen. Dafür stellen zunächst Nagy et al. aus einer erziehungswissenschaftlichen Perspektive erste Ergebnisse zur Lernausgangslage in schulisch vermittelten und berufsfeldspezifischen Kompetenzen von Auszubildenden dar. Zudem berichten Sie Zusammenhänge zwischen beiden Kompetenzbereichen. In dem zweiten Beitrag von Geißel et al. wird aus einer berufs- und technikpädagogischen Perspektive die Relevanz mathematischer Kompetenzen für den Erwerb berufsfachlicher Kompetenzen von Elektronikern, Mechatronikern, Fachinformatikern und in Bauberufen untersucht. Im dritten Beitrag von Winther und Sangmeister wird aus wirtschaftspädagogischer Perspektive ein Entwicklungsmodell ökonomischer Numeralität vorgestellt, in dem die Übergänge von allgemeinen kognitiven Fähigkeiten in das berufsspezifische Wissen dargestellt sind. Im vierten Beitrag werden aus einer naturwissenschaftsdidaktischen Perspektive erste Arbeiten zur Struktur beruflicher Kompetenzen für den Beruf des Chemielaboranten vorgestellt. Frank et al. verfolgen dabei vor allem die Frage der Dimensionalität sowie der Stabilität dieser Dimensionalität über den Verlauf der Ausbildung. Arrondiert wird das Symposium durch die Diskussion von Oliver Wilhelm, der die Beiträge zusammenfassend aus psychologischer Perspektive diskutiert.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
IG-Farben-Haus  
Raum 0.254





## ***Ausgangsleistungen von Auszubildenden in Berufen mit hohen mathematisch-naturwissenschaftlichen Anforderungen: Ergebnisse für schulisch vermittelte und berufsfeldspezifische Kompetenzen***

Die Frage nach der Bedeutung schulischer Kompetenzen für den Erfolg des Übertritts in die berufliche Erstausbildung hat spätestens mit Etablierung der großen Schulleistungsstudien hohe wissenschaftliche Relevanz erhalten. Rezente Literacy- und Kompetenzkonzepte betonen die Bedeutung der in der Schule erworbenen Kompetenzen für die berufliche Teilhabe, und sprechen diesen eine ursächliche Rolle für die Entwicklung berufsbezogenen Fähigkeiten zu. Bis heute fehlt es jedoch an einer soliden empirischen Untermauerung dieser Annahmen. Die Studie Mathematisch-naturwissenschaftliche Kompetenzen in der beruflichen Erstausbildung (ManKobE) widmet sich dieser Fragestellung unter besonderer Berücksichtigung mathematisch und/oder naturwissenschaftlich intensiver Ausbildungsbereiche (Industriemechaniker, KFZ-Mechatroniker, Elektroniker, Chemielaboranten, Biologielaboranten und Industriekaufleute). ManKobE ist als prospektive Large-Scale-Längsschnittstudie angelegt (drei Messzeitpunkte im Einjahresabstand). Die Instrumentierung der Studie umfasst Tests zur Erfassung schulisch erworbener Kompetenzen in den Bereichen Deutsch (C-Tests), Mathematik, Physik, Chemie und Biologie (Instrumente auf der Basis der länderübergreifenden Bildungsstandards). Diese werden u. a. um berufsfeldspezifische Tests ergänzt (Mathematik, Physik, Chemie und Biologie). Diese Tests aus Zwischenprüfungs- und Unterrichtsaufgaben zusammengestellt, die von der Industrie und Handelskammer entwickelt wurden. Sie umfassen Aufgaben, die mit Schulwissen gelöst werden können, jedoch in berufsfeldspezifischen Kontexten eingekleidet sind (Mathematik für Industriekaufleute, Physik für Mechaniker und Elektrotechniker, Chemie für Chemielaboranten und Biologie für Biolaboranten). Der vorliegende Beitrag widmet sich der Eingangsphase der Berufsausbildung. Er konzentriert sich auf die Daten der bis dato getesteten Ausbildungsgruppen (N = 2132; Industriekaufleute, Mechaniker, Mechatroniker, Elektroniker und Chemielaboranten). Im Zentrum stehen drei Fragen. Erstens soll der Leistungsstand der Auszubildenden auf den schulischen Leistungsmaßen berichtet werden. Der Einsatz der Bildungsstandardinstrumente ermöglichte auf der nationalen Metrik Leistungsvergleiche mit dem deutschen Durchschnitt am Ende des 9. Schuljahres. Zweitens werden Ergebnisse zur empirischen Trennbarkeit der schulischen Leistungstests (Bildungsstandard-Instrumente) und der neu entwickelten berufsfeldspezifischen Tests vorgestellt. Drittens werden die mittleren Leistungsprofile der untersuchten Ausbildungsgruppen berichtet. Insofern die Schulabgänger ihre Berufswahlentscheidungen gemäß ihrer individuellen Stärken und Schwächen treffen, sollten sich systematische Profilunterschiede ergeben. Die Befunde lassen sich wie folgt zusammenfassen. (1) Die Gesamtgruppe der untersuchten Auszubildenden wies in allen untersuchten schulischen mathematischen und naturwissenschaftlichen Domänen einen Leistungsstand auf, der auf oder oberhalb des nationalen Mittelwerts am Ende der Pflichtschulzeit lag. Die ermittelten Niveauunterschiede zwischen den unterschiedlichen Ausbildungsgängen konnten dabei größtenteils auf Unterschiede im Anteil der Auszubildenden mit Hochschulzugangsberechtigung zurückgeführt werden. Dieser Befund zeigt, dass die hier betrachteten Ausbildungsfelder Schülerinnen und Schüler anziehen, die sich durch relativ gute Leistungen im mathematisch-naturwissenschaftlichen Bereich auszeichnen. (2) Die nationalen Bildungsstandardtests ließen sich empirisch von den berufsfeldspezifischen Tests trennen. Die latenten messfehlerkorrigierten Korrelationen innerhalb der Domänen bewegten sich einem Bereich von  $r = 0.83$  (Mathematik) bis  $r = 0.66$  (Biologie). Die am Ende der Schulzeit erreichten individuellen Leistungsniveaus sind somit deutlich mit den berufsfeldspezifischen Eingangsleistungen korreliert, sind jedoch nicht mit diesen gleichzusetzen. (3) Die Ausbildungsgruppen unterschieden sich leicht in den Mittelwertprofilen der nationalen Bildungsstandardinstrumente. Die Profilunterschiede fielen für den Bereich der berufsfeldspezifischen Leistungen deutlich akzentuierter aus. Für beide Testkonzeptionen zeigte sich, dass die Auszubildenden ihre Stärken in den für die jeweiligen Ausbildungsberufe zentralen Domänen aufwiesen (Mechaniker und Elektrotechniker: Physik, Industriekaufleute: Mathematik und Chemielaboranten: Chemie). Diese Ergebnisse sprechen dafür, dass die individuellen Ausbildungswahlen im Zusammenhang mit den am Ende der Schulzeit erworbenen Stärken stehen. Die Profilunterschiede könnten jedoch auch – vor allem im Bereich der berufsfeldspezifischen Kompetenzen – zum Teil unterschiedlichen Lerngelegenheiten geschuldet sein. Insgesamt legen die Ergebnisse des ersten Messzeitpunkts der ManKobE-Studie ein ausreichendes bis gutes Niveau schulisch erworbener mathematisch-naturwissenschaftlicher Kompetenzen und eine gute Passung individueller Stärken an die spezifischen Anforderungen der Ausbildungsberufe für einen großen Teil der getesteten Auszubildenden und eine günstige Prognose des Ausbildungserfolgs nahe. Inwieweit diese Vermutung zutrifft und inwieweit sich zukünftige individuelle Lerngewinne vorhersagen lassen, wird sich in den späteren Erhebungen der ManKobE-Studie erweisen.



## ***Relevanz und prädiktive Kraft mathematischer Kompetenzen in der gewerblich-technischen Berufsbildung***

Sowohl Studien zur Eignungsdiagnostik als auch verschiedenen Studien zur Entwicklung berufsfachlicher Kompetenzen (im Überblick Abele, 2013) dokumentieren die hohe prädiktive Kraft mathematischer Fähigkeiten für die berufsfachliche Kompetenzentwicklung, wobei berufsspezifisch erhebliche Unterschiede beobachtet werden können (Lehmann/Seeber, 2007). Vor allem in gewerblich – technischen Berufen lassen sich z.T. auch bei Kontrolle des IQ substantielle Anteile der erklärten Varianz fachlicher Kompetenzen auf die mathematischen Leistungen zurückführen (Nickolaus et al., 2008; 2010; 2011; 2012). Dies scheint selbst dann der Fall zu sein, wenn innerhalb der beruflichen Anforderungen mathematische Leistungen nicht direkt relevant werden (Geißel et al., 2013). Denkbar wäre, dass dafür a) in den mathematischen Leistungsdaten inkorporierte metakognitive oder auch motivationale (Traitkomponente) Momente ursächlich sind oder auch b) die von Geißel et al. vorgelegten Befunde an spezifische Domänen gebunden sind, in welchen mathematikaffine Anforderungen relevant werden. In diesem Beitrag soll der Frage nachgegangen werden, inwieweit sich die von Geißel u.a vorgelegten Befunde auch in anderen Domänen bestätigen lassen. Vorgestellt werden Analyseergebnisse für ausgewählte gewerblich-technische Berufe (1) zu den mathematischen Anforderungen beruflicher Tätigkeiten, (2) den mathematischen Anforderungen in berufsfachlichen Kompetenztests und der prädiktiven Kraft mathematischer Leistungsdaten für die Entwicklung berufsfachlicher Kompetenzen. Nachgegangen wird dabei der Frage, inwieweit die mathematischen Fähigkeiten bei Kontrolle des IQ auch für fachliche Anforderungen prädiktiv werden, in welche keine mathematischen Anforderungen inkorporiert sind. Basis der Präsentationen sind Studien in vier Domänen (Elektroniker, Mechatroniker, Fachinformatiker und Bauberufe (Grundstufe)), die von Seiten der DFG und dem Bundesinstitut für Berufsbildung finanziert wurden. Die Stichprobengrößen variieren zwischen 150 und 680 Probanden, neben den mathematischen Fähigkeiten wurden der IQ und das berufsfachliche Vorwissen sowie motivationale Daten und ausgewählten Qualitätsmerkmale der didaktischen Settings in längsschnittlichen Designs erhoben.

## **Wie viel Bernoulli steckt in kaufmännischen Ausbildungsgängen? – Entwicklungsmodelle ökonomischer Numeralität**

Theoretischer Hintergrund: Der Erwerb professioneller Kompetenzen, die sich vorrangig über domänenspezifische Kenntnisse und Fertigkeiten in realen Arbeits- und Geschäftsprozessen konstituieren, wird durch mathematische, kontextbezogene Fähigkeiten unterstützt (vgl. Winther et al., 2013). In kaufmännischen Ausbildungsgängen werden mathematische Routinen und Algorithmen in Instruktionsprozessen an die wirtschaftlichen Fachinhalte gebunden; integrative Lehr-/Lernideen, die sich bis in die wirtschaftswissenschaftlichen Studiengänge und in die betrieblichen Weiterbildungskonzepte fortsetzen. Diese Grundbedingungen werden im Rahmen eines kaufmännischen Kompetenzmodells aufgegriffen, das zentral zwischen spezifisch fachlichen und cross-curricular allgemeinen Kompetenzstrukturen differenziert und unter einer Kompetenzentwicklungsperspektive den beruflichen Kompetenzerwerb durch Übergänge von basalen kognitiven Fähigkeiten hin zu domänenverbundenen und domänenspezifischen Denk- und Handlungsschemata beschreibt. Zu den domänenverbundenen Kompetenzstrukturen zählen ökonomische Literalitäts- und Numeralitätsannahmen, mit denen die für das (wirtschafts-)berufliche Handeln notwendigen kaufmännischen Grundfertigkeiten abgebildet werden. In Anlehnung an den International Adult Literacy Survey (IALS; OECD, 1995) sowie an die vom National Council on Economic Education herausgegebenen Standards (2010) werden text- und bildsprachliche Kenntnisse sowie das Verständnis von quantitativen Werten und Verhältnissen in wirtschaftlichen Kontexten erfasst.

Fragestellungen: Ein zentrales Forschungsdesiderat stellt die Frage dar, welchen Beitrag cross-curricular ökonomische Kompetenzen für den Auf- und Ausbau beruflicher Professionalität leisten (vgl. ManKobE; Nickolaus et al., 2013). Aus dem beruflichen Bereich ist bekannt, dass sich insbesondere mathematische Defizite als schwer überwindbare Barrieren der beruflichen Kompetenzentwicklung erweisen (für den gewerblich-technischen Bereich vgl. Nickolaus & Ziegler, 2005; Nickolaus, Gschwendtner, & Geißel, 2008; für den kaufmännischen Bereich vgl. u. a. Lehmann & Seeber, 2007; Winther, 2006). In welchem Maße und in welchen Stadien der beruflichen Professionalisierung dies besonders tragend ist, ist hingegen noch nicht hinreichend erforscht. Grundlage hierfür wären Kompetenzentwicklungsmodelle, die zentral auf die Enkulturation in eine Expertenkultur (Shavelson, 2010) abstellen. Dies impliziert u. a. die Abbildung der Übergangsprozesse von allgemeinen kognitiven Fähigkeiten zu domänenspezifischen wissensbasierten Entscheidungen sowie das Aufzeigen konkreter Enkulturationswege vor dem Hintergrund der Ausbildung von Arbeitsroutinen und informell erworbener Fähigkeiten.

Methode und Ergebnisse: Der Beitrag arbeitet ein Entwicklungsmodell ökonomischer Numeralität heraus, das empirisch validiert ist. Ökonomische Numeralität wird über simulationsbasierte Kompetenztests erfasst, mit denen sich kognitive Verstehensmuster sowie das Arbeitsverhalten der Auszubildenden beobachten und diagnostizieren lässt. Simulationsbasierte Verfahren lassen sich im Hinblick auf berufliche Kompetenzmessung über folgende Kriterien charakterisieren (vgl. Winther, 2010): (1) Eine authentische Abbildung komplexer berufstypischer Situationen, Zahlen, Fakten, Arbeitsmittel und Beziehungen, (2) die Abbildung vollständiger Arbeits- und Geschäftsprozesse sowie (3) eine Standardisierung der Arbeits- und Geschäftsprozesse durch Unternehmensmodellierung. Die besondere Herausforderung beruflicher Kompetenzmessverfahren besteht darin, dass die konstruierten Anforderungssituationen realen betrieblichen Vorgängen und Fragestellungen entnommen und damit zwangsläufig einer Realitätsdiskussion unterworfen werden – dies gilt insbesondere für mathematische Konzepte, die in der beruflichen Bildung vorrangig als unterstützend bewertet werden. Es werden Befunde präsentiert, die (1) die ökonomische Numeralität als signifikantes Korrelat der beruflichen Professionalisierung kennzeichnen und die (2) die Trajektorien ökonomischer Numeralität im Kontext authentischer Testszenarien beschreiben. Hierfür wurden, unter Nutzung eines Equivalent-Groups-Designs, vertikale Skalen (u. a. Young, 2006) administriert. So können neben DIF-Effekten, die sich vorrangig auf berufliche Vorerfahrungen zurückführen lassen auch Itemschwierigkeiten über verschiedene Ausbildungsstadien präsentiert werden (out-of-level-testing).

## **Struktur und Entwicklung beruflicher Kompetenzen von Chemielaboranten**

Die Modellierung und empirische Überprüfung von Kompetenzstrukturen im Bereich der beruflichen Ausbildung stellt erst seit kurzer Zeit einen Forschungsschwerpunkt innerhalb der beruflichen Bildung dar (Seeber et al., 2010). Während die Beschreibung allgemeinbildender Kompetenzen normativ erfolgt, werden berufliche Kompetenzen performanzorientiert modelliert (Erpenbeck, 2009). Empirische Erkenntnisse liegen vor allem für gewerblich-technische Berufe vor, wobei die Struktur und Entwicklung beruflicher Kompetenzen bisher für den Beruf des Chemielaboranten kaum untersucht ist. Dieser ist jedoch besonders interessant, um Informationen über Gemeinsamkeiten und Unterschiede allgemeiner und beruflicher Kompetenzen zu gewinnen. Denn im Gegensatz zu dem Großteil an Ausbildungsberufen weist der des Chemielaboranten eine eindeutige Bezugswissenschaft auf, die zudem auch in allgemeinen Bildungsgängen unterrichtet wird. Der Rahmenlehrplan für den Ausbildungsberuf besitzt daher auch starke inhaltliche Übereinstimmungen mit den in den Bildungsstandards angegebenen Kompetenzen für das Fach Chemie auf.

Aus Untersuchungen ist bekannt, dass der allgemeinbildenden Kompetenz für das Fach Chemie wahrscheinlich eine eindimensionale Kompetenzstruktur zugrunde liegt (Bernholt, 2010; Ferber, Emden, & Sumfleth, (im Druck); Ropohl, 2010). Analysen hinsichtlich der Anforderungsstruktur des Berufs (Eichhorn, 2007; Storz & Fries, 1997) ebenso wie der Rahmenlehrplan für den Ausbildungsberuf lassen vermuten, dass sich die beruflichen Kompetenzen von Chemielaboranten in Bezug zu den Schwerpunkten „Analyse von Stoffen“ und „Synthese von Stoffen“ ausbilden. Dass die Dimensionen beruflicher Kompetenzen im Wesentlichen den Anforderungsstrukturen des Berufes entsprechen, konnte für andere Ausbildungsberufe nachgewiesen werden (Nickolaus, Lazar, & Norwig, 2012). Es soll daher folgende Forschungsfrage untersucht werden:

(1) Kann die Kompetenzstruktur von Chemielaboranten entsprechend der theoretischen Annahme auch empirische in die beiden Dimensionen „Analyse von Stoffen“ und „Synthese von Stoffen“ differenziert werden?

Weiterhin ist aus anderen Studien bekannt, dass im Verlauf der Berufsausbildung Verschmelzungs- und Ausdifferenzierungsprozesse der Kompetenzstruktur in Bezug zu den curricularen Schwerpunktsetzungen stattfinden (z. B. für den Beruf des Kfz-Mechatronikers: Geschwendtner, 2010). In Bezug hierzu ergibt sich als zweite Forschungsfrage:

(2) Ist die Struktur der beruflichen Kompetenz abhängig vom Ausbildungsjahr der Lehrlinge?

Für die Erfassung der beruflichen Kompetenzen wurde ein Paper-Pencil-Test entwickelt. Die Testaufgaben spiegeln die typischen beruflichen Anforderungssituationen wider, welche im Rahmen der Pflichtqualifikationen der Ausbildung in Berufsschule und Betrieb thematisiert werden. Der Test wurde in Form eines Querschnittsdesigns Auszubildenden des 2. (N= 200), 3. (N= 150) und 4. Ausbildungsjahres (N= 80) vorgelegt. Es erfolgte keine Differenzierung des Tests für die Auszubildenden der verschiedenen Lehrjahre, da die im Rahmen der Ausbildung zu erwerbenden Pflichtqualifikationen i. d. R. mit der Beendigung des zweiten Ausbildungsjahres abgeschlossen sind und zudem oftmals bereits zuvor im Kontext der berufspraktischen Ausbildung erworben werden.

Die Struktur der beruflichen Fachkompetenz wird geprüft, indem ein- und mehrdimensionale IRT-Modelle mit Einfachstruktur (Between-Item-Mehrdimensionalität) bestimmt werden. Zusätzlich erfolgt die Spezifizierung von Zufallsmodellen, um das theoretisch abgeleitete Modell mit diesen anhand der Devianzmaße vergleichen zu können.

Auf Grundlage der Erkenntnisse in Bezug auf die Kompetenzstruktur sowie deren Entwicklung im Verlauf der Ausbildung können Konsequenzen hinsichtlich der inhaltlichen und organisatorischen Gestaltung der Berufsausbildung abgeleitet werden.

## **Unterrichtsstörungen im Spannungsfeld von Individuum, Klasse und Lehrperson - Eine multiperspektivische Betrachtung**

Chair(s): Päivi Taskinen (Friedrich-Schiller Universität Jena)

Diskutant/in: Martin Bonsen (Westfälische Wilhelms-Universität Münster)

Aus der Forschung zur effektiven Klassenführung ist bekannt, dass ein störungsarmer Unterricht eine Voraussetzung für effektive Lehr-Lern-Prozesse ist. Gleichzeitig berichten jedoch Lehrkräfte, dass störendes Verhalten von Schülerinnen und Schülern z. B. durch Nebengespräche, spielende Kinder oder Zwischenrufe im Unterricht relativ häufig ist. Ein störungsarmer Unterricht hat nicht nur auf das Lernen von Schülerinnen und Schülern Wirkungen. Die Forschung zu Belastung von Lehrkräften weist darauf hin, dass Unterrichtsstörungen mit einer hohen emotionalen Belastung von Lehrkräften einhergehen, welche ebenfalls die Unterrichtsqualität beeinträchtigen kann. Trotz dieser hohen Praxisrelevanz sind wenige konkrete, empirisch gesicherte Tatsachen über Wahrnehmungsunterschiede, Ursachen oder Folgen von gewöhnlichen Störungen im Unterricht bekannt. Das Ziel des Symposiums ist, Unterrichtsstörungen aus verschiedenen Perspektiven zu betrachten. Dazu stellt das Symposium vier aktuelle empirische Studien zur Unterrichtsstörung im Schulkontext vor. Der erste Beitrag nimmt die Wahrnehmung der Unterrichtsstörungen durch Lehrkräfte und Schülerinnen bzw. Schüler der Primarstufe in den Blick. Es wird überprüft, inwieweit die beiden Wahrnehmungsperspektiven in Bezug auf das Störausmaß und auf die Reaktion der Lehrkräfte darauf übereinstimmen. Der zweite Beitrag beschäftigt sich mit der Kompetenzentwicklung von Schülerinnen und Schülern in der Sekundarstufe. Es wird untersucht, inwieweit störendes Verhalten von Schülerinnen und Schülern einen Effekt auf die Mathematikkompetenz eines Individuums bzw. der Klasse hat. Der dritte Beitrag fokussiert den Einfluss der Peers auf die individuelle Entwicklung von problematischem Verhalten in der Klasse. Es wird im Laufe des siebten Schuljahrs untersucht, welchen Effekt das wahrgenommene Verhalten der Mitschülerinnen und Mitschüler auf das Problemverhalten eines Individuums hat. Der vierte Beitrag betrachtet undiszipliniertes Schülerverhalten als eine Störung des Schüler-Lehrer-Verhältnisses. Es wird in der gymnasialen Unterstufe untersucht, inwieweit die Qualität der Lehrer-Schüler-Beziehung durch Emotionen sowie Belastungserleben und bestimmte Verhaltensweisen der Lehrkräfte erklärt werden kann. Das Symposium kann zu Erkenntnissen aus der Forschung zum Classroom-Management und zum Belastungserleben von Lehrkräften anknüpfen. Der Fokus der Studien liegt auf Störungen im Rahmen des Unterrichts, wie sie im Schulalltag häufig beobachtet werden. Das Symposium ist multiperspektivisch angelegt, sodass der Blick auf die verschiedenen Akteure im Klassenzimmer gelenkt wird. Somit wird hier ein eher seltener, wenn auch ein sehr praxisrelevanter Zugang gewählt. Insgesamt vereint das Symposium Erkenntnisse aus der Erziehungswissenschaft und Psychologie und kann zu einer Synthese von Erkenntnissen beitragen.

- Abel, M. H. & Sewell, J. (1999). Stress and burnout in rural and urban secondary school teachers. *The Journal of Educational Research*, 92 (5), 287-293.
- Baumeister, R. F. & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117 (3), 497-529.
- Breidenstein, G. (2010). Überlegungen zu einer Theorie des Unterrichts. *Zeitschrift für Pädagogik*, 56, 869-887.
- Brophy, J. E. & Good, T. L. (1976). Die Lehrer-Schüler-Interaktion. München: Urban & Schwarzenberg.
- Demetriou, H., Wilson, E. & Winterbottom, M. (2009). The role of emotion in teaching: are there differences between male and female newly qualified teachers' approaches to teaching? *Educational Studies*, 35 (4), 449-473.
- Dishion, T. & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychiatry*, 62, 189-214.
- Döpfner, M., et al (1997). Deutschsprachige Konsensus-Versionen der Child Behavior Checklist (CBCL 4 - 18), der Teacher Report Form (TRF) und der Youth Self Report Form (YSR). *Kindheit und Entwicklung*, 6, 54-59
- Jackson, P. W. (1968). *Life in Classrooms*. New York: Holt, Rinehart & Winston.
- Hargreaves, A. (1998). The emotional practice of teaching. *Teaching and Teacher Education*, 14 (8), 835-854.
- Hargreaves, A. (2000). Mixed emotions: teachers' perceptions of their interactions with students. *Teaching and Teacher Education*, 16 (8), 811-826.
- Herzog, W. (2002). Zeitgemäße Erziehung. Die Konstruktion pädagogischer Wirklichkeit. Weilerswist: Velbrück.
- Herzog, W. (2011). Was dem Lehren und Lernen zugrunde liegt. Ein Mehrebenenmodell des Unterrichts. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 146-160). Bad Heilbrunn: Klinkhardt.
- Ingersoll, R. (2001). Teacher turnover and teacher shortages: an organizational analysis. *American Educational Research Journal*, 38 (3), 499-534
- Kiesler, A. (1999). Kommunikation unter Anwesenden. Studien über Interaktionssysteme. Frankfurt a. M.: Suhrkamp.
- Kiesner, J., Poulin, F. & Nicotra, E. (2003). Peer relations across contexts. Individual-network homophily and network inclusion in and after school. *Child Development*, 74, 1328-1343.
- Krapp, A. & Hascher, T. (2011). Lehreremotionen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrberuf* (S. 511-526). Münster: Waxmann.
- Luhmann, N. (1984). *Soziale Systeme. Grundriss einer allgemeinen Theorie*. Frankfurt a. M.: Suhrkamp.
- Makarova, E., Schönabächler, M.-T. & Herzog, W. (2008). Klassenmanagement und kulturelle Heterogenität: Projektphase 1 (Forschungsbericht Nr. 33). Bern: Abteilung Pädagogische Psychologie.
- Makarova, E., Schönabächler, M.-T. & Herzog, W. (2010). Klassenmanagement und kulturelle Heterogenität: Ergebnisse 2b -- Stellungnahme der Lehrpersonen (Forschungsbericht Nr. 39). Bern: Abteilung Pädagogische Psychologie.
- Mehan, H. (1998). *The Study of Social Interaction in Educational Settings: Accomplishments and Unresolved Issues*. Human Development, 41, 245-269.
- Müller, C. M., Begert, T., Gmünder, L. & Huber, C. (2012). Die „Freiburger Selbst- und Peerauskunftskaalen – Schulisches Problemverhalten“ – Entwicklung und Evaluation eines Verfahrens zur Verlaufsmessung von unterrichtsbezogenen Verhaltensproblemen. *Empirische Sonderpädagogik*, 1, 3-21.
- Nolting, H.-P. (2002). *Störungen in der Schulklasse*. Weinheim: Beltz.
- Riley, P. (2011). *Attachment theory and the teacher-student relationship*. London & New York: Routledge.
- Schönabächler, M.-T., Makarova, E. & Herzog, W. (2009). Klassenmanagement und kulturelle Heterogenität: Projektphase 2 (Forschungsbericht Nr. 36). Bern: Abteilung Pädagogische Psychologie.
- Schutz, P. A. & Zembylas, M. (Eds.) (2009). *Advances in teacher emotion research. The impact on teacher's lives*. New York: Springer.
- Singer, J. D. & Willett, J. B. (2003). *Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence*. New York: Oxford University Press.
- Spilt, J. L., Koomen, H. M. & Thijs, J. T. (2011) Teacher wellbeing: the importance of teacher-student relationships, *Educational Psychological Review*, 23, 457-477.
- Sutton, R. E., Mudrey-Camino, R. & Knight, C. (2009). Teachers' emotion regulation and classroom management. *Theory into Practice*, 48 (2), 130-137.
- Thomas, D. E., Bierman, K.L. & the Conduct Problems Prevention Research Group (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology*, 18, 471-487.
- Thomas, D. T., Bierman, K. L., Powers, C. J. & the Conduct Problems Prevention Research Group. (2011). The influence of classroom aggression and classroom climate on aggressive-disruptive behavior. *Child Development*, 82, 751-757.
- Vanderstraeten, R. (2001). The School Class as an Interaction Order. *British Journal of Sociology of Education*, 22, 263-277.
- Woolfolk Hoy, A. & Weinstein, C. S. (2006). Student and Teacher Perspectives on Classroom Management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management. Research, Practice, and Contemporary Issues* (pp. 181-219). Mahwah, NJ: Lawrence Erlbaum.





## ***Unterrichtsstörungen - ein multiperspektivischer Zugang***

Das Verständnis vom Unterricht als Interaktionssystem (vgl. Breidenstein, 2010; Brophy & Good, 1976; Jackson, 1968; Mehan, 1998; Vanderstraeten, 2001), unterstreicht die Bedeutung der gegenseitigen Wahrnehmung von Unterrichtsstörungen für den Aufbau von sozialer Ordnung im Unterricht (vgl. Herzog, 2002, 2011; Kieserling, 1999, S. 110ff.; Luhmann, 1984; Vanderstraeten, 2001). Somit sind Unterrichtsstörungen nicht im Fehlverhalten einzelner Schülerinnen und Schüler zu verorten, sondern als Ergebnis einer unzulänglich funktionierenden Interaktionsbasis des Unterrichts zu verstehen. Dabei grenzen wir den Begriff der Unterrichtsstörung auf Störungen der sozialen Ordnung des Unterrichts ein. Unterrichtsstörungen können sich demgemäß in Form von Regelverletzungen, Schwatzen, Dazwischenrufen, Missachtung der Lehrerautorität, Rempelen etc. äußern, wobei oft auch Disziplinprobleme miteingeschlossen werden (vgl. Nolting, 2002).

Das Defizit der bisherigen Forschung, die entweder die Schüler- oder die Lehrerperspektive fokussierte (Woolfolk Hoy & Weinstein 2006, p. 206), wird im vorliegenden Beitrag durch einen multiperspektivischen Zugang, der beide Perspektiven einander gegenüberstellt, behoben. Dazu analysieren wir die Wahrnehmung von Unterrichtsstörungen aus der Sicht der Lehrpersonen sowie aus der Sicht der Schülerinnen und Schüler derselben Schulklassen. Die erste Fragestellung lautet daher, inwiefern sich die Wahrnehmung von Unterrichtsstörungen in Abhängigkeit vom Ausmaß der Störung zwischen der Lehrperson und ihren Schülerinnen und Schülern unterscheidet. Die zweite Fragestellung betrifft die Reaktion der Lehrkräfte auf Unterrichtsstörungen. Dabei untersuchen wir, wie weit die Wahrnehmung der Lehrerreaktionen auf Unterrichtsstörungen durch die Schülerinnen und Schüler mit der Selbstwahrnehmung der Lehrperson übereinstimmt.

In methodischer Hinsicht verfolgt die Studie einen multimethodischen Ansatz, der sowohl quantitative als auch qualitative Daten zur Analyse der Forschungsfragen bezieht. Die Datenbasis der Analysen stammt aus einem Forschungsprojekt, das in fünf deutschsprachigen Kantonen der Schweiz durchgeführt wurde und zwei Phasen umfasste. In der ersten (quantitativen) Projektphase wurden Schülerinnen und Schüler (N = 4394) sowie deren Lehrpersonen (N = 225) der 5. Primarstufe mittels eines standardisierten Fragebogens befragt (Makarova, Schönbächler, & Herzog, 2008). Für die zweite (qualitative) Forschungsphase wurden 24 Klassen nach typologischen Kriterien ausgewählt. In diesen mittlerweile 6. Klassen gaben Schülerinnen und Schüler in Einzel- und Gruppeninterviews (N = 192) vertiefende Auskünfte zum Unterricht (Schönbächler, Makarova, & Herzog, 2009). Eine auf den Interviewanalysen beruhende Ergebniszusammenstellung wurde den Lehrpersonen (N = 20) der in der zweiten Forschungsphase beteiligten Klassen zur Stellungnahme vorgelegt (Makarova, Schönbächler, & Herzog, 2010).

Im Hinblick auf die erste Fragestellung zeigen die Ergebnisse eine weitgehende Übereinstimmung von Lehrer- und Schülerperspektive bei der Wahrnehmung von Unterrichtsstörungen, und zwar sowohl in Klassen mit niedrigem als auch in solchen mit hohem Störausmaß. Unterschiede zwischen der Lehrer- und Schülerwahrnehmung zeigen sich jedoch bei der zweiten Fragestellung. Die Lehrkräfte der Klassen mit hohem Störausmaß beurteilen hierbei ihre Reaktion auf Unterrichtsstörungen weit vorteilhafter als ihre Schülerinnen und Schüler. Zudem legen unsere Ergebnisse eine semantische Instabilität des Begriffs Unterrichtsstörungen nahe, in dem sie eine unterschiedliche Bedeutung von Unterrichtsstörung in Klassen mit hohem und tiefem Störausmaß belegen. Insgesamt erweist sich die im vorliegenden Beitrag angewandte Multiperspektivität in der Erfassung von Unterrichtsstörungen als gewinnbringend, da dadurch die Komplexität der sozialen Situation des Unterrichts veranschaulicht werden kann.

Päivi Taskinen (Friedrich-Schiller Universität Jena), Friederike Zimmermann (Christian-Albrechts-Universität zu Kiel), Kerstin Schütte (IPN), Olaf Köller (IPN)

## ***Unterrichtsstörungen durch Schülerinnen und Schüler: Nachteile für die Kompetenzen des Individuums oder der gesamten Klasse?***

Die Literatur zum Class-Room-Management weist darauf hin, dass ein störungsarmer Unterricht eine Voraussetzung für effektive Lehr-Lern-Prozesse ist. Empirische Studien zeigen, dass eine effektive Klassenführung in einem Zusammenhang mit einem höheren Lernerfolg von Schülerinnen und Schülern steht (z. B. Helmke & Schrader, 2006). Ein Erklärungsansatz liegt auf der Hand: Je weniger Lehrkräfte Zeit für Störungen aufwenden, desto mehr Zeit hat die Klasse für das Lernen zur Verfügung. Vor dem Hintergrund der Forschung ist insgesamt zu erwarten, dass ein störungsarmer Unterricht mit einem höheren Kompetenzzuwachs der Schülerinnen und Schüler einhergeht.

Der vorliegende Beitrag beschäftigt sich mit häufig vorkommenden Formen von Unterrichtsstörungen wie Dazwischenreden oder Nebengespräche auf der Ebene des Individuums und der Klasse. Das Ziel des Beitrags ist es, die Bedeutung von Unterrichtsstörungen für das Individuum und für die Klasse zunächst zu konkretisieren. Danach erfolgt die empirische Überprüfung der Frage: Trägt störendes Verhalten im Unterricht zur Erklärung der Höhe der aktuellen und zukünftigen mathematischen Kompetenz auf der Individual- und Klassenebene bei?

Die Untersuchung wurde in Klassen aus Bremen und Bremerhaven durchgeführt (NSchüler = 1001; NKlassen = 44). Zum 1. Messzeitpunkt besuchten die Schülerinnen und Schüler die fünfte Klassenstufe. Das störende Verhalten der Schülerinnen und Schüler wurde durch die Lehrkräfte anhand von fünf Items aus der Child Behavior Checklist (Arbeitsgruppe Deutsche Child Behavior Checklist, 1993) eingeschätzt. Zur Kompetenzerfassung wurden standardisierte Mathematiktests zum 1. Messzeitpunkt und zwei Jahre später (2. Messzeitpunkt) eingesetzt.

Die Überprüfung der Fragestellung erfolgte mit Hilfe eines Mehrebenenstrukturgleichungsmodells in Mplus mit Unterrichtsstörung und Kompetenz der Schülerinnen und Schülern auf der Within-Ebene und der Klassen auf der Between-Ebene. Als Einflussvariablen der Störung und der mathematischen Kompetenz wurden das Geschlecht auf der Individualebene und die Schulform auf der Klassenebene berücksichtigt.

Die Ergebnisse zeigen zunächst bereits bekannte Zusammenhänge: Jungen zeigen mehr störendes Verhalten im Unterricht und das Ausmaß der Störung ist in Gymnasialklassen ist vergleichsweise niedrig. Auch ist die mathematische Kompetenz in t1 ein sehr starker Prädiktor für die mathematische Kompetenz in t2. Dabei weisen Jungen eine höhere mathematische Kompetenz auf und die mathematische Kompetenz ist in Gymnasialklassen deutlich höher ausgeprägt als in anderen Schulformen.

In Bezug auf die Fragestellung zeigt sich: Zwischen der Unterrichtsstörung und der mathematischen Kompetenz zeigen sich auf der Klassenebene starke Zusammenhänge, wenn die Kontrollvariablen nicht berücksichtigt werden. Wird das Geschlecht der Schülerinnen und Schüler auf der Individualebene berücksichtigt, bleiben die Effekte signifikant: Störendes Verhalten im Unterricht zu t1 geht mit einer leicht geringeren mathematischen Kompetenz sowohl zu t1, als auch zu t2 einher. Somit weisen die Ergebnisse darauf hin, dass störendes Verhalten im Unterricht mit Einbußen in der individuellen Leistungsentwicklung einhergeht. Dies ist plausibel, da störendes Verhalten im Unterricht mit anderen schulischen oder außerschulischen Problemen einhergehen kann.

Auf der Klassenebene zeigt sich unter der Kontrolle des Bildungsganges: Das Ausmaß der Störung in der fünften Klasse steht in keinem Zusammenhang zur mathematischen Kompetenz in der fünften oder siebten Klasse. Auf der Klassenebene ist der Bildungsgang ein weit wichtigerer Erklärungsfaktor für die mathematische Kompetenz einer Klasse ist, als das Ausmaß an Störung. Ein Erklärungsansatz könnte der individuelle, situationsangepasste Umgang der Lehrkräfte mit Unterrichtsstörungen sein.

Christoph Michael Müller (Universität Freiburg), Janine Fleischli (Universität Freiburg), Verena Hofmann (Universität Freiburg), Felix Studer (Universität Freiburg)

## ***Effekte der Schulklassenzusammensetzung auf die Entwicklung individuellen schulischen Problemverhaltens***

Die Gleichaltrigen haben in der frühen Jugend eine zentrale Bedeutung für die Verhaltensentwicklung (Dishion & Tipsord, 2011). So vollziehen sich zwischen befreundeten Peers soziale Lernprozesse, die zu einer gegenseitigen Beeinflussung führen können. Peereinfluss vollzieht sich aber nicht nur in selbst gewählten Freundschaften - auch die Klassenkamerad(inn)en bilden eine wichtige Ressource für sozialen Austausch und den Aufbau sozialer Netzwerke (Kiesner et al., 2003). Es stellt sich daher die Frage, inwiefern das Niveau an Verhaltensproblemen unter den Klassenkamerad(inn)en einen Einfluss auf die individuelle Verhaltensentwicklung von Jugendlichen hat.

Dies wurde in der vorliegenden Studie hinsichtlich schulischen Problemverhaltens untersucht, welches hier Normverletzungen von Schüler(inne)n während der Interaktion mit der Lehrkraft im Klassenzimmer meint (z.B. unerlaubtes Aufstehen, Herumwerfen von Dingen, freche Bemerkungen etc.). Erste Ergebnisse auf Grundschulebene zum Einfluss der Klassenzusammensetzung auf die individuelle Entwicklung aggressiven Verhaltens (z.B. Thomas et al., 2006; 2011) ließen dabei erwarten, dass sich solche Effekte auch auf der Sekundarstufe I und in Bezug auf schulisches Problemverhalten nachweisen lassen.

Anhand einer Längsschnittstudie mit vier Messzeitpunkten während des 7. Schuljahrs (erstes Schuljahr auf der Sekundarstufe I) wurde überprüft, ob das Niveau an Problemverhalten unter den Peers zu Beginn des Schuljahrs ein Prädiktor für die weitere individuelle Verhaltensentwicklung ist. Weiter wurde bestimmt, ob sich Jugendliche auch an das sich über das Schuljahr hinweg stetig ändernde Niveau an Verhaltensproblemen unter den Klassenkamerad(inn)en anpassen.

### **Methode**

An der Studie nahm ein fast kompletter 7. Schuljahrgang von 825 Schülerinnen und Schülern aus allen acht Schulzentren einer spezifischen Region in der Schweiz teil. Die Jugendlichen bearbeiteten anonyme Selbst- und Peerauskünfte zu schulischem Problemverhalten (Müller et al., 2012). Dabei berichteten sie jeweils, an wie vielen der letzten 14 Tage sie selbst spezifische Problemverhaltensweisen gezeigt haben und wie viele ihrer Klassenkamerad(inn)en mindestens einmal in den letzten zwei Wochen dieses Verhalten an den Tag legten. Weiter wurden Konstrukte erfasst, die als Kontrollvariablen dienten (z.B. Geschlecht, SES, Migrationshintergrund, Schulnote, Impulshemmung, Verhalten der klassenexternen Freunde, Elternverhalten, Fürsorglichkeit der Lehrkraft, Bildungsgang). Die statistischen Analysen basierten auf Längsschnittmodellen, in denen alle Messzeitpunkte gleichzeitig einbezogen und die negativ binomiale Verteilung der Abhängigen Variable sowie die Mehrebenenstruktur der Daten berücksichtigt wurden (Ebene 1: Zeit; 2: Individuen; 3: Klasse; s. Singer & Willett, 2003).

### **Ergebnisse**

Schulisches Problemverhalten nahm im Verlauf der 7. Klasse insgesamt signifikant zu. Dabei zeigte sich, dass je höher der wahrgenommene Anteil an Klassenkamerad(inn)en mit schulischem Problemverhalten zu Beginn des Schuljahres war, desto mehr solches Verhalten wurde im weiteren Verlauf des Jahres individuell gezeigt (im Längsschnittmodell von t1-t4 sagte das Peerverhalten zu t1 individuelles Verhalten zu t2, t3 und t4 voraus). Weiter wurde der Entwicklungsverlauf auch durch das sich ständig ändernde Niveau an Verhaltensproblemen der Peers zum jeweils vorherigen Messzeitpunkt beeinflusst (im Längsschnittmodell von t1-t4 sagte das Peerverhalten zu t1 individuelles Verhalten zu t2 voraus und entsprechend t2->t3 und t3->t4). Weiterführend zeigten drei einzelne Kovarianzanalysen (t1->t2, t2->t3, t3->t4), bei denen das individuelle Verhalten zum jeweils nächsten Messzeitpunkt durch das Peerverhalten zum vorherigen Zeitpunkt vorhergesagt wurde (unter Kontrolle des individuellen Ausgangsverhaltens), dass der Effekt der Klassenzusammensetzung über alle Phasen des Schuljahrs hinweg stabil blieb.

### **Diskussion**

Schüler(innen) scheinen sich in ihrem eigenen Verhalten sowohl an der Ausprägung von schulischem Problemverhalten unter ihren Klassenkamerad(inn)en direkt nach dem Übertritt in die Sekundarstufe I als auch an eher kurzfristigen Änderungen des Peerverhaltens über das Schuljahr hinweg zu orientieren. Dieser Befund zeigt die entwicklungsrelevante Bedeutung von Prozessen der Zuteilung von Jugendlichen zu bestimmten Schulklassen auf und weist auf die Wichtigkeit des Klassenpeerkontextes bei der Planung von Präventionsmaßnahmen gegen Verhaltensprobleme hin. Angesichts der bekannten Unterschiede im Niveau an Verhaltensproblemen zwischen den verschiedenen Bildungsgängen können die Ergebnisse zudem bezüglich möglicher Effekte der Schulstruktur auf die Verhaltensentwicklung diskutiert werden.

## **Unterschiede im emotionalen Belastungserleben und im Unterrichtsverhalten als ein Resultat problematischer LehrerInnen-SchülerInnen-Beziehungen? Ergebnisse einer empirischen Untersuchung mit österreichischen Gymnasiallehrkräften**

Während die Forschungstradition zu Zufriedenheit, Stresserleben und Burn-Out im Lehrberuf eine relativ lange Geschichte aufweist, rücken Emotionen von Lehrkräften erst seit einigen Jahren – ausgelöst unter anderem durch den im Jahre 1998 veröffentlichten Artikel von Andy Hargreaves „The emotional practice of teaching“ – zunehmend in das Blickfeld des wissenschaftlichen Forschungsinteresses (z.B. Demetriou & Winterbottom, 2009; Hargreaves, 2000; Sutton, Mudrey-Camino, & Knight, 2009; Riley, 2011; im Überblick Krapp & Hascher, 2011 sowie Schutz & Zembylas, 2009). Studien zu Emotionen von Lehrkräften sind insofern als relevant einzustufen, da sie zum einen das Verhalten von Lehrkräften und somit die Qualität von Unterricht und Bildungsprozessen steuern; zum anderen beeinflussen sie die Lehrperson als Individuum, z. B. deren Wohlbefinden und stellen daher auch einen wesentlichen Prädiktor der LehrerInnengesundheits dar.

Die Quellen des Emotionserlebens von Lehrkräften sind vielfältig. Am häufigsten werden LehrerInnenemotionen jedoch von den Schülern und Schülerinnen, also durch LehrerInnen-SchülerInnen-Interaktionen, ausgelöst. Aus der Stressforschung ist bekannt, dass die Hauptursache von Lehrpersonenstress im undisziplinierten SchülerInnenverhalten liegt (Abel & Sewell, 1999). Zudem konnte Ingersoll (2001) belegen, dass Disziplinprobleme ein wesentlicher Prädiktor für den Berufsausstieg von Lehrpersonen sind. Undiszipliniertes SchülerInnenverhalten kann wiederum als eine Störung in der LehrerInnen-SchülerInnen-Beziehung betrachtet werden. Geht man nun davon aus, dass jeder Mensch von einem Grundbedürfnis nach positiven zwischenmenschlichen Beziehungen geleitet ist (Baumeister & Leary, 1995), so kann vermutet werden, dass Problematiken im Aufbau von positiven Beziehungen mit unangenehmen emotionalen Konsequenzen für die Lehrperson einhergehen (Spilt et al., 2011), wodurch in Folge auch der Unterricht negativ beeinflusst wird.

Mit dieser Thematik beschäftigt sich der vorliegende Beitrag: Es wird der Frage nachgegangen, inwieweit die Qualität der LehrerInnen-SchülerInnen-Beziehung (UV) das emotionale Erleben von Lehrkräften (operationalisiert durch: Freude/Angst/Ärger; IV), ihr Belastungserleben und Verhalten im Unterricht (operationalisiert durch: Belastung, Engagement hinsichtlich Unterrichtsgestaltung, SchülerInnenorientierung und Autonomiegewährung; AV) beeinflusst. Die LehrerInnen-SchülerInnen-Beziehung wird konzeptionell als ein mehrdimensionales Konstrukt verstanden, das sich durch eine affektive und eine behaviorale Komponente auszeichnet. Von einer problematischen Beziehung wird in diesem Beitrag dann gesprochen, wenn diese (1) durch eine geringe Nähe/Vertrautheit und einem hohen Konfliktpotenzial geprägt ist (= affektive Dimension), und (2) wenn sich dieses Verhältnis durch ein geringes Engagement der SchülerInnen (z.B. in Form von Mitarbeit und Aufgabenerledigung) sowie durch hohe Unterrichtsstörung auszeichnet (= behaviorale Dimension).

Zur Überprüfung der Fragestellung wurde eine schriftliche Befragung gewählt, an der 132 Lehrkräfte aus elf Gymnasien aus Österreich teilnahmen (66% weiblich; durchschnittliche Berufserfahrung=20.5 Jahre). Die Lehrkräfte bezogen sich in ihrer Einschätzung auf den Unterricht in der gymnasialen Unterstufe (= Sekundarstufe 1; Klassenstufen 5-8), und fokussierten dabei jeweils auf eine Schulklasse, da die Qualität der LehrerInnen-SchülerInnen-Beziehung zwischen unterschiedlichen Klassen variiert. Faktorenanalysen (explorativ / konfirmatorisch) bestätigen die mehrdimensionale Struktur des LehrerInnen-SchülerInnen-Verhältnisses. Reliabilitätsanalysen ergeben für alle eingesetzten Konstrukte zufriedenstellende Skalen- und Itemkennwerte.

Die genannten Beziehungen wurden mit Hilfe eines Strukturgleichungsmodells in Mplus überprüft. Die Ergebnisse bestätigen den vermuteten engen Zusammenhang zwischen der Qualität der LehrerInnen-SchülerInnen-Beziehung und den LehrerInnenemotionen. So geht beispielsweise ein hohes Konfliktpotenzial mit hohem LehrerInnenärger oder aber auch mit hoher LehrerInnenangst einher. Emotionen wiederum beeinflussen das Belastungserleben der Lehrkräfte und ihr Engagement im Hinblick auf die Unterrichtsgestaltung, während die Zusammenhänge zum selbstberichteten Unterrichtsverhalten in der Klasse schwächer ausfallen (z.B. im Hinblick auf das Ausmaß an gewährter Autonomie).

Aus diesen Ergebnissen kann die Schlussfolgerung gezogen werden, dass eine positive LehrerInnen-SchülerInnen-Beziehung als eine Ressource für positives emotionales Erleben von Lehrpersonen und auch als Grundlage für eine positive Unterrichtsgestaltung betrachtet werden kann. Kompetenz im Aufbau von positiven Beziehungen sollte daher in der LehrerInnenausbildung und -weiterbildung gefördert werden. Hierzu sind zum einen Fähigkeiten in der Klassenführung – als Grundlage für die Gestaltung positiver Lehrer-Schüler-Interaktionen – unabdingbar, zum anderen sollte auch die emotionale Kompetenz von Lehrkräften (z.B. der Umgang mit negativen Emotionen) gezielt geschult werden.



Chair(s): Miriam Vock (Universität Potsdam), Franzis Preckel (Universität Trier)

Diskutant/in: Martin Brunner (Institut für Schulqualität der Länder Berlin und Brandenburg)

Wie hochbegabte und leistungsstarke Schülerinnen und Schüler in der Schule wirksam gefördert werden können, ist eine zunehmend wichtige Frage in der empirischen Bildungs- und Unterrichtsforschung. Zum einen zeigen die internationalen Schulleistungsvergleiche der letzten Jahre, dass die Leistungsstärkeren in unserem Schulsystem offenbar weniger gut gefördert werden als in anderen Ländern. Zum anderen kann eine dauerhafte und umfassende schulische Unterforderung für die Schülerinnen und Schüler sehr ungünstige Konsequenzen für ihre Motivation sowie ihre weitere Leistungs- und Persönlichkeitsentwicklung haben.

In diesem Symposium betrachten wir aus unterschiedlichen Blickwinkeln und mit unterschiedlichen methodischen Zugängen Voraussetzungen und Effekte schulischer Begabtenförderung. Die ersten drei Beiträge gehen der Frage der Entwicklung von Leistung und Motivation Hochbegabter in unterschiedlichen Fördersettings nach. Vortrag 1 fokussiert aus eher psychologischer Perspektive die Zusammenhänge zwischen Selbstkonzept und Leistung in speziellen Begabtenklassen am Gymnasium; hierzu werden längsschnittliche Befunde von N = 1924 Schüler/innen berichtet. In Vortrag 2 werden Prädiktoren und Korrelate von misslingender Leistungsentwicklung (Underachievement) im Grundschulalter identifiziert. Vortrag 3 berichtet Befunde aus zwei Studien (N > 40.000), welche mithilfe von Propensity Score Matching die Effekte des Überspringens einer Klassenstufe auf Leistung, Motivation, soziale Einbettung und Wohlbefinden der Schüler/innen analysieren. Der vierte Vortrag schließlich betrachtet aus einer eher erziehungswissenschaftlichen Perspektive eine wesentliche Gelingensbedingung für eine erfolgreiche Begabtenförderung durch Akzeleration: Hier werden die Erfahrungen, das Wissen und die Haltungen von N = 316 Lehrkräften im Hinblick auf das Überspringen einer Klassenstufe untersucht. Dabei wird deutlich, dass ein beachtlicher Anteil der Lehrkräfte mit diesem Förderansatz inzwischen durchaus vertraut ist und ihm gegenüber positiv eingestellt ist.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
IG-Farben-Haus  
Raum 251





## ***Selbstkonzept und Leistung in gymnasialen Begabten- vs. Regelklassen***

### 1 Theoretischer Hintergrund

Das Fach Mathematik scheint für Schülerinnen und Schüler, die hochbegabt sind bzw. spezielle Begabtenklassen besuchen, einen besonderen Stellenwert zu haben. So berichten hochbegabte Schülerinnen und Schüler häufig ein höheres Interesse am Fach Mathematik als durchschnittlich begabte Schülerinnen und Schüler, während sich für das Fach Deutsch oft keine Unterschiede finden (Meier, Vogl, & Preckel, in Vorb.; Pruisken & Rost, 2005; Vlahovic-Stetic, Vidovic, & Arambasic, 1999). Weiterhin zeigen Schülerinnen und Schüler in Begabtenklassen im Fach Deutsch eine – im Vergleich zum Fach Mathematik – eher negative Noten- und auch Selbstkonzeptentwicklung (Schneider, Stumpf, Preckel, & Ziegler, 2012). Das Reciprocal Internal/External Frame of Reference Model (RI/E Model; Möller, Retelsdorf, Köller, & Marsh, 2011) bildet die Leistungs- und Selbstkonzeptentwicklung von Schülerinnen und Schülern im Längsschnitt ab. Konkret beschreibt das RI/E Model den funktionalen Zusammenhang zwischen mathematischen und verbalen Selbstkonzepten und den erbrachten Schulleistungen in Mathematik und einer verbalen Domäne (z. B. Deutsch). Hierbei rücken sowohl positive reziproke Effekte innerhalb einer Domäne (z. B. zwischen mathematischem Selbstkonzept und nachfolgender Mathematikleistung) als auch negative Kontrasteffekte zwischen Domänen (z. B. zwischen mathematischem Selbstkonzept und nachfolgender Deutschleistung) in den Fokus (Möller et al., 2011; Niepel, Preckel, & Brunner, 2013).

### 2 Fragestellung

Die vorliegende Studie untersucht das RI/E Model vergleichend für Schülerinnen und Schüler in regulären Gymnasial- und in speziellen gymnasialen Begabtenklassen. Ziel der Studie ist es zu untersuchen, ob sich die funktionalen längsschnittlichen Zusammenhänge zwischen Selbstkonzepten und Leistungen in den beiden Klassentypen systematisch unterscheiden. Finden sich in Begabtenklassen zum Beispiel stärkere negative Kontrasteffekte der Mathematik- auf die Deutschdomäne als in Regelklassen?

### 3 Methode

Stichprobe und Vorgehen: Verglichen werden zwei Stichproben: N = 1529 Schülerinnen und Schüler in regulären Gymnasialklassen und N = 395 Schülerinnen und Schüler in speziellen gymnasialen Begabtenklassen. Die insgesamt 1924 Schülerinnen und Schüler wurden über vier Messzeitpunkte hinweg untersucht: kurz nach dem Eintritt in das Gymnasium in der 5. Klasse (Messzeitpunkt 1), nach dem ersten Zwischenzeugnis in der 5. Klasse (Messzeitpunkt 2), nach dem Zwischenzeugnis in der 6. Klasse (Messzeitpunkt 3) und nach dem Zwischenzeugnis in der 8. Klasse (Messzeitpunkt 4). Das mittlere Alter der Schülerinnen und Schüler in den regulären Klassen betrug zwischen 10.69 Jahren (SD = 0.44; Messzeitpunkt 1) und 14.04 Jahren (SD = 0.46; Messzeitpunkt 4), in den Begabtenklassen zwischen 10.28 Jahren (SD = .61; Messzeitpunkt 1) und 13.71 Jahren (SD = .65; Messzeitpunkt 4).

Material: Das akademische Selbstkonzept in Mathematik und Deutsch wurde anhand von sechs Items erfasst, welche auf dem Self-Description Questionnaire (SDQ II; Kunter et al., 2002) basieren. Die Mathematik- und Deutschleistung der Schülerinnen und Schüler wurde anhand von selbstberichteten Halbjahresnoten erfasst.

### 4 Ergebnisse

Die Ergebnisse bestätigen die Vorhersagen des RI/E Modells für beide Klassentypen: Sowohl in den Regel- als auch in den Begabtenklassen fanden sich positive reziproke Effekte zwischen akademischem Selbstkonzept und nachfolgender Leistung, ebenso zwischen Leistung und nachfolgendem Selbstkonzept innerhalb einer Domäne. Negative Kontrasteffekte zeigten sich zwischen Leistungen in einer Domäne und nachfolgendem Selbstkonzept in der anderen und auch zwischen Selbstkonzept in einer Domäne und nachfolgender Leistung in der anderen. Weiterhin zeigten sich positive reziproke Effekte zwischen der Mathematik- und Deutschleistung sowie positive autoregressive Effekte.

### 5 Diskussion und Implikationen

Die funktionalen Zusammenhänge zwischen akademischen Selbstkonzepten und Leistungen unterschieden sich nicht – das RI/E Model kann über beide Subgruppen generalisiert werden. Leistungen und Selbstkonzepte scheinen vergleichbar relevant für die weitere Leistungs- und Selbstkonzeptentwicklung in regulären und in Begabtenklassen zu sein. Zur Erklärung unterschiedlicher Leistungs- und Selbstkonzeptentwicklungen in regulären und Begabtenklassen sollten daher verstärkt individuelle Trajektorien der Schülerinnen und Schüler als zentrale Analyseeinheit herangezogen werden.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., ... Weiß, M. (2002). PISA 2000: Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung (Band 72). Berlin: Max-Planck-Institut für Bildungsforschung.

Meier, E., Vogl, K. & Preckel, F. (in Vorb.). Who needs special advancement in school? Motivational characteristics of students in special gifted classes. Manuscript in preparation.

Möller, J., Retelsdorf, J., Köller, O. & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, 48, 1315–1346.

Niepel, C., Brunner, M. & Preckel, F. (2013). The longitudinal interplay of students' academic self-concepts and achievements: The reciprocal internal/external frame of reference model under scrutiny. Manuscript submitted for publication.

Pruisken, C. & Rost, D. H. (2005). Hochintelligent und besonders interessiert? Über Interessen und Hobbys hochbegabter Grundschul Kinder. *Psychologie in Erziehung und Unterricht*, 52, 100–110.

Vlahovic-Stetic, V., Vidovic, V. V. & Arambasic, L. (1999). Motivational characteristics in mathematical achievement: a study of gifted high-achieving, gifted underachieving and non-gifted pupils. *High Ability Studies*, 10(1), 37-49.

## Was unterscheidet überdurchschnittlich begabte von durchschnittlich begabten Underachievern?

1 Theoretischer Hintergrund: Underachievement ist definiert als Diskrepanz zwischen erwartbarer und tatsächlicher Leistung. Je nachdem, ob absolute, relative oder regressionsbasierte Operationalisierungen eingesetzt werden, ergeben sich unterschiedliche Prävalenzraten (Cheung & Rudowicz, 2003). Underachievement wird überwiegend im Kontext hoher Begabung beforscht – wohl auch deshalb, weil besonders Begabte eher mit Unterforderung konfrontiert sind. Dass Potenziale nicht entfaltet werden, hat in einer Leistungsgesellschaft negative Konsequenzen: Diskutiert werden neben zahlreichen anderen Merkmalen ein geringes (akademisches) Selbstkonzept und die Neigung zu internalisierenden und externalisierenden Problemen (vgl. Reis & McCoach, 2000). Hinzu kommt die Schwierigkeit der Identifikation: Lehrkräfte, die eher Leistung als Potenzial diagnostizieren können, erkennen (hochbegabte) Underachiever in der Regel nicht (Rost & Hanses, 1997).

2 Fragestellung: Da sich die meisten Studien zu Underachievement mit Hochbegabten befassen, ist unklar, inwieweit Prädiktoren und Korrelate generalisierbar sind. Unsere Studie verbindet verschiedene Ansätze (variablen- vs. personenzentriert), Operationalisierungen (Regressionsresiduen vs. direkte Ratings) und Perspektiven (Sicht der Lehrkräfte und der Eltern) des Phänomens. Überdurchschnittlich begabte werden mit durchschnittlich begabten Underachievern verglichen; eine Parallelisierung lege artis soll nach Abschluss der Datenerhebung erfolgen. Ferner werden sowohl die Regressionsresiduen als auch die Lehrer- und Elterneinschätzungen von Underachievement mittels verschiedener Prädiktoren vorhergesagt.

3 Methode: Stichprobe: Die Daten stammen aus der Normierungsstichprobe des Tests für (hoch-)intelligente Kinder T(H)INK; die Erhebungen werden Anfang 2014 abgeschlossen sein. Die bisherigen Ergebnisse basieren auf 280 Drittklässlern aus vier Bundesländern (53 % Mädchen, 14 % nicht deutsche Muttersprachler, Alter: M[SD] = 107,2 [6,05] Monate), für die IQ-Werte und Hauptfachnoten vorliegen. Für die Klassenstufen 2–4 werden gut 1500 Datensätze erwartet.

Material: Die Kinder bearbeiteten den T(H)INK (Baudson & Preckel, 2013), einen Gruppenintelligenztest für Grundschul Kinder ab Klasse 1, der verbales, numerisches und figural-bildhaftes Reasoning erfasst, eine adaptierte Kurzfassung des Fragebogens zur Erfassung emotionaler und sozialer Schulerfahrungen (FEES-K; Baudson & Preckel, 2011) und einen Fragebogen zu Need for Cognition (Preckel & Strobel, in Vorb.; Baudson, Strobel, & Preckel, 2012). Lehrkräfte und Eltern schätzten kognitive Fähigkeiten, Kreativität, Motivation, Sozial- und Arbeitsverhalten sowie Underachievement der Kinder ein. Die Eltern gaben außerdem Informationen zu Noten, Schlafverhalten und familiärer Situation ihres Kindes.

Operationalisierungen von Underachievement: Die studentisierten Residuen der Regression von durchschnittlicher Leistung in den Hauptfächern auf den IQ stellen ein kontinuierliches Maß für Under(bzw. Over)achievement dar. Die direkte Einschätzung durch Lehrkräfte und Eltern erfolgte über eine reliable 3-Item-Skala (Cronbach's  $\alpha$  [Lehrer] = .91, Cronbach's  $\alpha$  [Eltern] = .82).

4 Vorläufige Ergebnisse: Prävalenz und Gruppenunterschiede: 23 der 39 überdurchschnittlich begabten Kinder (IQ > 115) zeigen negative Regressionsresiduen, fünf davon mehr als eine Standardabweichung unter dem Mittelwert. Für die Vergleiche im nächsten Abschnitt wurden diesen 23 Kindern 80 durchschnittlich begabte Underachiever (85  $\leq$  IQ < 115) gegenübergestellt. Erstere haben ein signifikant höheres Selbstkonzept im Rechnen und Lesen; ihre Lehrkräfte schätzen ihre kognitive Fähigkeit, Kreativität, Motivation und Arbeitsverhalten höher ein, ihr Underachievement jedoch als geringer (alle Vergleiche außer Selbstkonzept Rechnen [ $p = .011$ ] auf 1 %- oder 0,1 %-Niveau signifikant). Ausmaß des Underachievements und Einschätzung der sozio-emotionalen Schulsituation sind jeweils vergleichbar.

Vorhersage von Underachievement: Residuen der Regression von Notenschnitt auf Intelligenz (Modell 1) sowie Lehrer- und Elterneinschätzung von Underachievement (Modell 2 und 3) wurden für die Gesamtgruppe mittels hierarchischer Regression vorhergesagt (1. Geschlecht des Kindes, 2. sozioökonomische Variablen, 3. Lehrer- bzw. Elterneinschätzungen von Motivation, Arbeitsverhalten, Kreativität und Sozialverhalten). Lehrerratings zu kognitive Fähigkeiten, Kreativität, Motivation, Sozial- und Arbeitsverhalten klärten 22,9 % der Residualvarianz und 23,2 % der Underachievement-Einschätzung auf. Geringere Motivationseinschätzungen sagten stärkeres Underachievement vorher ( $\beta = -.37$  bzw.  $-.42$ ); in Modell 2 waren außerdem Arbeitsverhalten ( $\beta = -.23$ ) und Kreativität ( $\beta = .20$ ) bedeutsam. Für Eltern spielt das Arbeitsverhalten ( $\beta = -.45$ ) eine stärkere Rolle als die Motivation ( $\beta = -.15$ ).

Baudson, T. G. & Preckel, F. (2011). Validierung einer modifizierten Kurzfassung des Fragebogens zur Erfassung emotionaler und sozialer Schulerfahrungen für die Klassen 1 bis 3. Vortrag bei der 11. Arbeitstagung der Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik (26.–28. September 2011), Saarbrücken.

Baudson, T. G. & Preckel, F. (2013). Development and validation of the German Test for (Highly) Intelligent Kids – T(H)INK. *European Journal of Psychological Assessment*, 29, 171–181.

Baudson, T. G., Strobel, A. & Preckel, F. (2012). Validierung einer neuen Need for Cognition (NFC)-Skala für Grundschulinnen und Grundschüler: Struktur, Messinvarianz und Zusammenhänge mit Intelligenz- und Leistungsmaßen. Vortrag beim 48. Kongress der Deutschen Gesellschaft für Psychologie (23.–27. September 2012), Bielefeld.

Cheung, C.-K. & Rudowicz, E. (2003). Underachievement and attributions among students attending schools stratified by student ability. *Social Psychology of Education*, 6, 303–323.

Preckel, F., Holling, H. & Vock, M. (2006). Academic underachievement: relationship with cognitive motivation, achievement motivation, and conscientiousness. *Psychology in the Schools*, 43, 401–411.

Preckel, F. & Strobel, A. (in Vorb.). Skala zur Erfassung kognitiver Motivation bei Grundschulkindern: Der Grundschul-NFC. Göttingen: Hogrefe.

Reis, S. M. & McCoach, D. B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly*, 44, 152–170.

Rost, D. H. & Hanses, P. (1997). Wer nichts leistet, ist nicht begabt? Zur Identifikation hochbegabter Underachiever durch Lehrkräfte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 29, 167–177.



## ***Auswirkungen des Überspringens einer Klassenstufe auf schulische Leistung, akademisches Selbstkonzept und sozio-emotionale Entwicklung***

1 Theoretischer Hintergrund: Das Überspringen einer Klassenstufe ist ein zentraler Ansatz der akzelerativen Begabtenförderung, bei dem eine verbesserte Passung zwischen Lernpotential und unterrichtlichen Lernanforderungen angestrebt wird (Lubinski & Benbow, 2000). Dabei soll keine künstliche Beschleunigung der natürlichen Entwicklung erzielt, sondern altersunabhängig das Maß an Herausforderung angestrebt werden, das den individuellen Voraussetzungen des Kindes und dessen Entwicklungsstand entspricht. Jedoch verlangt das kurzfristige selbstständige Aneignen des verpassten Unterrichtsstoffes ein hohes Maß an Leistungsfähigkeit und Motivation. Der Wechsel in eine neue Referenzgruppe mit Lern- und Reifevorsprung könnte das akademische Selbstkonzept der akzelerierten Schüler/innen beeinflussen (Marsh, Chessor, Craven, & Roche, 1995; Hooegeven, Van Hell, & Verhoeven, 2009). Zudem werden auch im sozio-emotionalen Bereich hohe Anforderungen an die Springerinnen und Springer gestellt: Die Eingliederung in einen neuen Klassenverband und der dauerhafte Altersunterschied zu den Klassenkameraden sowie der damit potentiell einhergehende „Exotenstatus“ erfordern hohe soziale Anpassungsleistungen.

2 Fragestellung: Im vorliegenden Beitrag wird deshalb der Fragestellung nachgegangen, ob Schülerinnen und Schüler, die eine Klassenstufe überspringen, in Folge dessen und im Vergleich mit nicht gesprungenen Schülerinnen und Schülern eine veränderte Leistungsentwicklung zeigen. Zudem wird die Frage nach den emotional-motivationalen und sozialen Auswirkungen des Überspringens einer Klassenstufe untersucht. Dabei wird geprüft, ob ein Big-Fish-Little-Pond-Effekt bei den gesprungenen Schülerinnen und Schülern festzustellen ist und es demnach zu Einbußen im akademischen Selbstkonzept kommt. Des Weiteren werden Effekte des Springens auf soziales Selbstkonzept, Schulzufriedenheit der Schüler/innen und auf das Schüler-Lehrer-Verhältnis analysiert.

3 Methode: Empirische Studien, die belastbare Aussagen über die Effekte des Überspringens treffen, sind bisher rar. Die vorhandenen Befunde unterliegen methodischen Einschränkungen: Neben kleinen Stichprobengrößen und fehlenden Vorher-Messungen besteht ein zentrales Problem oft im Fehlen adäquater Vergleichsgruppen (Vock, Preckel, & Holling, 2007). Im vorliegenden Projekt wird daher der Ansatz des Propensity Score Matchings gewählt (Stuart, 2010) und unter Rückgriff auf eine große, unausgelesene Stichprobe eine adäquate Vergleichsgruppe konstruiert. Dabei werden neben demografischen Variablen und zahlreichen familiären Hintergrundmerkmalen auch Maße kognitiver Grundfähigkeiten als Kontrollvariablen berücksichtigt. Als Datengrundlage stehen Fragebogen- und Testdaten von insgesamt über 40.000 Schülerinnen und Schülern aus Studien zur Erhebung der länderübergreifenden Bildungsstandards des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin zur Verfügung. Des Weiteren werden die Daten der ELEMENT-Erhebung (Lehmann & Lenkeit, 2008) zur Erfassung der Mathematik- und Leseleistung von etwa 4.000 Schülerinnen und Schülern an Berliner Grundschulen in den Jahren 2003 bis 2005 zur Re-Analyse genutzt. Diese Daten wurden vom Forschungsdatenzentrum des IQB zur Verfügung gestellt.

4 Ergebnisse: Hinsichtlich der Leistungsentwicklung akzelerierter Schülerinnen und Schüler bestätigen die Befunde das Bild, das bisherige empirische Arbeiten vorwiegend aufzeigen (Kulik, 2004; Park, Lubinski & Benbow, 2012; Steenbergen-Hu & Moon, 2011): Schülerinnen und Schüler, die eine Klassenstufe übersprungen haben, fallen in ihrer schulischen Leistung im Vergleich mit nicht gesprungenen Schülerinnen und Schülern der neuen Klassenstufe nicht ab. Signifikante Leistungsvorsprünge zeigen sich jedoch überwiegend im Vergleich mit den nicht akzelerierten Schülerinnen und Schülern der gleichen Alterskohorte, die in der ehemaligen Klassenstufe zurückbleiben. Hinsichtlich des Effekts der akzelerativen Maßnahme auf das akademische Selbstkonzept zeigt sich jedoch ein weniger einheitliches Bild: Ergebnisse anhand des Ländervergleichs zeigen überwiegend negative Effekte auf, während Analysen anhand der ELEMENT-Daten teilweise positive Effekte aufdecken konnten. Effekte des Überspringens auf Schulzufriedenheit und das Verhältnis zwischen Schüler und Lehrer konnten nicht gefunden werden. Lediglich in Bezug auf das soziale Selbstkonzept der gesprungenen Schülerinnen und Schüler ließen sich übereinstimmend signifikante negative Effekte kleiner bis mittlerer Höhe aufzeigen.

Hooegeven, L., van Hell, J. G. & Verhoeven, L. (2009). Self-concept and social status of accelerated and nonaccelerated students in the first 2 years of secondary school in the Netherlands. *Gifted Child Quarterly*, 53(1), 50–67.

Kulik, J. A. (2004). Meta-analytic studies of acceleration. In N. Colangelo, S. G. Assouline & M. U. M. Gross (Eds.), *A nation deceived: How schools hold back America's brightest students. The Templeton national report on acceleration* (pp. 13–22). Iowa City: University of Iowa.

Lehmann, R. H. & Lenkeit, J. (2008). Element. Erhebung zum Lese- und Mathematikverständnis. Entwicklung in den Jahrgangsstufen 4 bis 6 in Berlin: Abschlussbericht über die Untersuchungen 2003, 2004 und 2005 an Berliner Grundschulen und grundständigen Gymnasien. Berlin: Humboldt-Universität zu Berlin. [http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/element6\\_bericht\\_komplett.pdf](http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/element6_bericht_komplett.pdf)

Lubinski, D. & Benbow, C. P. (2000). States of excellence. *American Psychologist*, 55(1), 137–150.

Marsh, H. W., Chessor, D., Craven, R. & Roche, L. (1995). The effect of gifted and talented programs on academic self-concept: The big fish strikes again. *American Educational Research Journal*, 32(2), 285–319.

Park, G., Lubinski, D. & Benbow, C. P. (2012). When less is more: Effects of grade skipping on adult STEM productivity among mathematically precocious adolescents. *Journal of Educational Psychology*. Advance online publication.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21.

Steenbergen-Hu, S. & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly*, 55(1), 39–53.

Vock, M., Preckel, F. & Holling, H. (2007). Förderung Hochbegabter in der Schule—Evaluationsbefunde und Wirksamkeit von Maßnahmen. Göttingen: Hogrefe.



## ***Begabtenförderung aus Lehrersicht: Wissen, Erfahrung und Einstellungen zum Überspringen einer Klassenstufe***

1 Theoretischer Hintergrund: Hochbegabte Schülerinnen und Schüler sind den Gleichaltrigen in ihrer Klasse durch ihr erhöhtes Lerntempo oft weit voraus, so dass das Unterrichtsniveau nicht ihrem Potenzial und ihrem Lernbedürfnis entspricht. Eine Möglichkeit, diese Kinder angemessener zu fördern, besteht im Überspringen einer Klassenstufe. Empirische Studien zeigen, dass sich das Überspringen von Klassenstufen positiv auf die Leistungsentwicklung von hochbegabten Schülerinnen und Schülern auswirkt (Park, Lubinski, & Benbow, 2013; Kretschmann, Vock & Lüdtke, 2013) und zumindest keine negativen Konsequenzen für die sozioemotionale Entwicklung mit sich bringt (Hoogeveen, van Hell, & Verhoeven, 2011; Steenbergen-Hu & Moon, 2011). Zum Gelingen dieser Begabtenförderungsmaßnahme tragen die begleitenden Lehrkräfte maßgeblich bei. Offenbar stehen Lehrkräfte der Maßnahme jedoch ambivalent gegenüber, selbst wenn sie Begabtenförderung insgesamt positiv sehen (McCoach & Siegle, 2007; Siegle, Wilson, & Little, 2013). Andererseits gibt es Hinweise darauf, dass das Interesse deutscher Lehrkräfte für das Überspringen in den letzten Jahren gestiegen ist, da Springen mittlerweile in allen Bundesländern möglich ist (Heinbokel, 2009).

2 Fragestellung: Mit Blick auf die zunehmende Flexibilisierung des deutschen Schulsystems, das sich verstärkt an individuellen Entwicklungsverläufen orientiert (Heinbokel, 2009; Liebers, Prengel, & Bieber, 2008), untersucht dieser Beitrag, wie umfangreich der Wissensstand und die Erfahrungen der Lehrkräfte mit dem Unterrichten von Springer/innen sind und welche Einstellungen sie zum Überspringen einer Klassenstufe haben. Darüber hinaus betrachten wir die wechselseitigen Einflüsse von Kenntnissen, Erfahrungen und Einstellungen von Lehrkräften zum Überspringen einer Klassenstufe.

3 Methode: Im Rahmen der Panel Study at the Research School 'Education and Capabilities' in North Rhine-Westphalia (PARS; Bos et al., 2012) nahmen N = 316 Lehrkräfte im Frühjahr 2010 an einer Befragung teil, die auch Fragen zum Überspringen von Klassenstufen enthielt. Die Lehrkräfte berichteten, ob sie bereits Erfahrung mit Schüler/innen besitzen, die eine Klassenstufe übersprungen haben, wie sie diese Erfahrung beurteilen und ob sie bereits selbst zum Überspringen geraten haben (3 Items). Die Erfassung der Einstellungen und die Selbsteinschätzung des Wissensstandes erfolgte über reliable Skalen (Cronbach's  $\alpha = .86$ ; 4 Items bzw. Cronbach's  $\alpha = .89$ ; 10 Items). Weitere 17 Items erfragten, welche spezifischen Auswirkungen auf soziale, motivationale, emotionale und leistungsbezogene Aspekte des Schülererlebens und -verhaltens die Lehrkräfte vermuteten.

4 Ergebnisse: Mehr als 70 Prozent der Lehrkräfte in unserer Stichprobe haben bereits Erfahrung mit dem Unterrichten von Schüler/innen, die eine Klassenstufe übersprungen haben. Diese Erfahrungen werden überwiegend positiv bewertet. Außerdem sehen die Lehrkräfte das Überspringen als eher positive Interventionsmaßnahme. Explorative Faktorenanalysen zeigen, dass Lehrkräfte bei der Bewertung der Maßnahme zwischen sozialen, affektiv-motivationalen und leistungsbezogenen Auswirkungen des Springens differenzieren. Die Mehrzahl der Lehrkräfte schätzt ihr Wissen zum Überspringen als zu gering ein, um Eltern kompetent bei der Entscheidung beraten zu können. Dieser Befund ist unserer Ansicht nach alarmierend, da unsere regressionsanalytischen Ergebnisse auch darauf hinweisen, dass der Wissensstand statistisch signifikant mit den Einstellungen zum Überspringen assoziiert ist. Unsere Ergebnisse unterstreichen daher den Bedarf an Weiterbildungen zu dieser Form der Begabtenförderung.

Bos, W., Stubbe, T. C., Buddeberg, M., Dohe, C., Kasper, D., Müller, S. & Walzebug, A. (2012). Framework for the Panel Study at the Research School 'Education and Capabilities' in North Rhine-Westphalia (PARS). Manuskript in Vorbereitung.

Heinbokel, A. (2009). Handbuch Akzeleration. Münster: LIT Verlag.

Hoogeveen, L., van Hell, J. G. & Verhoeven, L. (2011). Social-emotional characteristics of gifted accelerated and non-accelerated students in the Netherlands. *British Journal of Educational Psychology*, 82, 585–605.

Kretschmann, J. Vock, M. & Lüdtke, O. (2013). Acceleration in elementary school: Using propensity score matching to estimate the effects on academic achievement. Manuscript submitted for publication.

Liebers, K., Prengel, A. & Bieber, G. (2008). Die flexible Schuleingangsphase. Evaluationen zur Neugestaltung des Anfangsunterrichts. Weinheim: Beltz.

McCoach, D. B. & Siegle, D. (2007). What predicts teachers' attitudes toward the gifted? *Gifted Child Quarterly*, 51, 246–255.

Park, G., Lubinski, D. & Benbow, C. P. (2013). When less is more: Effects of grade skipping on adult STEM productivity among mathematically precocious adolescents. *Journal of Educational Psychology*, 105, 176–198.

Siegle, D., Wilson, H. E. & Little, C. A. (2013). A sample of gifted and talented educators' attitudes about academic acceleration. *Journal of Advanced Academics*, 24, 27–51.

Steenbergen-Hu, S. & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly*, 55, 39–53.





## ***Pygmalion im Klassenzimmer: Erwartungseffekte von Lehrkräften auf die Mathematikleistungen ihrer Schüler***

**Chair(s):** Alena Friedrich (Eberhard Karls Universität Tübingen)

Die auch als Pygmalion benannten Erwartungseffekte von Lehrkräften können die akademische Leistung von Schülern beeinflussen (z.B. Cooper, 1979; Jussim & Eccles, 1995; Madon, Jussim, & Eccles, 1997). Die Erwartungen, die Lehrkräfte entwickeln, basieren auf dem Wissen über ihre Schüler durch vorherige Noten oder im Klassenzimmer gezeigte Leistung, aber auch auf Vorurteilen oder Stereotypen (Good, 1987; Jussim, Eccles, & Madon, 1996; Reyna, 2000, 2008). Viele empirische Studien haben die Vorhersagen des Pygmalioneffekts bestätigt, auch wenn die gefundenen Effektstärken eher klein ausfielen. Zur Erfassung der Schülerleistungen wurden in bisherigen Studien zu Erwartungseffekten meist ausschließlich Tests (z.B. Kuklinski & Weinstein, 2001) oder Noten (z.B. Tiedemann, 2000) verwendet, jedoch selten beide Maße simultan eingesetzt. Da Tests und Noten als Leistungsmaß jeweils spezifische Vor- und Nachteile besitzen, profitiert man vom simultanen Einsatz beider Maße und kann so deren Ergebnisse vergleichen. Zudem untersuchten die meisten Studien Erwartungseffekte von Lehrkräften nur auf der Schülerebene (d.h. verbessert sich ein einzelner Schüler?), aber nicht auf der Klassenebene. Es wurde also selten untersucht, ob die Leistung ganzer Klassen mit den Erwartungen einer Lehrkraft in Zusammenhang steht. Des Weiteren verwenden manche bisherige Studien nur relativ kleine Stichproben oder vernachlässigen die Mehrebenenstruktur in ihrer Auswertung.

Die vorliegende Studie untersucht den Pygmalioneffekt daher in einer Längsschnittstudie, anhand einer großen Stichprobe in normalen Schulklassen und differenziert zwischen zwei Leistungsmaßen (Noten und standardisierte Leistungstests) sowie der Schüler- und Klassenebene. Zudem wurde das Selbstkonzept der Schüler als möglicher Mediator zwischen Lehrererwartungseffekten und Leistung untersucht. Daten von 73 Lehrkräften und ihren 1,289 Schülern der fünften Klasse wurden analysiert. Die Studie erstreckte sich über eine Zeitdauer von sechs Monaten und bestand aus drei Messzeitpunkten (T1, T2 und T3). Die Lehrkräfte wurden zu T1 nach ihren Erwartungen über die Leistungen/Kompetenzen der Schüler befragt (Beispielitem: „Der Schüler kann selbst schwierige Matheaufgaben lösen“), das Schülerelbstkonzept wurde zu T2 erfasst (Beispielitem: „Ich kann selbst schwierige Matheaufgaben lösen“). Zu T3 wurden die Schülerleistungen (Note und Test) erhoben. Die mittleren Lehrererwartungen auf Klassenebene wurden durch Aggregation der individuellen Lehrereinschätzungen gebildet. Als Kontrollvariablen wurden das Geschlecht, das Alter, die kognitiven Fähigkeiten, die frühere Schulleistung und das frühere Selbstkonzept der Schüler berücksichtigt. Die untersuchten Forschungsfragen lauteten: 1) Sind Lehrererwartungen mit der Leistungsentwicklung von Schülern in Tests und Noten assoziiert? 2) Wird der Zusammenhang zwischen Lehrererwartungen und Leistung mediiert durch das Selbstkonzept der Schüler? 3) Sind Lehrererwartungen auf der Klassenebene über Erwartungseffekte auf der individuellen Ebene signifikant mit der Leistungsentwicklung von Schülern assoziiert? Zur Beantwortung der drei Forschungsfragen wurden Mehrebenen-Regressionsanalysen und Mehrebenen-Mediationsanalysen mit Mplus berechnet. Das Resultat: Erstens, Pygmalioneffekte ließen sich auf der individuellen Schülerebene für beide Leistungsmaße nachweisen ( $b = .11$  und  $.13$ ). Über die Effekte von Geschlecht, Alter, kognitiven Fähigkeiten und früheren Schulleistungen hinaus hängen die Erwartungen der Lehrkräfte demnach mit den Noten und dem Testergebnis der Schüler ein halbes Jahr später zusammen. Zweitens, Mehrebenen-Mediationsanalysen zeigten, dass Lehrererwartungseffekte teilweise durch das Selbstkonzept von Schülern mediiert wurden. Drittens, die durchschnittliche Lehrererwartung auf der Klassenebene zeigte sich nicht signifikant mit der Schulleistung assoziiert, wenn u.a. für die vorherige Schulleistung der Schüler kontrolliert wurde. Die drei Ergebnisse werden in Bezug auf ihre theoretischen und praktischen Implikationen für die Wirkung und Bedeutung von Erwartungen und der Lehrerbildung diskutiert.

Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research*, 49(3), 389-410. doi: 10.3102/00346543049003389

Good, T. L. (1987). Two decades of research on teacher expectations: Findings and future directions. *Journal of Teacher Education*, 38(4), 32-47. doi: 10.1177/002248718703800406

Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 28. (pp. 281-388). San Diego, CA: Academic Press.

Jussim, L., & Eccles, J. S. (1995). Naturally occurring interpersonal expectancies. In N. Eisenberg (Ed.), *Social development* (pp. 74-108). Thousand Oaks, CA: Sage Publications.

Kuklinski, M. R., & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development*, 72(5), 1554-1578. doi: 10.1111/1467-8624.00365

Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, 72(4), 791-809. doi: 10.1037/0022-3514.72.4.791

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review*, 12(1), 85-110. doi: 10.1023/A:1009037101170

Reyna, C. (2008). Ian is intelligent but Leshau is lazy: Antecedents and consequences of attributional stereotypes in the classroom. *European Journal of Psychology of Education*, 23(4), 439-458. doi: 10.1007/BF03172752

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1), 144-151. doi: 10.1037/0022-0663.92.1.144





## „Die Klasse bringt`s!“ – Kompositionseffekte der Schulklasse auf die Veränderung der akademischen Aspiration in der Sekundarstufe I

Die Kultivierung schulischer Aspirationen kann als ein wichtiger Bestandteil der pädagogischen Arbeit in Bildungsinstitutionen gesehen werden. Studien verweisen nicht zuletzt darauf, welche Bedeutung Aspirationen bei Bildungsentscheidungen und dem Erlangen eines Bildungsabschlusses zukommt (z.B. Palardy, 2013; Wiebke & Blossfeld, 2007). Mit dem postulierten Wandel von einer Industrie- zu einer Wissensgesellschaft (Europäischer Rat, 2000), bei gleichzeitiger augenscheinlicher geringer Studienbeginn- und Studienabschlussrate (OECD, 2013), kann auch vonseiten der Bildungspolitik die Förderung und der langfristige Erhalt von Bildungsaspiration als wünschenswert angesehen werden. Die – in Österreich, als Reaktion auf die ersten PISA Ergebnisse im Schuljahr eingeführte Schulform „Neue Mittelschule“ (neue Schulform der Sekundarstufe I – Schule der 10-14jährigen) setzt sich zum Ziel eine Schule für alle zu sein und Marginalisierungsprozesse zu minimalisieren. Dabei sollen vor allem auch die schulische Aspiration und Bildungsmotivation langfristig aufgebaut und erhalten bleiben. Vielfach zeigt sich aus der Forschung, dass, noch stärker als die individuelle Leistung, das akademische Selbstkonzept, also die persönliche Wahrnehmung der eigenen Fähigkeiten im fachspezifischen Kontext (Byrne & Shavelson, 1986), einen sehr einflussreichen Prädiktor darstellt, um Karriere- und Bildungsaspirationen vorherzusagen (Guay, Larose & Boivin, 2004; Marsh & Yeung, 1997). Dennoch muss davon ausgegangen werden, dass sich Aspirationen nicht nur aus individuellen Merkmalen der einzelnen SchülerInnen erklären lassen, sondern auch durch Faktoren der schulischen Zusammensetzung und des schulischen Kontextes. Im Rahmen des Big-fish-little-pond-Effekts (Marsh, 1987), lässt sich die negative Wirkung eines allgemein hohen Selbstkonzepts in der gleichaltrigen Bezugsgruppe auf die Bildungsaspiration einzelner SchülerInnen durch soziale Vergleichsprozesse erklären. Obwohl weiters kontextuelle Barrieren als bedeutende Einflussfaktoren auf Karrierewahl und Entwicklung gefunden wurden, legt die bisherige Forschung nur wenig Aufmerksamkeit auf kontextuelle Schul- und Peereffekte (Nagengast & Marsh, 2011). Dies erweist sich insofern interessant, als bereits im Rahmen des „Equality of Educational Opportunity Reports“ von Coleman und Kollegen (1966) eine Wirkung der sozio-ökonomischen Zusammensetzung der KlassenkollegInnen auf die schulische Leistung, die schulischen Aspirationen sowie auf das weitere Bildungsverhalten am Ende der High-School gezeigt werden konnte. Der Beitrag geht der Frage nach, wie sich die schulischen Aspirationen von niederösterreichischen SchülerInnen in dem inklusiven Setting Neue Mittelschule im Verlauf ihrer Schulkarriere in der Sekundarstufe I entwickeln und welche Einflussfaktoren für Veränderungen auszumachen sind. Dabei interessieren nicht nur intraindividuelle Einflussfaktoren, sondern auch welche Bedeutung Kompositionseffekten durch die Bezugsgruppe von KlassenkollegInnen zukommt. Als Analysegrundlage für diesen Beitrag dienen Daten der NOESIS-Längsschnitterhebung (Niederösterreichische Mittelschule in der Schulentwicklung, [www.noesis-projekt.at](http://www.noesis-projekt.at)), in der seit dem Schuljahr 2010/11 drei SchülerInnenkohorten von der 4. Klasse Volksschule bis zum Ende ihrer Pflichtschulzeit begleitet werden. In die Analyse fließen die ersten Erhebungszeitpunkte 2 bis 4 der ersten Kohorte ein (Schulstufe 5 bis 7, N=923). Die Ergebnisse zeigen, dass die akademische Aspiration der SchülerInnen in der neuen Mittelschule mit der Zeit steigt. Mit dieser Entwicklung parallel/in Verbindung, kann vor allem eine positive Veränderung der Bildungsmotivation und des Selbstkonzeptes im Unterrichtsgegenstand Deutsch, aber auch vermehrte kulturelle Aktivitäten außerhalb der Schule gesehen werden. Im Sinne eines Kompositionseffektes scheinen SchülerInnen, die eine Klasse besuchen, in welcher Motivation, Selbstkonzept und außerschulische Aktivitäten im zeitlichen Verlauf steigen, auch eine erhöhte schulische Aspiration aufzuweisen. Interessant zeigt sich im Rahmen der Panelanalyse zudem, dass diese Kompositionseffekte nicht allein auf die Klassenzusammensetzung der SchülerInnen rückführbar sind, sondern auch die subjektive Sichtweise des aktuellen Standes der eigenen schulischen Fähigkeiten eine signifikante Rolle einnimmt. Die Ergebnisse verweisen damit nicht zuletzt darauf, dass sowohl die Zusammensetzung der SchülerInnen, als auch der Vergleich mit der Peergruppe im Setting des Klassenzimmers einen nicht zu vernachlässigenden Stellenwert bei der Entwicklung von schulischen Aspirationen einnimmt. Im Rahmen der Neuen Mittelschule kann damit in der „Klasse“ und den Peers eine mögliche Ressource zum Aufbau und Erhalt von akademischer Aspiration gesehen werden.

- Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-concept. IN: Journal of Educational Psychology, Vol. 78, pp. 474–481.
- Coleman, J.S., Campbell, E. Q., Hobson, J. M., Mood, A. M., Weinfeld, F. D. (1966). Equality of educational Opportunity. Washington, D.C.: U.S. Department of Health, Education, and Welfare.
- Europäischer Rat (2000). Memorandum über Lebenslanges Lernen. Online abrufbar unter: <http://www.bologna-berlin2003.de/pdf/MemorandumDe.pdf> [28.10.2013].
- Guay, F., Larose, S., & Boivin, M. (2004). Academic self-concept and educational attainment level: A 10-year longitudinal study. IN: Self and Identity, Vol. 3, pp. 53–68.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. IN: Journal of Educational Psychology, Vol. 79, pp. 280–295.
- Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. IN: American Educational Research Journal, Vol. 34, pp. 691–720.
- Nagengast, Benjamin & Marsh, Herbert W. (2012). Big Fish in Little Ponds Aspire More: Mediation and Cross-Cultural Generalizability of School-Average Ability Effects on Self-Concepts and Career Aspirations in Science. IN: Journal of Educational Psychology, Vol. 104, No. 4, pp. 1033-1053.
- OECD (2013): Education at a Glance 2013. Online abrufbar unter: [http://www.oecd-ilibrary.org/education/education-at-a-glance-2013\\_eag-2013-en](http://www.oecd-ilibrary.org/education/education-at-a-glance-2013_eag-2013-en) [28.10.2013].
- Palardy, Gregory J. (2013). High School Socioeconomic Segregation and Student Attainment. IN: American Educational Research Journal, vol. 50, No.4, pp. 714-754.
- Wiebke, P. & Blossfeld, H.-P. (2007). Schichtspezifische Präferenzen oder sozioökonomisches Entscheidungskalkül? Zur Rolle elterlicher Bildungsaspirationen im Entscheidungsprozess beim Übergang von der Grundschule in die Sekundarstufe. IN: Zeitschrift für Pädagogik, Jg. 53, Nr. 4, S. 491-508.



**Die Effekte von Schulführung auf evidenzbasiertes Handeln von Lehrkräften**

Einleitung/Fragestellung: Das Konzept der Evidenzbasierung ist mit der Erwartung verknüpft, dass Lehrkräfte Daten bspw. aus Evaluationen im Rahmen von Entscheidungen nutzen. Studien zeigen jedoch, dass Lehrkräfte oftmals nur bedingt evidenzbasiert handeln (Zlatkin-Troitschanskaia, Seidel & Stump, 2013). In diesem Zusammenhang ist die Frage von Bedeutung wie evidenzbasiertes Handeln systematisch gefördert werden kann. Insbesondere gemäß Forschungsbefunden des englischsprachigen Raumes spielt hierbei die Schulleitung bzw. die Führung einer Schule eine maßgebliche Rolle (Levin & Datnow, 2012). Diese Forschungsergebnisse beruhen zumeist auf qualitativen Fallstudien mit einem entsprechend geringen Stichprobenumfang. Um jedoch präzise Aussagen treffen zu können, sind quantitative Studien mit größeren Stichproben notwendig. Hier setzt das Projekt ‚Evidenzbasiertes Handeln im schulischen Mehrebenensystem - Bedingungen, Prozesse und Wirkungen‘ (EviS) an. (Projektnummer: 01JG1010A; Laufzeit: September 2010 bis August 2016 im Rahmen des BMBF-Forschungsschwerpunkts Steuerung im Bildungssystem (SteBis); <http://www.wipaed.uni-mainz.de/evs/>). Der nachstehend skizzierte Vortrag zielt auf folgende Frage ab: Welche Effekte haben transformationale Führung, Förderung von Kooperation, Förderung von Partizipation in Entscheidungen und die systematische Bereitstellung von Daten seitens der Schulleitung auf evidenzbasiertes Handeln von Lehrkräften?

Theoretischer Hintergrund: Mit dem Ziel ein differenziertes theoretisches Modell zu entwickeln wurden für die unabhängigen Konstrukte wie transformationale Führung sowohl schulspezifische Literatur (Leithwood & Sun, 2012; Leithwood & Jantzi, 2005) als auch nicht-schulbezogene Literatur herangezogen (Bass, 1985; Bass & Riggio, 2006; Felfe, 2006). Darüber hinaus wurden theoretische Ansätze aus dem Bereich evidenzbasierten Handelns mit einem Fokus auf Führung berücksichtigt (Wayman, Spring, Lemke & Lehr, 2012; Cosner, 2012). Da das in der Studie abhängige Konstrukt evidenzbasiertes Handeln in der bisherigen Literatur relativ wenig präzisiert ist, erfolgt der Einbezug umfangreicher deutsch- und englischsprachiger Literatur. Dazu gehören theoretische Konzepte und empirische Befunde aus den Bereichen evidenzbasiertes Management (Rousseau, 2006; Yates & Potworowski, 2012), Entscheidungsfindung (LeBoeuf & Shafir, 2012; Kahnemann, Slovic & Tversky, 2005), evidenzbasierte Medizin (Satterfield et al., 2009; Virgilio, Chiapa & Palmarozzi, 2007), Wissenschaftstheorie (Mittelstraß, 2005) und empirische Bildungsforschung (Schildkamp & Lai, 2013; van Ackeren et al., 2011; Goodnow & Wayman, 2009; Huber, 2008).

Methode: Im Rahmen einer Querschnittserhebung im Jahr 2011 wurden standardisierte Fragebögen an 153 allgemein- und berufsbildenden Schulen in Rheinland-Pfalz eingesetzt. Ergebnis sind Daten von 2.640 Lehrkräften sowie 297 (stellvertretenden) SchulleiterInnen. Die hierbei verwendeten Messinstrumente wurden zuvor mehrfach validiert. Die Förderung evidenzbasierten Handelns durch die Schulleitung wurde mittels vier Skalen gemessen: Transformational Führung, Unterstützung von Kooperation unter Lehrkräften, Unterstützung von Partizipation in Entscheidungen von Lehrkräften und Bereitstellung von Daten (unabhängige Konstrukte). Die dazugehörigen Skalen entstammen dem Multifactor Leadership Questionnaire von Felfe (2006) sowie von Bass und Avolio (1995) und dem Evidence-Based Management Questionnaire von Stumm, Dormann und Mohr (2010). Die Skala evidenzbasiertes Handeln (abhängiges Konstrukt) wurde ebenfalls von Stumm, Dormann und Mohr (2010) entwickelt. Die statistischen Analysen erfolgen durch den Einsatz von SPSS und MPLus.

Ergebnisse: Die eingesetzten Skalen wurden faktor- sowie reliabilitätsanalytisch geprüft und ihre weitgehend hohe Güte konnte bestätigt werden (Cronbachs Alpha > 0,7). Es können sehr interessante Befunde gezeigt werden: Die vier Skalen zum Führungshandeln der Schulleitung können gemeinsam 50% der Varianz der Skala Evidenzbasiertes Handeln der Lehrkräfte erklären. Die Skala Transformationale Führung hat dabei den größten Anteil an dieser Varianzaufklärung. Insgesamt können die Ergebnisse als wichtiges Indiz dafür gewertet werden, dass das Führungshandeln der Schulleitung einen bedeutsamen Einfluss auf das evidenzbasierte Handeln von Lehrkräften hat. Im Rahmen des Vortrags werden die ermittelten Ergebnisse sowie weitere Nebenbefunde insbesondere mit Blick auf ihre internationale Vergleichbarkeit hin kritisch diskutiert. Ergänzend wird erörtert, welche Implikationen diese Ergebnisse in Bezug auf die Verbesserung der Aus- und Weiterbildung von SchulleiterInnen haben. Weiterhin werden auf Grundlage der Forschungsbefunde Implikationen für die weitere Forschung diskutiert.

- Ackeren, I. van; Zlatkin-Troitschanskaia, O.; Binnewies, C.; Clausen, M.; Dormann, C.; Preisendörfer, P. et al. (2011). Evidenzbasierte Schulentwicklung. Ein Forschungsüberblick aus interdisziplinärer Perspektive. In: Die deutsche Schule 103 (2), S. 170–184.
- Bass, B. M. (1985). Leadership and performance beyond expectations. New York: Free Press.
- Bass, B. M.; Avolio, B. (1995). MLQ Multifactor Leadership Questionnaire: Technical Report. Redwood City, CA: Mind Garden.
- Bass, B. M.; Riggio, R. E. (2006). Transformational leadership. 2. Aufl. New York: Taylor & Francis Group.
- Cosner, S. (2012). Leading the Ongoing Development of Collaborative Data Practices: Advancing a Schema for Diagnosis and Intervention. In: Leadership and Policy in Schools 11 (1), S. 26–65.
- Felfe, J. (2006). Validierung einer deutschen Version des "Multifactor Leadership Questionnaire" (MLQ Form 5 x Short) von Bass und Avolio (1995). In: Zeitschrift für Arbeits- und Organisationspsychologie 50 (2), S. 61–78.
- Goodnow, E. J.; Wayman, J. C. (2009). The Intersection between Transformational Leadership and Data Use in Schools. Department of Educational Administration. The University of Austin at Texas. Convention of the University Council for Educational Administration. Anaheim, CA., 2009.
- Huber, S. G. (2008). Steuerungshandeln schulischer Führungskräfte aus Sicht der Schulleitungsforschung. In: R. Langer (Hg.): 'Warum tun die das?'. Governanceanalysen zum Steuerungshandeln in der Schulentwicklung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 95–126.
- Kahnemann, D.; Slovic, P.; Tversky, A. (Hg.) (2005). Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.
- LeBoeuf, R. A.; Shafir, E. (2012). Decision Making. In: K. Holyoak und R. G. Morrison (Hg.): The Oxford Handbook of Thinking and Reasoning. New York: Oxford University Press, S. 301–321.
- Leithwood, K.; Jantzi, D. (2005). Transformational leadership. In: B. Davies (Hg.): The essentials of school leadership. Thousand Oaks, CA: Sage Publications, S. 31–43.
- Leithwood, K.; Sun, J. (2012). The nature and effects of transformational school leadership: A meta-analytic review of unpublished research. In: Educational Administration Quarterly 48 (3), S. 387–423.
- Levin, J. A.; Datnow, A. (2012). The Principal Role in Data Driven Decision Making: Using case study data to develop multi-mediator models of educational reform. In: School Effectiveness and School Improvement 23 (2), S. 179–201.
- Mittelstraß, J. (2005): Evidenz. In: J. Mittelstraß (Hg.): Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart: J. B. Metzler, S. 438–439.
- Rousseau, D. M. (2006). Is there such a thing as "evidence-based Management"? In: Academy of Management Review 31 (2), S. 256–269.
- Satterfield, J. M.; Spring, B.; Brownson, R. C.; Mullen, E. J.; Newhouse, R. P.; Walker, B. B.; Whitlock, E. P. (2009). Toward a Transdisciplinary Model of Evidence-Based Practice. In: Milbank Quarterly 87 (2), S. 368–390.
- Schildkamp, K.; Lai, M. K. (2013). Conclusions and a Data Use Framework. In: K. Schildkamp, M. K. Lai und L. M. Earl (Hg.): Data-based Decision Making in Education. Challenges and Opportunities. Dordrecht, Heidelberg, London, New York: Springer, S. 177–191.
- Stumm, S.; Dormann, C.; Mohr, N. (2010). Evidenzbasiertes Management als Schlüssel für den Transfer zwischen Wissenschaft und unternehmerischer Praxis. 47. Kongress der Deutschen Gesellschaft für Psychologie. Bremen, 2010.
- Virgilio, R. F.; Chiapa, A. L.; Palmarozzi, E. A. (2007). Evidence-Based Medicine, Part 1. An Introduction to Creating an Answerable Question and Searching the Evidence. In: The Journal of the American Osteopathic Association 107 (8), S. 295–297.
- Wayman, J. C.; Spring, S. D.; Lemke, M. D.; Lehr, M. D. (2012,5). Using Data to Inform Practice: Effective Principal Leadership Strategies. Department of Educational Administration. The University of Austin at Texas. Annual Meeting of the American Educational Research Association. Vancouver, 2012,5.
- Yates, J. F.; Potworowski, G. A. (2012). Evidence-Based Decision making. In: D. M. Rousseau (Hg.): The Oxford Handbook of Evidence-Based Management. New York: Oxford University Press, S. 198–222.
- Zlatkin-Troitschanskaia, O.; Seidel, J.; Stump, M. (2013). Teachers' evidence-based actions. A comparison of different school types. In: K. Beck und O. Zlatkin-Troitschanskaia (Hg.): From Diagnostics to Learning Success. Proceedings in Vocational Education and Training. Rotterdam: Sense Publishers, S. 281–294.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G092



## **Vorstellungen von 13-15-jährigen Schülern über Staat**

Der Begriff Staat ist im politischen Diskurs in vielerlei Hinsicht von essentieller Bedeutung. So stellt er als politische Strukturform den Rahmen politisch vergleichender Analysen, wird als Akteur auf der politischen Handlungsbühne thematisiert oder als gesellschaftliche Ordnung, die es politisch zu erklären gilt, expliziert. Neuere Diskurse verweisen zudem auf die Überwindung idealtypisch verstandener Staatsformationen und rücken dabei die funktionale Bedeutung des Staates in den Fokus (vgl. Genschel & Zangl, 2007, S. 10f). Analog zu seiner Bedeutungsdimension im fachwissenschaftlichen Kontext spielt der Staat ebenso in politischen Lehr-Lernsituationen eine entsprechend omnipräsente Rolle. Dabei scheint der Staat besonders in der Vermittlungspraxis als Begriff jenseits eines Bezugsaspekts, wie beispielsweise der Charakterisierung als Sozialstaat, jedoch kaum eigenständig thematisiert zu werden.

Ausgehend von der Schieflage zwischen referentieller Bedeutung und eigenständiger Thematisierung, stellt das vorliegende Forschungsprojekt der Arbeitsgruppe „Staatlichkeit im Wandel und Politische Bildung“ im Rahmen des Sonderforschungsbereichs 597 (Staatlichkeit im Wandel) die Frage nach der Perspektive der Lernenden als Bedingung politischer Lehr-Lernprozesse: Welche Vorstellungen von Staat haben 13-15-jährige Schüler und welche Schlussfolgerungen lassen sich daraus für politische Vermittlungsbemühungen ziehen?

Ausgehend von dem Anfangsverdacht, dass eine problematische Vermittlungspraxis auch brüchige Vorstellungsbilder auf der Ebene der Lernenden nahe legt, ist es Ziel der Studie, Vorstellungen über Staat aus Lernericht zu exponieren und in einen Zusammenhang mit Erkenntnissen der Domäne Politik zu bringen, um so Lerngelegenheiten für politische Lernprozesse auszuloten. Die Studie lässt sich dementsprechend innerhalb der politikdidaktischen Lehr-Lernforschung verorten, weil Gegenstände (Staat) sowie Verstehensweisen (Vorstellungen) im Mittelpunkt des Erkenntnisinteresses stehen und damit die Bedingungen tatsächlichen Politikunterrichts. Sowohl die Perspektive der Fachlichkeit als auch diejenige der Alltäglichkeit werden als Bezugspunkte politischer Vermittlungsprozesse verstanden, in Beziehung zueinander gesetzt und fruchtbare Differenzen für den Erkenntnisgewinn im komplexen Forschungs- und Lehr-Lernbereich Staat genutzt (vgl. Kattmann, 2005, S. 166).

Gewählt wurde eine Methodik, die eine Beobachtung bzw. Rekonstruktion subjektiver Deutungen ermöglicht. Die Erhebung der Vorstellungen von 123 Schülern erfolgte mithilfe leitfadengestützter Interviews, die durch drei zentrale Erzählimpulse strukturiert wurden: erstens zu den allgemeinen Assoziationen der Probanden mit dem Begriff Staat, zweitens zur Darstellung der Staatsaufgaben aus Perspektive der Probanden und drittens zur Bewertung der Performanz des Staates in Bezug auf die Erfüllung der zuvor jeweils beschriebenen Aufgaben. Die fünfzehn- bis zwanzig minütigen Interviews wurden im Anschluss in redigierter und paraphrasierter Form transkribiert. Die quantitative Auswertung erfolgte computergestützt mit MAXQDA. In einem Zweifach-Codierungsverfahren wurden die verschiedenen Dimensionen im Verstehen von Staat herausgearbeitet (vgl. Benninghaus, 2002; Schnell, Hill, & Esser, 2011). Insgesamt zeigen die Ergebnisse, dass der Staatsbegriff erst in der Interviewsituation von den Schülern erschlossen wird. Dabei liegen vielfältige Verstehensweisen zugrunde, oftmals jedoch ohne explizite kausale Verknüpfungen zwischen den Vorstellungen. Neben der Auseinandersetzung mit den Auswertungsergebnissen soll der auf der Grundlage der Befunde überlegte Vermöglichungsspielraum für Lehr-Lernprozesse und deren Erforschung in der Diskussion begrenzt oder erweitert werden.

Benninghaus, Hans (2002). Deskriptive Statistik. Eine Einführung für Sozialwissenschaftler. 9. überarb. Aufl., Wiesbaden: Westdeutscher Verlag.

Genschel, Phillip & Zangl, Bernhard (2007). Die Zerfaserung von Staatlichkeit und die Zentralität des Staates. In: Aus Politik- und Zeitgeschichte, Heft 20 -21, S. 10 – 16.

Kattmann, Ulrich (2005). Lernen mit anthropomorphen Vorstellungen? – Ergebnisse von Untersuchungen zur Didaktischen Rekonstruktion in der Biologie. Zeitschrift für Didaktik der Naturwissenschaften. Jg. 11, 2005, 164 – 175.

Schnell, Rainer, Hill, Paul B., & Esser, Elke (2011). Methoden der empirischen Sozialforschung. 9. Aufl., München: Oldenbourg.

Dienstag,

04.03.2014,

10:15 - 12:00,

PEG-Gebäude

Raum 1.G092



## **Statistik aus argumentationstheoretischer Perspektive am Beispiel der Berichterstattung über die erste PISA-Studie in deutschen Printorganen**

In dem Beitrag werden Ergebnisse einer empirischen Studie berichtet, deren Ziel es war methodisch-methodologische Problemfelder bei der Rezeption von Ergebnissen aus der empirischen Bildungsforschung in den Medien zu identifizieren. Der Beitrag plädiert dafür Statistik aus einer entscheidungs- und argumentationstheoretischen Perspektive zu betrachten. Klassischerweise werden statistische Fehlschlüsse innerhalb eines numerisch-mathematischen Kontextes als fehlerhafte Anwendungen formaler Rechenregeln bei der Analyse sozialwissenschaftlicher Daten begriffen (vgl. bspw. Stelzl, 2005). Dabei wird verkannt, dass insbesondere die sozialwissenschaftliche Statistik schon in ihren Anfängen im Wesentlichen ein Instrument zur politischen Entscheidungsfindung war (vgl. Kennedy, 1993; Zeisel, 2008) und davon ausgehend zunehmend eine wichtige Rolle als Argumentationsgrundlage in öffentlichen Diskussionen eingenommen hat. Ziel des Beitrages ist es aufzuzeigen, wie eine argumentationstheoretische Perspektive bei der Analyse von statistischen Argumenten in öffentlichen Diskussionen dabei helfen kann Fehlschlüsse zu identifizieren die „vor“ Problemen und Fehlern entstehen die sich auf spezifische statistische Verfahren beziehen. Dabei wird Bezug genommen auf ein weit gefasstes Konzept von Statistical Literacy, welches insbesondere die nicht-numerischen Aspekte von Statistik betont, sowie auf das Konzept der (logischen) Fallacies, dass auf die Rolle von Statistiken in Kommunikationszusammenhängen fokussiert. Der Begriff der Statistical Literacy entwickelte sich ausgehend von Bestrebungen, den Fokus der Statistikausbildung weg von spezifischen komplexen Verfahren hin zu einem allgemeinen Verständnis der Grundprinzipien und der Vermittlung der alltäglichen Anwendbarkeit und Beurteilungskompetenz zu legen (vgl. Moore, 1990; delMas, 2004). Ohne numerisch-mathematische Kompetenzen zu marginalisieren, richtet sich das Augenmerk dabei insbesondere auf kognitive und logische Strukturen, die statistischem Denken und Urteilen zu Grunde liegen (Ben-Zvi & Garfield, 2004; Gal, 2004; Pfannkuch & Wild, 2004). Der Begriff der Fallacies wird zwar auch im Zusammenhang mit Statistik (Statistical Fallacies) verwendet, lässt sich jedoch auch aus eine argumentationstheoretischen Tradition betrachten, deren Ursprünge in der formalen Logik zu finden sind (siehe zusammenfassend Walton & Brinton 1997). In neuerer Zeit steht eine weniger formale – pragma-dialektische – Auffassung des Fallacy-Begriffs im Zentrum der Fallacy-Theorie, die die Rolle von Argumenten in konkreten Kommunikationszusammenhängen betont und so die Kontextabhängigkeit von formalen und nicht-formalen Fehlschlüssen besonders berücksichtigt (van Eemeren & Grootendorst, 2004; Woods & Walton, 1989; Walton, 1995). Der Beitrag greift auf Ergebnisse einer Inhaltsanalyse der Berichterstattung über die erste PISA-Studie in deutschen Printorganen zurück (Reith, 2014 im Erscheinen). Die Analyse folgte dabei einem Ansatz, der qualitative und quantitative Elemente der Inhaltsanalyse verbindet (vgl. Mayring, 2010; Früh, 2007). Formale Kategorien auf Artekelebene wie „Ressort“ oder „Thema“ wurden quantitativ ausgezählt, die Darstellung der methodisch-methodologischen Problemfelder beschränkt sich jedoch auf eine theorie- und empiriegeleitete Explikation (vgl. Früh, 2001). Ziel dieser Analyse war die Identifizierung methodisch-methodologischer Problemfelder bei der Rezeption sozialwissenschaftlicher Forschungsergebnisse in den Medien. Als zentrale Problemfelder haben sich das Auseinanderfallen von Begriffen im Alltagssprachgebrauch und dem wissenschaftlichen Sprachgebrauch, abduktive Fehlschlüsse, fehlerhafte Kausalinterpretationen, unzulässige oder vorschnelle Verallgemeinerungen sowie Probleme bei der Interpretation absoluter und relativer Häufigkeiten und Wahrscheinlichkeiten identifizieren lassen. Anhand eines dieser Problemfelder soll aufgezeigt werden, wie der Rückgriff auf den Fallacy-Begriff zu einer erweiterten Perspektive auf Statistiken führen kann, die in öffentlichen Diskussionen als Argumentationsgrundlage verwendet werden.

- BEN-ZVI, D.; GARFIELD, J. (2004): Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In: BEN-ZVI, D. und GARFIELD, J. (Hg.): The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Dordrecht. Springer, S. 3–15.
- DELMAS, R. C. (2004): A Comparison of Mathematical and Statistical Reasoning. In: BEN-ZVI, D. und GARFIELD, J. (Hg.): The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Dordrecht. Springer, S. 79–95.
- FRÜH, W. (2001): Kategorienexploration bei der Inhaltsanalyse. Basiswissensgeleitete offene Kategorienfindung (BoK). In: WIRTH, W. und LAUF, E. (Hg.): Inhaltsanalyse. Perspektiven, Probleme, Potentiale. Köln. Halem, S. 117–139.
- FRÜH, W. (2007): Inhaltsanalyse. Theorie und Praxis. Konstanz: UVK Verl.-Ges.
- GAL, I. (2004): Statistical Literacy. Meanings, Components, Responsibilities. In: BENZVI, D. und GARFIELD, J. (Hg.): The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Dordrecht. Springer, S.47–78.
- KENNEDY, G. (1993): Einladung zur Statistik. Frankfurt/Main: Campus-Verl.
- MAYRING, P. (2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim: Beltz.
- MOORE, D. S. (1990): Uncertainty. In: STEEN, L. A. (Hg.): On the shoulders of giants. New approaches to numeracy. Washington. National Academy Press, S. 95–137.
- PFANNKUCH, M.; WILD, C. (2004): Towards an Understanding of Statistical Thinking. In: BEN-ZVI, D. und GARFIELD, J. (Hg.): The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Dordrecht. Springer, S. 17–46.
- REITH, F. (2014): Ungewisse Gewissheiten. Vermittlung sozialwissenschaftlicher Ergebnisse in den Medien. Marburg. Tectum. Im Erscheinen.
- STELZL, I. (2005): Fehler und Fallen der Statistik. Münster, New York, München, Berlin. Waxmann.
- VAN EEMEREN, F. H.; GROOTENDORST, R. (2004): A systematic theory of argumentation. The pragma-dialectical approach. Cambridge: Cambridge University Press.
- WALTON, D. N. (1995): A pragmatic theory of fallacy. Tuscaloosa: University of Alabama Press.
- WALTON, D. N.; BRINTON, A. (1997): Historical foundations of informal logic. Aldershot, Hants, England: Ashgate (Avebury series in philosophy)
- WOODS, J.; WALTON, D. N. (1989): Fallacies. Selected papers 1972-1982. Dordrecht: Foris Publ.
- ZEISEL, H. (2008): Zur Geschichte der Soziographie. In: JAHODA, M., LAZARSFELD, P. F. und ZEISEL, H. (Hg.): Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit ; mit einem Anh. zur Geschichte der Soziographie. 1. Aufl., [21. Dr.]. Frankfurt am Main. Suhrkamp, S. 113–148.





## **Kalibrierung adaptiver Itempools**

Chair(s): Linda Gräfe (Friedrich-Schiller-Universität Jena)

### Einleitung und Fragestellung

Ein wesentlicher Bestandteil bei der Entwicklung neuer Tests stellt die Kalibrierung von Itemparametern dar. Bei diesem Schritt der Testentwicklung werden die Itemparameter der im Test verwendeten Aufgaben bestimmt, so dass sie anschließend in die Haupttestung als fixe Parameter eingehen können. Eine besondere Bedeutung kommt der Kalibrierung des Itempools bei der Entwicklung und Anwendung adaptiver Tests zu (Frey, 2012). Dabei werden den untersuchten Probanden im Testverlauf Aufgaben passend zu ihrer aktuellen Fähigkeitsschätzung vorgelegt. Gerade bei dieser Art des Testens kann es in den Randbereichen der Fähigkeitsskala zu massiven Über- bzw. Unterschätzungen der Personenparameter kommen (Döbler, 2012). In Lehrbüchern zu Testtheorie und Testentwicklung finden sich lediglich allgemeine Forderungen nach einer für den Untersuchungszweck angemessenen Stichprobengröße (z.B. Rost, 2004; Hambleton & Swaminathan, 1985), ohne diese näher zu konkretisieren. Offen bleibt dabei auch, wie viele Items den Personen vorgelegt werden müssten und damit ebenfalls, wie viele Datenpunkte insgesamt benötigt werden, um eine angemessene Schätzung der Itemparameter zu gewährleisten. Um diese offenen Punkte näher zu klären, wird in der vorliegenden Studie untersucht, wie sich (a) die Stichprobengröße und (b) die Anzahl an vorgelegten Items pro Person auf die Genauigkeit der Itemparameterschätzungen auswirken.

### Methode

Die oben aufgeworfene Frage wurde mit einer Simulationsstudie untersucht. Durch die Verwendung eines zweifaktoriellen Versuchsplans wurden sowohl die Stichprobengröße (50, 100, 1000, 10000) als auch die Anzahl der vorgelegten Items pro Person (10, 50, 100) variiert. Mit Hilfe des Statistikprogramms R wurden 500 Replikationen geschätzt und analysiert. Der Untersuchung lag ein 2PL-Modell zugrunde. Als Maße der Genauigkeit der geschätzten Itemparameter wurden Bias, Mean Square Error (MSE) und Schiefe als abhängige Variablen untersucht. Darüber hinaus wurde die Reliabilität in den unterschiedlichen Versuchsbedingungen berechnet und analysiert.

### Ergebnisse

Die Reliabilität nimmt mit einer zunehmenden Zahl an beantworteten Items pro Person erwartungsgemäß zu. Dies ist unabhängig von der Stichprobengröße. Betrachtet man nur dieses Kriterium als Maß für Genauigkeit, so kommt man schnell zum Schluss, jeder Person möglichst viele Items vorzulegen. Die Größe der Stichprobe spielt dabei kaum eine Rolle. Bezieht man allerdings Bias und MSE der Diskriminations- und Schwierigkeitsparameter mit ein, so zeigt sich eine Abhängigkeit von der Stichprobengröße. Beide Kriterien nehmen mit zunehmender Personenzahl ab. Bei der Betrachtung der Schiefe der Verteilungen der einzelnen geschätzten Itemparameter zeigt sich insbesondere für die Schwierigkeitsparameter, dass diese in den Extrembereichen stark von der zu erwartenden Schiefe abweichen. Gerade bei kleineren Stichprobengrößen bleibt nur ein kleiner Abschnitt der Schwierigkeitsskala, der für die Verwendung in einem adaptiven Test uneingeschränkt geeignet ist.

### Diskussion

Die durchgeführte Studie bietet Testentwicklern konkrete Anhaltspunkte, welchen Effekt die Stichprobengröße und die Anzahl vorgelegter Items auf die Genauigkeit der Itemparameterschätzung haben. Dadurch ergeben sich greifbarere Kriterien für die Planung von Kalibrierungsstudien. Gleichzeitig eröffnen sich für die Kalibrierung eines adaptiven Tests neue Fragen, da die Befunde zur Schiefeverteilung der Itemparameterschätzer große Abweichungen in den anschließenden Messungen von Personenparametern nach sich ziehen können. Weitere Untersuchungen sind nötig, um neue Lösungen zu finden, mit denen entweder die notwendige Genauigkeit bei der Kalibrierung des Itempools eines adaptiven Tests erzielt oder die Schiefeproblematik bei der adaptiven Itemvorgabe berücksichtigt wird.

Doebler, A. (2012). The Problem of Bias in Person Parameter Estimation in Adaptive Testing. *Applied Psychological Measurement*, 36(4), 255–270.

Frey, A. (2012). Adaptives Testen. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (2. Auflage, S. 275–293). Berlin, Heidelberg: Springer.

Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer.

Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion*. Bern: Huber.

Dienstag,  
04.03.2014,  
10:15 - 12:00,  
PEG-Gebäude  
Raum 1.G102



## ***Einfluss von Positionseffekten auf die Schätzung der Itemparameter bei der Entwicklung eines computerisierten adaptiven Tests***

Bei der Kalibrierung von Kompetenztests zeigen sich regelmäßig sogenannte Positionseffekte (z.B. Hartig & Buchholz, 2012). Diese bestehen darin, dass die Lösungshäufigkeit der gleichen Items zunehmend niedriger ausfällt, je später sie im Testverlauf vorgegeben werden. Bei herkömmlichen linearen Tests wird dem Problem häufig durch die Verwendung von balancierten Testheftdesigns (Frey, Hartig & Rupp, 2009) begegnet, wodurch zum Schwierigkeitsparameter jedes Items der Mittelwert aller Positionseffekte addiert wird. Die resultierenden Itemparameter können mit wenigen Einschränkungen für weitere Analysen genutzt werden. Ein substanzielleres Problem als bei linearen Tests liegt indes vor, wenn Itemparameter im Rahmen eines computerisierten adaptiven Tests (CAT) genutzt werden sollen. Dies deshalb, da hierbei die geschätzten Itemparameter als fest angenommen werden ohne zu berücksichtigen an welcher Position die einzelnen Items bei der adaptiven Testung vorgegeben werden. Im Rahmen der Entwicklung von mehreren CATs im Projekt „Messung allgemeiner Kompetenzen-adaptiv“ wurden deshalb bei der Auswertung der Kalibrierungsstudie zwei Fragestellungen untersucht.

1) Lassen sich Positionseffekte identifizieren?

2) Sind etwaige Positionseffekte (a) global oder (b) itemspezifisch?

Bei der Kalibrierungsstudie wurden einer Stichprobe von  $N = 1632$  Schülerinnen und Schülern beruflicher Schulen unter anderem 73 Items zur Messung der Lesekompetenz zur Bearbeitung vorgelegt. Die Zuteilung der Leseitems zu den Testpersonen erfolgte auf Basis eines unvollständigen balancierten Testheftdesigns mit 73 verschiedenen Testheften und jeweils 9 Positionen. Dieses Design wurde in vollständig permutierter Form mit zwei weiteren Designs mit jeweils 12 Positionen verbunden, so dass die Leseitems an insgesamt 33 Positionen vorgegeben werden konnten. Die Analyse der Positionseffekte erfolgte auf Basis der Item-Response-Theorie mit dem Programm ConQuest 3.0. Konkret wurden drei Modelle verglichen.

Modell 1: Ein eindimensionales Rasch-Modell ohne Positionseffekte (in ConQuest: model item).

Modell 2: Ein Multifacetten-Rasch-Modell (z.B. Linacre, 1994) mit der zusätzlichen Facette Position zur Modellierung globaler Positionseffekte (in ConQuest: model item+position).

Modell 3: Ein Multifacetten-Rasch-Modell mit globalem Positionseffekten und itemspezifischen Positionseffekten (in ConQuest: model item+position+item\*position).

Das Modell 2 zeigte eine signifikant bessere Modellpassung als das Modell 1 ( $p < .001$ ). Auch die Informationskriterien sprechen eindeutig für das Modell 2 (BIC = 10344.47, AIC = 9880.28; CAIC = 9889.97) im Vergleich zum Modell 1 (BIC = 10537.18, AIC = 10132.37, CAIC = 10139.69). Die Positionseffekte steigen über die Positionen fast perfekt linear an und haben eine Spannweite von 0.451 Logits. Die Hinzunahme itemspezifischer Positionseffekte im Modell 3 resultierte indes in keiner signifikanten Verbesserung der Modellpassung ( $p = .999$ ). Auch die Informationskriterien fallen mit BIC = 14983.05, AIC = 10718.98 und CAIC = 12205.04 für das Modell 3 deutlich schlechter aus als für das Modell 2. Entsprechend wurde das Modell 2 als endgültiges Modell gewählt.

Zusammenfassend kann geschlossen werden, dass in den vorliegenden Daten globale itemunspezifische Positionseffekte enthalten sind. Diese können beim adaptiven Testen berücksichtigt werden, indem bei der Itemauswahl zur Schwierigkeit des Items der geschätzte Positionseffekt der gerade aktuellen Position addiert wird.

Frey, A., Hartig, J., & Rupp, A. (2009). An NCME Instructional Module on Booklet Designs in Large-Scale Assessments of Student Achievement: Theory and Practice. *Educational Measurement: Issues and Practice*, 28, 39-53.

Hartig, J., & Buchholz, J. (2012). A multilevel item response model for item position effects and individual persistence. *Psychological Test and Assessment Modeling*, 54, 418-431.

Linacre, J.M. (1994). *Many-Facet Rasch Measurement*. Chicago. MESA Press.

## ***Effekte der Bearbeitungszeit auf die Lösungswahrscheinlichkeit unter Berücksichtigung moderierender Einflüsse auf Personen- und Itemebene***

Für die Rolle der Bearbeitungszeit bei der Bearbeitung eines Items können unterschiedliche, entgegengesetzte Annahmen getroffen werden. Langsamere Bearbeitung könnte mit der Wahrscheinlichkeit, eine korrekte Antwort zu erzielen, positiv assoziiert sein, da geringeres Tempo eine sorgfältigere Bearbeitung erlaubt. Auf der anderen Seite kann die Beziehung negativ sein, wenn eine flüssige und somit schnellere Bearbeitung eine höhere Fähigkeitsausprägung anzeigt.

Das Ziel der vorliegenden Studie war die Klärung der Frage, unter welchen Bedingungen die eine oder andere Interpretation bei Aufgaben aus der Domäne Reasoning zutrifft. Unter Bezugnahme auf die Theorie automatisierter und kontrollierter Informationsverarbeitung (vgl. Schneider & Chein, 2003), formulierten wir die Hypothese, dass die Stärke und Richtung des Effektes der Bearbeitungszeit von der Itemschwierigkeit und der Personenfähigkeit abhängt (vgl. Goldhammer, Naumann, Stelter, Tóth, Rölke, & Klieme, in press). Das heißt, bei einfachen Items, die von fähigen Personen bearbeitet werden, sollte sich ein negativer Effekt der Bearbeitungszeit zeigen und umgekehrt. Darüber hinaus prüften wir, ob die Itemschwierigkeit der Matrizenaufgaben durch die Anzahl anzuwendender Regeln (vgl. der Taxonomie von Carpenter, Just & Shell, 1990) erklärt werden kann. Dadurch sollte bestimmt werden, ob die moderierende Rolle der Itemschwierigkeit zumindest teilweise auf dieses Itemmerkmal zurückgeführt werden kann.

Insgesamt nahmen 230 Personen im Alter zwischen 19 und 40 Jahren an der Studie teil und bearbeiteten eine computerbasierte Version des Raven's Advanced Progressive Matrices-Test. Analysen unter Anwendung von generalisierten linearen gemischten Modellen demonstrierten, dass der Effekt der Bearbeitungszeit insgesamt negativ ausfiel. Jedoch zeigte sich wie erwartet, dass dieser signifikant über Personen und Items variiert. Für einfache Items und fähige Personen war der Effekt der Bearbeitungszeit stark negativ, für schwere Items und weniger fähige Personen war er deutlich weniger negativ oder sogar positiv. Die Anzahl der Regeln innerhalb einer Matrizenaufgabe (unterschiedliche Typen von Regeln oder unterschiedliche Instanzen desselben Typus) erwies sich als signifikanter Prädiktor der Itemschwierigkeit. Insbesondere zeigte sich ein positiver Interaktionseffekt zwischen der Anzahl von Regeln und dem Effekt der Bearbeitungszeit: Der Effekt der Bearbeitungszeit fiel mit zunehmender Anzahl von anzuwendenden Regeln und damit steigender Aufgabenkomplexität weniger negativ aus.

Die vorliegenden Befunde replizieren die von Goldhammer et al. (in press) für die Domänen Lesen und Problemlösen gewonnenen Befunde für den Bereich kognitiver Grundfähigkeiten. Die Ergebnisse belegen abermals, dass die Bearbeitungszeit von Testaufgaben keine feste Bedeutung besitzt. Vielmehr kann die Bearbeitungszeit nur relativ zu Itemschwierigkeit und Personenfähigkeit, und zu Itemcharakteristiken, die die Itemschwierigkeit beeinflussen, interpretiert werden.

Carpenter, P. A., Just, M. A. & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. & Klieme, E. (in press). The time on task effect in reading and problem solving is moderated by item difficulty and ability: Insights from computer-based large-scale assessment. *Journal of Educational Psychology*.

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, 27, 525-559.



## ***Arbeitest Du noch oder machst Du schon schlapp? – Eine empirische Untersuchung zur Robustheit von Schulformunterschieden in nicht vollständig rotierten Testheft-Designs***

Zur Schätzung von Leistungsunterschieden zwischen Gruppen und deren Korrelate werden in der Bildungsforschung üblicherweise Matrix-Designs eingesetzt (z.B. OECD, 2009). Der Vorteil dieser Designs besteht in einer breiten Abdeckung von Leistungsbereichen ohne dabei die Testteilnehmer/innen zu überlasten, da jede/r nur einen Bruchteil der insgesamt verfügbaren Items bearbeitet (Gonzalez & Rutkowski, 2010). Jedoch können designbedingte Störfaktoren, wie beispielsweise die Inhalte vorangestellter Items (Kontexteffekte; Leary & Dorans, 1985) oder deren Position (Positionseffekte; Kingston & Dorans, 1982, 1984), die Leistungsschätzungen erheblich verzerren. Zur Balancierung designbedingter Störfaktoren erfolgt deshalb häufig eine Rotation der Matrix-Designs: Einzelne Gruppen von Probanden erhalten Testhefte mit unterschiedlichen Aufgaben, die an verschiedenen Positionen präsentiert werden (Gonzalez & Rutkowski, 2010; Prenzel et al., 2007). Eventuelle Gruppenunterschiede werden dann als mittlere Unterschiede über alle Kontexte und Positionen interpretiert. In der Praxis unterscheidet man zwischen rotierten Matrix-Designs, bei denen eine vollständige bzw. eine unvollständige Balance der Position erfolgt (Frey & Bernhardt, 2012). Da die Interpretation der Ergebnisse sowohl von der Variation der Design-Artefakte über die Gruppen als auch von der Vollständigkeit der Rotation abhängt, sind Leistungsschätzungen bei der Verwendung von unvollständig rotierten Matrix-Designs insbesondere dann verzerrt, wenn Designs-Artefakte (z.B. Positionseffekte) erheblich zwischen den untersuchten Gruppen variieren. Daran anknüpfend verfolgte diese Studie das Ziel, die Robustheit von Schulformunterschieden in nicht perfekt rotierten (domänen-gemischten) Matrix-Designs genauer zu erforschen. Dabei wurde der Forschungsfrage nachgegangen, ob Schulformunterschiede in Abhängigkeit der Item-Cluster bzw. der Cluster-Position variieren und wenn ja, ob die Auswertung von positionsspezifischen Gruppenunterschieden eine adäquate Alternative in unvollständig rotierten Testheft-Designs darstellt. Für die Beantwortung dieser Forschungsfrage wurde eine Stichprobe bestehend aus N = 3183 Berliner Schüler/innen analysiert, die im Rahmen der BERLIN-Studie an Schulleistungstests teilnahmen. Dabei diente die Schulform der Schüler/innen als unabhängige Variable und untergliederte sich in der untersuchten Stichprobe in 29 Hauptschulen, 23 Realschulen, 35 Gymnasien und 29 Gesamtschulen. Als abhängige Variable wurde die Schulleistung der Schüler/innen in den drei Domänen Lesen, Mathematik und Naturwissenschaften, operationalisiert über ihr Antwortverhalten in vier PISA-2006-Testheften (OECD, 2009), erfasst. Für die Beantwortung der Forschungsfrage wurden drei Strukturgleichungsmodelle (Bollen, 1989; Bollen & Long, 1993) aufgestellt, deren jeweilige Passung auf die beobachteten Daten mit Hilfe von Wald-Tests (Wald, 1943) statistisch getestet wurde. Im ersten Modell wurde nur der Haupteffekt der Schulform zugelassen, während das zweite Modell um eine Schulform  $\times$  Item-Cluster-Interaktion erweitert wurde. Das dritte Modell enthielt stattdessen die Schulform  $\times$  Cluster-Positionen-Interaktion. Tatsächlich wick in den Domänen Lesen und Mathematik nur das letztere Modell nicht signifikant von den beobachteten Daten ab ( $\chi^2(3) = 2.18, p = .54$  für Lesen bzw.  $\chi^2(6) = 9.79, p = .13$  für Mathe). So waren Hauptschüler deutlich stärker von Positionseffekten betroffen als Schüler der übrigen Schulformen, was sich in fast einem Drittel Standardabweichung Leistungsunterschied von der ersten bis zur letzten Position widerspiegelte. Dagegen zeigten im Bereich der naturwissenschaftlichen Domäne alle drei Modelle eine sehr gute Passung auf die beobachtete Datenstruktur ( $ps \geq .45$ ). Zusammenfassend wirkten sich Design-Artefakte, in diesem Fall Positionseffekte, substantiell auf Schätzungen in unvollständig rotierten Testheft-Designs aus. Die durchgeführte Studie lieferte Belege dafür, dass Schulformunterschiede in Abhängigkeit der Cluster-Position variierten. Dies wurde als Hinweis darauf gedeutet wurde, dass positionsspezifische Gruppenunterschiede eine große Rolle für Large-Scale-Studien zu spielen scheinen und Design-Artefakte insbesondere für unvollständig rotierte Matrix-Designs problematisch sind. Limitationen (z.B. keine bestätigende Evidenzen in der naturwissenschaftlichen Domäne) und praktische Implikationen dieser Studie (z.B. Ergebnisinterpretation in Abhängigkeit der Design-Matrix) sollen im Vortrag näher diskutiert werden.

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Bollen, K. A., & Long, S. J. (1993). Testing structural equation models. Newbury Park: Sage Publications.

Frey, A., & Bernhardt, R. (2012). On the importance of using balanced booklet designs in PISA. Psychological Test and Assessment Modeling, 54(4), 397-417.

Gonzalez, E., & Rutkowski, L. (2010). Principles of multiplematrix booklet designs and parameter recovery in large scale assessments. IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments, 3, 125-156.

Kingston, N. M., & Dorans, N. J. (1982). The effect of the position of an item within a test on item responding behavior: An analysis based on item response theory. GRE Board Professional Report 79-12bP. Princeton NJ: Educational Testing Service.

Kingston, N. M., & Dorans, N. J. (1984). Item location effects and their implications for IRT equating and adaptive testing. Applied Psychological Measurement, 8, 147-154.

Leary, L. F., & Dorans, N. J. (1985). Implications for altering the context in which test items appear: A historical perspective on an immediate concern. Review of Educational Research, 55, 387-413.

OECD (2009). PISA 2006 Technical Report. Paris: OECD Publishing.

Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., & Pekrun, R. (2007). PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich. Münster: Waxmann.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 54(3), 426-482.



## ***Eine Untersuchung der latenten Struktur von Rosenbergs Selbstwertkala im Rahmen von Factor Mixture Modeling***

Hintergrund: Der Selbstwert wird durch den Wert, den Personen sich selbst zumessen definiert und stellt somit die evaluative Komponente des Wissens über sich selbst dar; hierbei bedeutet ein hoher Selbstwert eine positive globale Evaluation des eigenen Selbst und vice versa (Baumeister, Campbell, Krueger & Vohs, 2003). Der Selbstwert wird in der pädagogisch/psychologischen Forschung häufig mittels der Selbstwertkala von Rosenberg gemessen. Die als unidimensional und kontinuierlich konzipierte Skala besteht aus 10 Items, die zur Hälfte positiv (z.B. "Alles in allem bin ich mit mir selbst zufrieden") und zur Hälfte negativ (z.B. "Ich fühle mich von Zeit zu Zeit richtig nutzlos") formuliert sind. In faktoranalytischen Studien zur Dimensionalität der Skala fand sich häufig, dass ein unidimensionales Modell die erhobenen Daten nur unzureichend abbildet. Erst durch die Modellierung eines zweiten, auf die negativ formulierten Items bezogenen Faktors, zeigen konfirmatorische Faktorenanalysen (KFA) einen akzeptablen Fit. Wenngleich einige Autoren vorschlagen, den zweiten Faktor inhaltlich (als spezifische Facette des Selbstwerts) zu interpretieren (z.B. Weems et al., 2003; Kaufman et al., 1991), weisen die meisten Forscher diese Interpretation zurück und gehen von einem Methodenfaktor aus, der sich auf die negative Itemformulierung bezieht (z.B. Marsh, 1996; DiStefano & Motl, 2006). Die Grundidee ist dabei, dass einige Versuchspersonen bei der Skalenbearbeitung unachtsam sind, wobei beispielsweise solche mit einem hohen Selbstwert (versehentlich) negativen Statements zustimmen würden. In Bezug auf die Modellierung stellt sich hierbei die Frage, ob die KFA die Methode der Wahl ist. Wenn ein Teil der Versuchspersonen die Kreuze auf der falschen Seite der Antwortskala setzt, so deutet dies vielmehr auf latente Klassen hin, die im Rahmen von KFA nicht modelliert werden können.

Fragestellung: In dieser Studie soll der Frage nachgegangen werden, welche latente Struktur der Selbstwertkala von Rosenberg zugrunde liegt, was unseres Wissens nach in bisheriger Forschung ausschließlich im Rahmen von (konfirmatorischen) Faktorenanalysen untersucht wurde. Hierbei gehen wir von zwei Klassen der Itembeantwortung aus.

Methode: Im Rahmen dieser Studie ziehen wir Daten der Rosenberg Selbstwertkala von Fünftklässlern (n=275) der Pilotstudie der National Educational Panel Study (Blossfeld, Rossbach & von Maurice, 2011; www.neps-data.de) heran. Wir schlagen ein Modell vor, in dem wir den Selbstwert als latente kontinuierliche Variable und das Antwortverhalten als zweite, latente kategoriale Variable mit zwei Klassen einführen. Unsere Hypothese ist dabei, dass eine erste Gruppe von Versuchspersonen inhaltlich konform das Antwortverhalten wechselt, wenn sie statt positiv formulierter negativ formulierte Items beantwortet. Eine zweite Gruppe wechselt das Antwortverhalten nicht, antwortet also inhaltlich inkonsistent. Die Modellierung eines solchen Modells (Muthén, 2008; Masyn, Henderson, & Greenbaum, 2010) ist im Rahmen einer Factor Mixture Analyse (FMA) möglich, in der gleichzeitig kategoriale und kontinuierliche Variablen modelliert werden. Der Fit dieser FMA wird mit klassischen KFA Modellen verglichen.

Ergebnisse/Diskussion: Anhand der Informationskriterien AIC und BIC zeigt sich, dass die FMA die Daten sowohl besser abbildet, als eine eindimensionale, aber auch als eine zweidimensionale KFA. Die Mehrzahl der Versuchspersonen antworten formulierungsunabhängig inhaltlich konsistent (Klasse A, etwa 80 %). In dieser Klasse fanden sich erwartungskonform positive Faktorladungen für die positiv formulierten Items und negative für die negativ formulierten Items. In Klasse B (etwa 20 %) fanden sich ausschließlich positive Faktorladungen. Diese Ergebnisse interpretieren wir dahingehend, dass Rosenbergs Selbstwertkala ein eindimensionales Merkmal (den globalen Selbstwert) misst, etwa ein Fünftel der Versuchspersonen auf die negativ formulierten Items jedoch nicht erwartungskonform, sondern gegenteilig antwortet. Dieses Ergebnis zeigt zum einen, dass ein solches Antwortverhalten bei der Verwendung von Fragebögen mit positiv und negativ formulierten Items überprüft werden sollte. Zum anderen lässt sich mit solchen Modellierungen in Zukunft auch untersuchen, ob sich die Versuchspersonen der beiden Klassen z.B. in ihrer Lesefähigkeit oder ihrem Alter unterscheiden, worauf bereits andere Forschungsergebnisse einen ersten Hinweis geben (z.B. Marsh, 1996).

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does High Self-Esteem Cause Better Performance, Interpersonal Success, Happiness, or Healthier Lifestyles? *Psychological Science in the Public Interest*, 4, 1-44.  
Blossfeld, H.-P., Rossbach, H. G., & von Maurice, J. (2011). Education as a Lifelong Process : The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft (Sonderheft)*. 14. Wiesbaden: VS Verlag für Sozialwissenschaften.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, 13(3), 440-464.

Kaufman, P., Rasinski, K. A., Lee, R., & West, J. (1991). National Education Longitudinal Study of 1988. Quality of the responses of eighth-grade students in NELS88. Washington, DC: U.S. Department of Education.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70(4), 810-819. doi: 10.1037/0022-3514.70.4.810.

Masyn, K. E., Henderson, C. E., & Greenbaum, P. E. (2010). Exploring the Latent Structures of Psychological Constructs in Social Development Using the Dimensional-Categorical Spectrum. *Social Development*, 16(3), 470-493.

Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In G. R. Hancock & K. M. Samuelson (Eds.), *Advances in latent variable mixture models* (S. 1-24). Charlotte, NC: Information Age Publishing.

Weems, G. H., Onwuegbuzie, A. J., Schreiber, J. B., & Eggers, S. J. (2003). Characteristics of Respondents Who Respond Differently to Positively and Negatively Worded Items on Rating Scales. *Assessment & Evaluation in Higher Education*, 28(6), 587-607.



***Between- and Within-Differentiation in Higher Education: New Dimensions of Social Inequality?***

Chair(s): Birgit Becker (Goethe-Universität Frankfurt)

In Germany as in many European countries participation in higher education has expanded considerably accompanied with institutional differentiation. Besides the vocationally vs. academically oriented tiers of higher education (Fachhochschule and university) a further division of different tertiary levels was introduced by Bachelor and Master programmes in the course of the so-called 'Bologna process'. Additionally, pathways into higher education have become more diverse and student experience has broadened e.g. by studying abroad or participating in the labour market while enrolled. It is an open theoretical and empirical question, whether and in what respect this processes of differentiation and pluralisation were accompanied by a reduction or an increase in social inequality. In my presentation I summarise the findings of a number of recent sociological studies on (a) the consequences of differentiation on social equalities in access to higher education and (b) social inequalities in students' lives. Several studies show that educational expansion has led to the inclusion of less privileged students into higher education. These students now enter higher education that had not entered higher education in the past. However, higher-tier institutions and master programmes have remained a stronghold of the privileged. Besides these differences in access, students from the upper classes are more likely to choose the most prestigious subjects, they are less affected by the necessity of working during their studies and sign up more often for study programmes abroad. Hence, children from more privileged homes are still able to assert their advantageous positions.

Dienstag,  
04.03.2014,  
13:00 - 14:00,  
Casino-Gebäude  
Raum 823





Chair(s): Nina Jude (DIPF Frankfurt)

Diskutant/in: Oliver Lüdtke (IPN Kiel und ZIB)

Mit PISA 2012 wurde Ende 2013 die fünfte Erhebungsrunde dieser internationalen Schulleistungsstudie vorgestellt. Nach dem PISA-Schock, der auf die Veröffentlichung der ersten Ergebnisse vor zwölf Jahren folgte, gehört PISA mittlerweile zum deutschen Bildungsalltag und liefert regelmäßig Indikatoren über die Leistungsfähigkeit von Schulsystemen weltweit. Wissen wir also inzwischen nicht alles, was es daraus zu erfahren gibt?

Dieses Symposium widmet sich aus vier Perspektiven der Frage, warum die PISA-Studie nach wie vor hoch aktuell und eine umfassende Quelle von Daten zu Merkmalen von Bildungssystemen ist. Anhand von Daten aus PISA 2012 sowie zweier daran angekoppelter Begleitstudien in Deutschland werden verschiedene Aspekte des Kompetenzerwerbs und der Kompetenzmessung herausgearbeitet, die ohne PISA noch weitgehend unbekannt wären:

Aus bildungspolitischer Sicht können die Befunde aus PISA 2012 Auskunft darüber geben, inwieweit die Ziele von Maßnahmen, die mit der Einführung länderübergreifender Bildungsstandards ab 2003 umgesetzt wurden, inzwischen erreicht worden sind: ein insgesamt höheres Leistungsniveau, deutlich weniger Schülerinnen und Schüler auf den niedrigsten Kompetenzstufen sowie die Verringerung von Disparitäten. Der erste Einzelbeitrag ordnet die fünfte PISA-Studie in den Kontext deutscher Bildungspolitik seit PISA 2000 ein und diskutiert Implikationen für künftige PISA-Erhebungsrunden.

Ein aus methodischer Perspektive relevanter Aspekt von PISA 2012 ist die teilweise am Computer durchgeführte Testung von Kompetenzen (CBA = Computer-Based Assessment). Der zweite Einzelbeitrag berichtet Befunde aus einer CBA-Begleitstudie, die im Anschluss an die PISA-Erhebung 2012 von einem Teil der Schülerinnen und Schüler bearbeitet wurde. Untersucht wird die Vergleichbarkeit von PISA-Leseaufgaben, dazu werden die Stabilität von Trendaussagen sowie Implikationen für individuelle Veränderungswerte in Längsschnittstudien analysiert.

Seit PISA 2003 wird in Deutschland eine zusätzliche Stichprobe von zwei vollständigen neunten Klassen in jeder PISA-Schule getestet. Dies ermöglicht vertiefende Einblicke in Schüler- und Lehrerperspektiven insbesondere auf Unterricht. Im dritten Einzelbeitrag wird der Zusammenhang zwischen Unterrichtsmerkmalen und übergeordneten Bildungszielen anhand dieser Neunte-Klassen-Stichprobe vorgestellt. Solche Bildungsziele sind neben Wissen und kognitiven Fähigkeiten beispielsweise auch die Freude von Jugendlichen an Mathematik, das Erkennen der Relevanz von Mathematik für ihr eigenes Leben oder die Entwicklung einer zielstrebigem und gewissenhaften Arbeitsweise.

Deutschland ist einer von nur wenigen PISA-Teilnehmerstaaten, in denen es gesonderte Schulen für Kinder mit sonderpädagogischem Förderbedarf gibt. Im Rahmen von PISA wird für die Schülerinnen und Schüler an diesen Förderschulen ein verkürztes, nicht aber einfacheres, Testheft eingesetzt. Anhand eines Oversamples von zusätzlichen Förderschulen in Deutschland wurden verschiedene weitere Anpassungen von Testmaterialien variiert, wie etwa einfachere Testhefte. Ergebnisse aus dieser Untersuchung werden im vierten Einzelbeitrag vorgestellt.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
Casino-Gebäude  
Raum 823

## ***PISA 2012: Ergebnisse des 5. Internationalen Vergleichs und Perspektiven für Deutschland***

Mit PISA 2012 geht das Programme for International Student Assessment bereits in die fünfte Runde. Inzwischen ist PISA fester Bestandteil des Bildungsmonitorings und der Bildungsberichterstattung in Deutschland. Über die vordergründige „Routine“ der Prozeduren und Rezeption der Ergebnisse hinaus bringt PISA 2012 jedoch auch einige Neuerungen: Zum ersten Mal wurden curricular übergreifende Kompetenzen im Problemlösen am Computer getestet. Zudem ermöglicht ein erweitertes Fragebogendesign, ein besonders breites Spektrum an Einstellungen, Verhaltensweisen sowie motivationalen und emotionalen Orientierungen von Schülern, Lehrpersonen und Schulleitungen zu erheben. Diese Weiterentwicklungen liefern relevante und interessante Befunde für Bildungspolitik und Bildungspraxis.

Insbesondere in Bezug auf Veränderungen in der Qualität von Bildungsergebnissen in den letzten zwölf Jahren bietet PISA 2012 aufschlussreiche Grundlagen, um die Entwicklungen von Schulen einzuschätzen. Diese lassen sich auf der Systemebene verknüpfen mit Ergebnisse anderer umfangreicher Studien zur Kompetenzmessung in verschiedenen Alterskohorten, etwa dem Programme for the International Assessment of Adult Competencies (PIAAC) oder dem Ländervergleich für die Bildungsstandards zum Mittleren Schulabschluss.

Dieser erste Beitrag des Symposiums PISA 2012 – Aktuelle Perspektiven zur fünften Erhebungsrunde nimmt eine Meta-Perspektive ein und diskutiert den Gewinn einer kontinuierlichen Beteiligung an internationalen Large-Scale-Assessments am Beispiel der PISA-Studie 2012. Ausgehend von den zentralen Ergebnissen für Deutschland werden Fortschritte der Sekundarschülerinnen und –schüler in Deutschland, die Entwicklung von Qualitätsaspekten sowie die bestehenden Herausforderungen für das deutsche Bildungssystem aufgezeigt.

## ***Lassen sich Unterschiede zwischen papier- und computerbasiert gemessener PISA-Lesekompetenz durch Log- und Prozessdaten erklären?***

Falsifikationsversuche der Testäquivalenzhypothese bilden die Grundlage für den Übergang von papierbasierter (PBA) zu computerbasierter Kompetenzdiagnostik (CBA) und sind auch im Kontext von Bildungsvergleichsstudien relevant, bspw. im Hinblick auf die Stabilität von Trendergebnissen oder im Hinblick auf individuelle Veränderungswerte in Längsschnittstudien. Im Vortrag wird die Vergleichbarkeit von PISA-Leseaufgaben berichtet und basierend auf Log- und Prozessdaten (LPD) wird versucht, die gefundenen Unterschiede zu erklären.

Unterschiede zwischen papier- und computerbasiert gemessener PISA-Lesekompetenz

Den theoretischen Rahmen liefert die Beschreibung von Messeigenschaften, wodurch CBA und PBA als Konglomerat verschiedener Eigenschaften betrachtet werden können. Beispielsweise konnte bei CBA nur innerhalb von Units frei zwischen mehreren Seiten (Textstimulus und anschließende Fragen) navigiert werden, während bei PBA die Aufgabenbearbeitungssequenz innerhalb und zwischen Units frei wählbar war.

Methode

Bei einer Stichprobe von N=852 15-Jährigen wurden in einer nationalen Ergänzungsstudie zu PISA 2012 ein oder zwei intakte Cluster administriert. Die Probanden wurden randomisiert den Bedingungen CBA oder PBA sowie einer Teilstichprobe den Bedingungen CBA-PBA oder PBA-CBA zugeordnet.

Hypothesen 1

Zunächst wurde geprüft, ob (1) die Itemschwierigkeiten zwischen PBA und CBA identisch sind (Messmodellinvarianz), oder ob die Itemschwierigkeitsänderungen im Messmodell als (2) fixierter oder (3) zufälliger Effekt modelliert werden müssen. Darauf aufbauend wurde (4) erwartet, dass sich die latente Cross-Mode-Korrelation nicht von 1.0 unterscheidet.

Ergebnisse 1

Hypothese (1) musste verworfen werden ( $p < 0.01$ ). Die CBA-Itemschwierigkeiten konnten unter Kontrolle der Position und einer globalen Schwierigkeitsverschiebung jedoch nicht von den PBA-Itemschwierigkeiten unterschieden werden (Hypothese 2). Es wurden keine interindividuellen Unterschiede im globalen Modeffekt gefunden (Hypothese 3 verworfen). Für das finale Messmodell mit Positions- und Modeffekt konnte Hypothese (4) beibehalten werden, da die latente Korrelation von 0.911 nicht signifikant von 1.0 verschieden war ( $p = 0.72$ ).

Erklärung der Unterschiede durch Log- und Prozessdaten (LPD)

Durch Randomisierung konnte die Schwierigkeitsverschiebung eindeutig als Modeffekt interpretieren werden. Eine weitere Erklärung durch LPD erfolgt, um den Effekt mit den Messeigenschaften in Beziehung zu setzen. Dafür wurden LPD-Indikatoren ausgewählt, welche sowohl für CBA als auch PBA gebildet werden können: Aufgabenbearbeitungssequenz (SEQ), benötigte Antwortzeit (AZ) zum Notieren der Lösung (bspw. tippen) sowie Aufgabenbearbeitungszeiten (BZ) pro Item (inkl. approximierter Lesezeit in der ersten Aufgabe jeder Unit). Die LPD für PBA wurden mit digitalen Kugelschreibern erfasst. Die BZ pro Item wurden aus der Zeitdifferenz zur jeweils vorangegangenen Antwort approximiert.

Hypothesen 2

Unterschiede der SEQ wurden zwischen PBA und CBA erwartet, resultierend (5) in einem verringerten Modeffekt, wenn PBA-Antworten die nicht der sequentiellen Bearbeitung entsprechen ignoriert werden. Es wurde (6) erwartet, dass die AZ für CBA geringer ausfällt als bei PBA, insbesondere bei Textantworten. Für die BZ als abhängige Variable wurde (7) ebenfalls ein Modeffekt erwartet, mit geringeren BZ für CBA. Konkret wurde angenommen, dass sich neben der BZ auch die Lesezeit zwischen CBA und PBA unterscheidet, d.h., (8) größere Unterschiede in der BZ der jeweils ersten Aufgabe jeder Unit. Schließlich wurde (9) erwartet, dass sich der Modeffekt bzgl. der Aufgabenschwierigkeiten reduziert, wenn nur ein Subsample, gebildet mit vergleichbaren beobachteten BZ, analysiert wird.

Ergebnisse 2

Anders als erwartet wurden nur geringe Unterschiede in der SEQ zwischen CBA und PBA gefunden, so dass die Prüfung von Hypothese (5) nur eingeschränkt möglich war. Statistisch bedeutsame Unterschiede bzgl. unterschiedlicher AZ konnten nicht gefunden werden (Hypothese 6) wurde verworfen. Die SEQ und AZ spielen für die Erklärung von Modeffekten deshalb eine untergeordnete Rolle. Die Hypothesen (7) und (8) konnten bestätigt werden. Die Ergebnisse der Prüfung von Hypothese (9) waren abhängig vom methodischen Vorgehen, was abschließend zusammen mit den Implikationen der Befunde für die computerbasierte Diagnostik von Lesekompetenz diskutiert wird.

## ***Merkmale des Mathematikunterrichts in Deutschland und ihr Zusammenhang mit kognitiven sowie motivational-affektiven Lernergebnissen in PISA 2012***

### Theoretischer Hintergrund

Der Bildungsauftrag der Schule beschränkt sich nicht auf die Vermittlung von Wissen und kognitiven Fähigkeiten, sondern bezieht auch die Entwicklung positiver motivationaler und affektiver Orientierungen mit ein (vgl. Prenzel, 2011). Für das Fach Mathematik gehört dazu beispielsweise, dass Jugendliche Freude an Mathematik entwickeln und die Relevanz von Mathematik für ihr eigenes Leben und ihre Zukunft erfahren. Ferner sollten Jugendliche herausfinden, was sie sich selbst in Bezug auf ihre mathematischen Fähigkeiten zutrauen können (Selbstwirksamkeitserwartung), und Gelegenheit haben, eine zielstrebige und gewissenhafte Arbeitshaltung zu entwickeln.

Für das Erreichen solcher mehrdimensionaler Bildungsziele im Fach Mathematik ist der Mathematikunterricht der entscheidende Ort, da die Lehrkraft im Unterricht systematisch Lerngelegenheiten zur Verfügung stellt. Für den Lernprozess der Schülerinnen und Schüler ist jedoch entscheidend, wie diese Lerngelegenheiten wahrgenommen und genutzt werden. Erst die aktive und konstruktive Auseinandersetzung mit den Lerngelegenheiten ermöglicht den Aufbau kognitiver und motivational-affektiver Kompetenzen (Angebots-Nutzungs-Modell; Helmke & Weinert, 1997).

Im vorliegenden Beitrag wird auf Grundlage der bei PISA 2012 erhobenen Daten der Frage nachgegangen, mit welchen Merkmalen sich der Mathematikunterricht in Deutschland charakterisieren lässt und inwiefern das Auftreten dieser Merkmale mit dem Erreichen mehrdimensionaler Bildungsziele einhergeht.

Eine Charakterisierung von Unterricht durch das einfache Nebeneinanderstellen als wirksam geltender Merkmale greift jedoch zu kurz. Entscheidend ist vielmehr, wie im Unterricht unterschiedliche Faktoren arrangiert werden und zusammenwirken (Oser & Baeriswyl, 2001; Seidel, Prenzel, Wittwer & Schwindt, 2007). Auf der Grundlage theoretischer Überlegungen wie auch empirischer Befunde lässt sich qualitativ hochwertiger Mathematikunterricht dadurch beschreiben, dass er kognitiv anspruchsvoll und gleichzeitig störungsarm sowie zielorientiert ist und auf die individuellen Bedürfnisse der Schülerinnen und Schüler eingeht (vgl. Schiepe-Tiska, Reiss, Obersteiner, Heine, Seidel & Prenzel, 2013). Im vorliegenden Beitrag wird deshalb versucht, anhand relevanter Merkmale Unterrichtsmuster zu identifizieren und deren Zusammenhang mit motivational-affektiven sowie kognitiven Lernergebnissen zu untersuchen.

Fragestellungen: Der Beitrag beantwortet zunächst die Frage, welche Muster des Mathematikunterrichts in Deutschland unterschieden werden können und wie häufig die identifizierten Unterrichtsmuster in deutschen Klassenzimmern vorkommen. Darüber hinaus wird der Frage nachgegangen, ob es Zusammenhänge zwischen den gefundenen Unterrichtsmustern und den motivational-affektiven und kognitiven Lernergebnissen gibt. Dabei ist vor allem von Interesse, ob ein Muster von Unterricht identifiziert werden kann, das mit beiden Lernergebnissen positiv zusammenhängt.

Methode: Die Fragestellungen werden einerseits mit Hilfe der internationalen PISA-Stichprobe untersucht, bei der sich die fünfzehnjährigen Schülerinnen und Schüler in Deutschland auf unterschiedliche Klassenstufen verteilen. Um zuverlässigere Aussagen auf Klassenebene treffen zu können, werden andererseits die im Rahmen der nationalen Ergänzung der PISA-Studie zusätzlich erhobenen Daten aus vollständigen neunten Klassen untersucht. Insgesamt bearbeiteten N = 9998 Schülerinnen und Schüler den PISA-Mathematiktest und beantworteten darüber hinaus im Rahmen des Schülerfragebogens Fragen zum Mathematikunterricht (Basisdimensionen der Unterrichtsqualität, Unterrichtspraktiken, Vertrautheit mit mathematischen Begriffen und Aufgabenstellungen) sowie zu ihren motivational-affektiven Orientierungen.

Ergebnisse: Zur Beschreibung von Unterrichtsmustern wurden Items ausgewählt, die nach den theoretischen Überlegungen und empirischen Befunden aus früheren Studien wesentliche Qualitätsmerkmale des Mathematikunterrichts erfassen sollten. Anhand dieser Items konnten mit Hilfe der Latent-Class-Analyse für die internationale Stichprobe drei Unterrichtsmuster für Deutschland identifiziert werden, die sich in ihrer Auftretenshäufigkeit unterscheiden. Mittelwertvergleiche zeigten, dass diese Muster unterschiedlich stark mit der mathematischen Kompetenz, der Freude an Mathematik sowie der mathematikbezogenen Selbstwirksamkeitserwartung zusammenhängen. Für die Stichprobe der vollständigen neunten Klassen, deren Daten derzeit ausgewertet werden, wird ein ähnliches Bild erwartet. Dabei sollten sich auch systematische Unterschiede zwischen Schularten zeigen. Die Ergebnisse werden im Vortrag vorgestellt.

Helmke, A. & Weinert, F. E. (1997). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241-251). Weinheim: Beltz.

Oser, F. & Baeriswyl, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 1031-1065). Washington, D.C.: American Educational Research Association.

Prenzel, M. (2011). Empirische Bildungsforschung morgen: Reichen unsere bisherigen Forschungsansätze aus? In M. Gläser-Zikuda, T. Seidel, C. Rohlf, A. Gröschner & S. Ziegelbauer (Hrsg.), *Mixed methods in der empirischen Bildungsforschung* (S. 273-286). Münster: Waxmann.

Schiepe-Tiska, A., Reiss, K., Obersteiner, A., Heine, J.-H., Seidel, T. & Prenzel, M. (2013). Mathematikunterricht in Deutschland: Befunde aus PISA 2012. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen* (S. 123-154). Münster: Waxmann.

Seidel, T., Prenzel, M., Wittwer, J. & Schwindt, K. (2007). Unterricht in den Naturwissenschaften. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 147-179). Münster: Waxmann.

## **Kompetenzen von Schülerinnen und Schüler mit besonderem Förderbedarf. Ergebnisse aus dem PISA 2012 Förderschul-Oversample**

Die internationalen PISA-Regelungen sehen vor, SchülerInnen mit besonderem Förderbedarf (SPF) der Zielpopulation zugehörig anzusehen, wenn sie in der Lage sind, ohne fremde Unterstützung den Test adäquat zu bearbeiten. Liegen keine körperlichen, geistigen oder emotionalen Beeinträchtigungen oder unzureichende Kenntnisse der Testsprache vor, werden SchülerInnen aus den entsprechenden Sonder- und Förderschulen sowie Inklusionsklassen in die Testung mit einbezogen. Im PISA 2012 Förderschul-Oversample wurde die Stichprobe der SchülerInnen mit SPF erstmals stark erweitert. Ziel des Forschungsprojekts ist es, die Kompetenzen dieser Schülerschaft unter Berücksichtigung der relevanten Kontextbedingungen anhand einer größeren Stichprobe nachzuzeichnen und in Beziehung zu setzen zu den Kompetenzen von SchülerInnen anderer Schularten. Von zentraler Bedeutung sind dabei Fragen bezüglich der standardisierten, reliablen und validen Kompetenzmessung, der strukturellen Vergleichbarkeit der Kompetenztestungen bei SchülerInnen mit und ohne SPF sowie der Nutzen spezifischer Testadaptationen. Nachdem bereits über das aufgabenbezogene Aufmerksamkeitsverhalten während der Testung berichtet wurde, wird im Vortrag der Frage nachgegangen, inwieweit die Verkürzung der Testzeit und die Implementation leichterer Testaufgaben die Kompetenzmessung beeinflussen und den Bedürfnissen der insgesamt heterogenen Gruppe sonderpädagogisch förderbedürftiger SchülerInnen entgegenkommen.

### Adaptionen in Large-Scale Assessments

Um bei einer Testung von SchülerInnen mit und ohne SPF valide Vergleiche zwischen den Gruppen ermöglichen zu können, sollten im Rahmen von Large-Scale Assessments die unterschiedlichen Bedingungen und besonderen Bedürfnissen der Schülerschaft einbezogen werden (Heydrich et al., 2013). Von zentraler Bedeutung ist die Berücksichtigung der durch die Behinderung bedingten Einschränkungen durch eine Korrektur konstrukt-irrelevanter Leistungshindernisse im Test. Bei diesen sogenannten Adaptationen werden die regulären Leistungstests in der Art der Präsentation, des Beantwortungsmodus, des Testablaufes oder des Testsettings modifiziert. Konstruktrelevante Merkmale wie etwa die Testschwierigkeit sollten nicht verändert werden, um die Validität des Tests nicht zu beeinflussen (Thurlow, 2010). Zentral bei der Adaptation von Tests ist, dass die dahinterliegenden Messkonstrukte nicht verändert, die besonderen Bedürfnisse präzise berücksichtigt, durch die Adaptation für die entsprechende Testgruppe differenzielle Effekte erzielt und durch den Einsatz der adaptierten Testmaterialien der Standardtestung vergleichbare Schlussfolgerungen gezogen werden (Thurlow et al., 2006). Die Befunde zu den differentiellen Effekten beim Einsatz von Adaptationen sind insgesamt jedoch uneinheitlich (Cormier et al., 2010; Thomspson et al., 2002). Konsistent sind die Ergebnisse lediglich hinsichtlich von Adaptationen des Testsettings, die für die Testung von SPF-SchülerInnen in besonderem Maß geeignet zu sein scheinen (Sireci et al., 2006). PISA 2012 Förderschul-Oversample

Um zu untersuchen, ob sich die Kompetenzen 15jähriger SchülerInnen mit SPF in PISA in Abhängigkeit der Testhefte systematisch unterscheiden, kamen im nationalen PISA 2012 Förderschul-Oversample bei 929 SchülerInnen drei verschiedene Testhefte systematisch variiert zum Einsatz: (1) das reguläre Testheft (120 Minuten Testzeit, eine Pause), (2) ein Testheft mit einfacheren Aufgaben (120 Minuten Testzeit, eine Pause), (3) ein verkürztes Testheft (sogenanntes UH-Booklet) (60 Minuten Testzeit, 4 Pausen). Die Stichprobenziehung sowie die Schätzung der Kompetenzen wurde analog zu den regulären PISA Prozeduren vorgenommen. Die Kompetenzen der SchülerInnen mit SPF besitzen die gleiche Metrik bzw. Skala wie die der regulären PISA SchülerInnen. Analysiert und berichtet werden die Befunde zur Anzahl der bearbeiteten, bzw. ausgelassenen und nicht-erreichten Aufgaben sowie die Kompetenzunterschiede zwischen den Testgruppen. Dabei wird auf der Grundlage der Beobachtungsstudie (Müller, Sälzer, Mang, Gebhardt & Prenzel, 2013) angenommen, dass sich die Adaptation des Testsettings beim UH-Testheft negativ auf die Anzahl der bearbeiteten Testaufgaben auswirkt. Der relative Anteil der on-task Verhaltensweisen bei den UH getesteten Jugendlichen in den ersten 60 Minuten der Testung lag signifikant unter dem Anteil der Kinder mit dem regulären ( $H=28.448$ ,  $p=.023^*$ ) bzw. dem einfacheren ( $H=37.499$ ,  $p=.002^{**}$ ) Testheft. Der Anteil der Kinder, die bereits in der ersten Teststunde nicht das Testmaterial, sondern von der Lehrperson zur Verfügung gestelltes Material bearbeitete, war bei den UH getesteten Kindern gegenüber den Kindern mit regulärem ( $H=-61.454$ ,  $p=.000^{**}$ ) bzw. einfacherem ( $H=-56.830$ ,  $p=.000^{**}$ ) Testheft deutlich höher.

Cormier, D., Altman, J., Shyyan, V., & Thurlow, M. D. C. (2010). A summary of the research on the effects of test accommodations: 2007-2008. Technical Report 56. Verfügbar über <http://www.cehd.umn.edu/NCEO/onlinepubs/Tech56/TechnicalReport56.pdf> (08.10.2013).

Heydrich, J., Weinert, S., Nusser, L., Artelt, C. & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: Challenges and approaches within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online (JERO)*, 5(2), S. 217-240.

Müller, K., Sälzer, S., Mang, J., Gebhardt, M. & Prenzel, M. (2013). Aufgabenbezogenes Aufmerksamkeitsverhalten von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. Verhaltensbeobachtungen in Large-scale Assessments. Vortrag auf der Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) am IPN, Kiel am 13.03.2013.

Sireci, S.G., Li, S., & Scarpati, S. (2006). The Effects of Test Accommodation on Test Performance: A Review of the Literature. Center for Educational Assessment, Research Report No. 485. Amherst, MA: School of Education, University of Massachusetts Amherst.

Thompson, S.J., Blount, A., & Thurlow, M.L. (2002). A summary of research on the effects of test accommodations—1999 through 2001. Minneapolis, MN: National Center on Educational Outcomes.

Thurlow, M. L., Thompson, S. J. & Lazarus, S. S. (2006). Considerations for the Administration of Tests to Special Needs Students: Accommodations, Modifications, and More. In: S. M. Downing & T. M. Haladyna (Ed.), *Handbook of Test Development*. New York/London: Routledge, pp. 653-673.

Thurlow, M.L. (2010). Large Scale Assessment and Accountability and Students with Special Needs. In: Peterson, P., Baker, E. & McGaw, B (Ed.), *International Encyclopedia of Education (Third Edition)*, pp 752-758.





Chair(s): Friedrich Linderkamp (Bergische Universität Wuppertal)

Diskutant/in: Friedrich Linderkamp (Bergische Universität Wuppertal)

2006 wurde auf der UNO-Generalversammlung ein Übereinkommen über die Rechte von Menschen mit Behinderungen, die sogenannte Behindertenrechtskonvention (BRK), verabschiedet. Inhaltlich wird dabei die »Achtung vor der Unterschiedlichkeit von Menschen mit Behinderungen und die Akzeptanz dieser Menschen als Teil der menschlichen Vielfalt und der Menschheit« betont. 2009 hat die Bundesrepublik Deutschland die BRK ratifiziert und sich damit verpflichtet, ihr Bildungssystem gemäß dieser Idee der „Inklusion“ umzugestalten. Wie kann dies gelingen?

Statt der Unterrichtung homogener Leistungsgruppen in einem segregierenden Schulsystem soll den Bedürfnissen einer heterogenen Schülerschaft individualisierend und differenzierend begegnet werden.

Dies stellt hohe Ansprüche an das neue inklusive Bildungssystem und zwar in struktureller Hinsicht aber auch im Hinblick auf grundlegende pädagogische Einstellungen und Haltungen sowie pädagogisches Handeln vor allem auch in den so zentralen Bereichen des Beurteilens und Unterrichtens.

Genau diese Felder deckt das Symposium mit seinen vier Einzelbeiträgen ab. Prof. Dr. Hillenbrand & Dr. Melzer dokumentieren das empirisch fassbare Aufgabenprofil von Lehrkräften der Sonderpädagogik in inklusiven Settings. Diese Studie ist von ihrer empirischen Qualität her die erste ihrer Art in Deutschland.

Prof. Dr. Wilbert, Karolina Urton und Prof. Dr. Hennemann steuern zu den sonderpädagogischen Beiträgen eine empirische Arbeit mit psychologischer Thematik bei, denn sie erfassen die Selbstwirksamkeit von Schulleitungen und deren Kollegien und setzen dies mit der Einstellung zur Inklusion in Beziehung. Die Studie beeindruckt mit einer großen Datengrundlage von 79 Schulleitungen und 501 Lehrkräften.

Prof. Dr. Grosche kann anschließend zeigen, wie Leistungsbeurteilungen von Schülern auch nicht-kategorial und ohne soziale Bezugsnorm dennoch mit hoher methodischer Güte lernverlaufsdiagnostisch möglich sind. Prof. Dr. Huber wird schließlich auf Grundlage eines PC-gestützten Experimentaldesigns und mit einer beachtlichen Stichprobe von 480 Schülern erstmals empirisch fundiert zeigen, dass das Lehrkraftfeedback einen signifikanten und ursächlichen Einfluss auf soziale Integrationsprozesse bei Grundschulkindern nehmen kann. Mithin bietet das Symposium eine hohe thematische und forschungsmethodische Vielfalt und nimmt inklusionsrelevante Themen sowohl sonderpädagogisch als auch psychologisch kritisch in den Blick.

- Avolio, B.J., Sosik, J.J., Jung, D.I. & Berson, Y. (2003). Leadership models, methods and applications: Small steps and giant leaps. In W.C. Borman, R.J. Klimoski, D.J. Ilgen & I. B. Weiner (Eds.), *Handbook of Psychology: Industrial and Organizational Psychology* (Vol. 12, pp. 277-307). New York: John Wiley & Sons.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter et al. (Hrsg.), *Professionelle Kompetenz von Lehrkräften* (S. 29-53). Münster u.a.: Waxmann.
- Beretvas, S. N., & Chung, H. (2008). An evaluation of modified R<sup>2</sup>-change effect size indices for single-subject experimental designs. *Evidence-Based Communication Assessment and Intervention*, 2 (3), 120-128.
- Bless, G. & Mohr, K. (2007). Die Effekte von Sonderunterricht und gemeinsamem Unterricht auf die Entwicklung von Kindern mit Lernbehinderungen. In J. Borchert, H. Goetze, J. Walter & F. B. Wember (Hrsg.), *Sonderpädagogik des Lernens* (S. 375-383). Göttingen [u.a.]: Hogrefe.
- Chang, L. (2003). Variable Effects of Children's Aggression, Social Withdrawal, and Prosocial Leadership as Functions of Teacher Beliefs and Behaviors. *Child Development*, 74 (2), 535-548.
- Christ, T. J. (2006). Short-term estimates of growth using curriculum-based measurement of oral reading fluency: estimating standard error of the slope to construct confidence intervals. *School Psychology Review*, 35 (1), 128-133.
- Döbert, H. & Weißhaupt, H. (Hrsg.) (2013). *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. Münster: Waxmann/European
- Feinman, S. (1992). Social Referencing and Conformity. In S. Feinman (Hrsg.), *Social referencing and the social construction of reality in infancy* (S. 229-268). New York: Plenum Press.
- Franzkowiak, T. (2012). *Meine Schüler, deine Schüler, unsere Schüler. Welche Kompetenzen erwarten Grundschullehrerinnen von den sonderpädagogischen Lehrkräften im Gemeinsamen Unterricht? Gemeinsam leben*, 20 (1), 12-19.
- Grosche, M. (2011). Effekte einer direkt-instruktiven Förderung der Lesegenauigkeit. *Empirische Sonderpädagogik*, 3 (2), 147-161.
- Grosche, M., & Volpe, R. J. (2013). Response-to-intervention (RTI) as a model to facilitate inclusion for students with learning and behaviour problems. *European Journal of Special Needs Education*, 28 (3), 254-269.
- Grünke, M. (Hrsg.). (2012). *Kontrollierte Einzelfallforschung. Empirische Sonderpädagogik, Themenheft*, 3 (3/4), 211-347.
- Heinrich, M., Urban, M. & Werning, R. (2013). Grundlagen, Handlungsstrategien und Forschungsperspektiven für die Ausbildung und Professionalisierung von Fachkräften für inklusive Schulen. In H. Döbert & H. Weißhaupt (Hrsg.), *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. (S. 69-134). Münster: Waxmann.
- Hillenbrand, C., Melzer, C. & Hagen, T. (2013). Bildung schulischer Fachkräfte für inklusive Bildungssysteme. In H. Döbert & H. Weißhaupt (Hrsg.), *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. (S. 33-68). Münster: Waxmann.
- Huber, C. & Wilbert, J. (2012). Soziale Ausgrenzung von Schülern mit sonderpädagogischem Förderbedarf und niedrigen Schulleistungen im gemeinsamen Unterricht. *Empirische Sonderpädagogik* (2), 147-165.
- INTASC (2011). *Model standards for beginning teacher licensing and development: A resource for state dialogue*. Washington: Council of Chief State Officers. Verfügbar unter: [http://www.ccsso.org/Resources/Publications/INTASC\\_Model\\_Core\\_Teaching\\_Standards\\_A\\_Resource\\_for\\_State\\_Dialogue\\_%28April\\_2011%29.html](http://www.ccsso.org/Resources/Publications/INTASC_Model_Core_Teaching_Standards_A_Resource_for_State_Dialogue_%28April_2011%29.html) [05.01.13].
- Jansen, F., Streit, U., & Fuchs, A. (2007). Lesen und Rechtschreiben lernen nach dem IntraActPlus-Konzept. Heidelberg: Springer.
- Kavale, K. A. & Forness, S. R. (May, 1996). Social skill deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities*, 29 (3), 226-237. Verfügbar unter <http://dx.doi.org/10.1177/002221949602900301>.
- Kern, H. J. (1997). *Einzelfallforschung*. Weinheim: Beltz.
- Klauer, K. J. (2006). Erfassung des Lernfortschritts durch curriculumbasierte Messung. *Heilpädagogische Forschung*, 32 (1), 16-26.
- McAuliffe, M. D., Hubbard, J. A. & Romano, L. J. (Jul, 2009). The role of teacher cognition and behavior in children's peer relations. *Journal of Abnormal Child Psychology: An official publication of the International Society for Research in Child and Adolescent Psychopathology*, 37 (5), 665-677.
- Melzer, C. & Hillenbrand, C. (2013a). Aufgaben sonderpädagogischer Lehrkräfte für die inklusive Bildung: empirische Befunde internationaler Studien. *Zeitschrift für Heilpädagogik*, 64 (5), S. 194-202.
- Melzer, C. & Hillenbrand, C. (2013b). Aufgabenverteilung von pädagogischen Fachkräften in inklusiv arbeitenden Grundschulen Modellprojekt „ISI“. *Forschungsbericht*. Universität Oldenburg, eingereicht am 05.08.2013 (70 Seiten).
- Moser, V. (2013). *Professionsforschung als Unterrichtsforschung*. H. Döbert & H. Weißhaupt (Hrsg.), *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. (S. 135-146). Münster: Waxmann.
- Moser, V., Schäfer, L. & Redlich, H. (2011). Kompetenzen und Beliefs von Förderschullehrkräften in inklusiven Settings. In: B. Lütje-Klose, M. T. Langner, B. Serke & M. Urban (Hrsg.), *Inklusion in Bildungsinstitutionen. Eine Herausforderung an die Heil- und Sonderpädagogik*. (S. 143-149). Bad Heilbrunn: Klinkhardt.
- Reiser, H. (2001). *Untersuchungen zu den Aufgaben und dem Berufsbild von Sonderschullehrerinnen im Zentrum für Erziehungshilfe der Stadt Frankfurt am Main. Behindertenpädagogik*, 40 (4), 471-494.
- Ricken, G. (Hrsg.) (2011). *Trends in der sonderpädagogischen Diagnostik. Empirische Sonderpädagogik, Themenheft*, 3 (3), 188-271.
- Terhart, E. (2010). *Heterogenität der Schüler – Professionalität der Lehrer: Ansprüche und Wirklichkeiten*. In S. L. Ellgar-Rüttgardt & G. Wachtel (Hrsg.), *Pädagogische Professionalität und Behinderung. Herausforderungen aus historischer, nationaler und internationaler Perspektive* (S. 89-104). Stuttgart: Kohlhammer.
- Webster, M. & Foschi, M. (1992). *Social Referencing and Theories of Status and Social Interaction*. In S. Feinman (Hrsg.), *Social referencing and the social construction of reality in infancy* (S. 269-296). New York: Plenum Press.
- Weiß, S., Kollmannsberger, M. & Kiel, E. (2013). Sind Förderschullehrkräfte anders? Eine vergleichende Einschätzung von Expertinnen und Experten aus Regel- und Förderschulen. *Empirische Sonderpädagogik*, (2), S. 167-186.
- White, K. J. & Jones, K. (2000). Effects of Teacher Feedback on the Reputations and Peer Perceptions of Children with Behavior Problems. *Journal of Experimental Child Psychology*, 76 (4), 302-326.

Dienstag,

04.03.2014,

14:15 - 16:00,

IG-Farben-Haus

Raum 311



## **Aufgaben sonderpädagogischer Fachkräfte in inklusiven Bildungssystemen**

Im Zuge des Aufbaus eines umfassenden inklusiven Bildungssystems verändert sich die berufliche Realität für alle Lehrkräfte. In aktuellen Expertisen werden die veränderten Aufgaben für pädagogische Fachkräfte in inklusiven Settings belegt (Döbert & Weishaupt, 2013), insbesondere eine notwendige Revision der Lehrerbildung, wie sie in zwei Gutachten für das Deutsche Institut für Internationale Pädagogische Forschung entwickelt werden (Heinrich, Urban, & Werning, 2013; Hillenbrand, Melzer, & Hagen, 2013) erscheint unumgänglich. Grundlage hierfür sollten die empirisch nachweisbaren Aufgaben von Lehrkräften in inklusiven Settings sein, um eine Passung zwischen Lehrerbildung und praktischen Anforderungen herzustellen. Hier besteht ein gravierendes Forschungsdefizit, das aus unterschiedlichen theoretischen Positionen beschrieben wird (Hillenbrand, Melzer, & Hagen, 2013, S. 44, Moser, 2013, S. 143; Weiß, Kollmannsberger, & Kiel, 2013, S. 170). Das Aufgabenprofil sowohl von sonderpädagogischen Fachkräften wie auch von Lehrkräften der allgemeinen Schulen im Kontext inklusiver Settings ist in Deutschland bislang empirisch weitgehend unerforscht (Melzer & Hillenbrand, 2013, S. 195; Moser, Schäfer, & Redlich, 2011, S. 144), die Diskussion basiert v.a. auf Erfahrungsberichten aus Integrationsversuchen oder theoretischen Ableitungen (Melzer & Hillenbrand, 2013a, S. 194). Lediglich zwei empirische Studien liegen vor: eine Befragung mittels Interview (drei Lehrkräfte, drei Sozialpädagogen) (Reiser, 2001) und eine Befragung zu den Erwartungen von 21 Grundschullehrkräften an sonderpädagogische Fachkräfte (Franzkowiak, 2012). Eine aktuelle Studie thematisiert Kompetenzprofile von Lehrkräften allgemeiner Schulen sowie der Förderschule durch Gruppendiskussionen (N=220), allerdings ohne direkten Bezug zur Inklusion (Weiß, Kollmannsberger, & Kiel, 2013). Die internationale Befundlage ist deutlich besser einzuschätzen: In einem Literaturreview (Melzer & Hillenbrand, 2013) konnten insgesamt 14 empirische Studien ermittelt werden, die insgesamt 57 verschiedene Aufgaben in elf Aufgabenbereichen identifizieren (ebd., S. 197), die überwiegend qualitativ gewonnen wurden. Es fehlt folglich auch international an quantitativ-empirischen Studien. Die theoretische Grundlage der empirischen Forschung zu Aufgaben von Lehrkräften in inklusiven Settings kann auf die vorliegenden Professionalisierungstheorien zurückgreifen: Lehrkräfte üben eine Profession aus, die mittels kompetenztheoretischer Professionalisierungsmodelle beschrieben werden kann (Terhart, 2010, S. 91). Für die vorzustellenden, empirisch-quantitativen Untersuchungen bildet das in empirischen Studien bereits bewährte COACTIV-Modell von Baumert und Kunter (2011, S. 32) eine geeignete Theoriebasis. Professionelle Handlungskompetenz gliedert sich demnach in Professionswissen, Überzeugungen und Werthaltungen, motivationale Orientierungen und selbstregulative Fähigkeiten. Das Professionswissen wiederum setzt sich aus Fachwissen, fachdidaktischem Wissen (diagnostische Kompetenz, Erklärungswissen), Beratungswissen, pädagogischem Wissen (Bezugsnormorientierung, Wissen über Lernprozesse) und Organisations- sowie Interaktionswissen zusammen. Das Professionswissen von Lehrkräften in inklusiven Bildungssystemen, insbesondere von sonderpädagogischen Fachkräften, erfordert veränderte Inhalte zur Bewältigung der Aufgaben: Welche Aufgaben bewältigen Lehrkräfte für Sonderpädagogik in inklusiven Settings? Auf der Basis des COACTIV-Modells und der internationalen Befunde berichtet der Beitrag aus einer landesweiten Erhebung und dokumentiert das empirisch fassbare Aufgabenprofil von Lehrkräften der Sonderpädagogik in inklusiven Settings. Die Studie erhebt die Einschätzung von sonderpädagogischen Lehrkräften in inklusiven Settings mittels eines Surveys im Bundesland Hessen. Das Erhebungsinstrument umfasst Aufgabenbeschreibungen aus allen elf Bereichen des internationalen Literaturreviews, deren Relevanz die Befragten einschätzen sollen. Es wurde in zwei kleineren Pilotstudien (sonderpädagogische Fachkräfte im mobilen Dienst Niedersachsen, N=65; sonderpädagogische Fachkräfte inklusiver Grundschulen in Berlin, N=85) überprüft (Melzer & Hillenbrand, 2013b). Das Instrument belegt mit Hilfe einer Klassenanalyse Zusammenhänge zwischen der Einschätzung der Aufgaben und dem Einsatzfeld der Befragten und konnte auf dieser Basis weiterentwickelt werden. Das revidierte Instrument dient in der vorzustellenden Studie, durchgeführt in Kooperation mit dem Hessischen Kultusministerium, der Ermittlung des Aufgabenprofils von sonderpädagogischen Lehrkräften in inklusiven Settings. Die Studie basiert auf einer Grundgesamtheit von N=1600 Befragten und wird aktuell durchgeführt. Der Symposiumsbeitrag stellt, neben einem Überblick über die deskriptiven Daten, wichtige Ergebnisse der Faktorenanalyse vor und zeigt relevante Subgruppenunterschiede (z.B. Einsatzorte und -formen) auf. Die Konsequenzen für die Lehrerbildung in allen drei Phasen werden diskutiert.

## ***Der Zusammenhang zwischen der Einstellung zur Inklusion und der Selbstwirksamkeit von Schulleitungen und deren Kollegien***

Der Weg hin zu einem inklusiven Schulsystem ist ein Transformationsprozess, der grundlegende Veränderung der schulischen Organisationsstrukturen erfordert. In diesem Zusammenhang stellen die psychologischen Variablen Einstellung, Überzeugung und das Belastungserleben der Lehrerinnen und Lehrer und deren Schulleitungen wesentliche Aspekte der professionellen Handlungskompetenz dar (Baumert & Kunter, 2006; Schaarschmidt, 2005) und tragen maßgeblich zum Gelingen einer inklusiver Schulentwicklung bei (Savolainen, Engelbrecht, Nel, & Malinen, 2012). Untersuchungen konnten zeigen, dass die individuelle Selbstwirksamkeit und die kollektive Wirksamkeit die Einstellung von Lehrkräften zur Inklusion positiv beeinflussen (u.a. Brownell & Pajares, 1999; Jordan, Kircaali-Iftar, & Diamond, 1993; Leyser, Zeiger, & Romi, 2011). Darüber hinaus ist die (Selbst-)Wirksamkeitsüberzeugung von Lehrerinnen und Lehrern auch als eine bedeutsame personale Ressource für den Umgang mit Belastungen zu sehen (Brouwers, Evers, & Tomic, 2001; Schwarzer & Warner, 2011). Es lässt sich vermuten, dass das Belastungserleben auf Individuums- wie auch Kollegiumsebene ein wichtiger Faktor für eine gelingende inklusive Schulentwicklung darstellt, da eine hohe schulische Qualität nur von Lehrkräften gewährleistet werden kann, die sich durch Zufriedenheit, Engagement und Widerstandsfähigkeit auszeichnen (Schaarschmidt, 2005).

Schulentwicklungsprozesse erfordern Veränderungen auf Gruppenebene deren Entwicklung und Steuerung eine zentrale Aufgabe von Schulleitungen ist (Avolio, Sosik, Jung, & Berson, 2003; Edmonds, 1979, Rost, 1995). Eine höhere Selbstwirksamkeit der Schulleitung führt demnach zur besseren Bewältigung von Changeprozessen im Gesamtkollegium, was sich wiederum positiv auf die Schülerinnen und Schüler auswirkt (Dimmock & Hattie, 1996).

Vor diesem Hintergrund geht die vorliegende Studie der Frage nach, welche Rolle Schulleitungen im aktuellen Transformationsprozess hin zu einem inklusiven Schulsystem spielen. Genauer soll zunächst der Frage nachgegangen werden, inwiefern es a) bedeutsame Unterschiede zwischen den Kollegien hinsichtlich der Einstellungen gibt und ob b) diese Unterschiede mit den jeweiligen Einstellungen und Haltungen der Schulleitungen korrespondieren.

Dazu wurden Daten an einer Stichprobe von 79 Schulleitungen und deren Kollegien (N = 501) in inklusiven Primarstufen erhoben. Beide Gruppen wurden mit den gleichen standardisierter Fragebögen zur Einstellung zur Inklusion, Selbstwirksamkeit und kollektiven Wirksamkeit, Arbeitsbelastung und Burn-out Risiko befragt. Die Analysen erfolgten mittels Mehrebenenmodellierungen um der Datenstruktur zu entsprechen.

Die Ergebnisse zeigen, dass sich Lehrerkollegien hinsichtlich der Einstellung zur Inklusion sowie der individuellen und kollektiven (Selbst-)wirksamkeit unterscheiden. Entsprechend unserer Annahmen zeigt sich eine bedeutsame Varianzaufklärung der Einstellungen zur Inklusion sowie der kollektiven Wirksamkeit der Kollegien durch die Schulleitung. Ein Zusammenhang zwischen dem Burnout-Risiko, der Kollegiumszugehörigkeit und der Schulleitung ließ sich hingegen nicht abbilden.

## **Auswertung individueller Lernverläufe zum Treffen von Förderentscheidungen**

Im Zuge der Umsetzung von Inklusion müssen Lehrkräfte ihren Unterricht mehr denn je differenzieren und individualisieren. Dies geschieht am zuverlässigsten auf der Basis einer fundierten Diagnostik. Eine besonders geeignete Form zur Verknüpfung von Diagnostik, Förderung und Evaluation ist die Lernverlaufsdiagnostik, die prozessbezogen statt summativ-klassifizierend und entwicklungs sensitiv statt statusorientiert ist (Klauer, 2006; Ricken, 2011). Dazu finden in relativ kurzen Abständen (2-4 Mal pro Monat) regelmäßige Datenerhebungen mittels sehr kurzer Paralleltests (1-3 Min.) statt, um individuelle Lernverläufe von Schülerinnen und Schülern zu erfassen.

Bislang liegen nur wenige früh einsetzbare Verfahren zur Lernverlaufsdiagnostik vor (Grosche, 2011). Auch die Auswertung steht aufgrund großer Messfehler noch vor Schwierigkeiten (Christ, 2006). Daher entwickelten wir einen Lernverlaufstest für das synthetisierende Lesen, dessen Verläufe durch hierarchische Regressionen für Einzelanalysen modelliert werden (Beretvas & Chung, 2008). Die vorliegende Studie soll zeigen, ob durch diesen Test anfängliche Probleme im Lesenlernen frühzeitig, lernprozessbegleitend und veränderungssensitiv erfasst werden können.

In drei ersten Klassen identifizierten wir durch den Lernverlaufstest vier Kinder mit besonders schwachen Leseleistungen. Sie erhielten eine acht-wöchige direkt-instruktive Leseförderung durch das Programm Intra-Act-Plus (Jansen, Streit, & Fuchs, 2007). Die Effektivität der Förderung wurde ebenfalls durch den Lernverlaufstest überprüft. Als Forschungsdesign verwendeten wir einen multiplen Grundraten-Versuchsplan. Die experimentelle Absicherung der Leseförderung erfolgte pro Kind über drei verschiedene Test- und Fördermaterialien (Kern, 1997). Insgesamt liegen pro Kind zwischen 20 und 30 Messzeitpunkte vor. Ausgewertet werden die Lernverläufe durch verschiedene deskriptive und inferenzstatistische Verfahren (Grosche, 2011; Grünke, 2012).

Die Ergebnisse zeigen, dass die Lernverlaufsdiagnostik des synthetisierenden Lesens reliabel (split-half .91; Retest 88.  $< r > .91$ ) und valide ist (Korrelation zu WLLP  $r = .69$ ). Ebenso sind starke Lernzuwächse änderungssensitiv nachweisbar ( $.87 < \text{PAND} > .95$ ;  $51 < \text{Phi}^2 > 80$ ;  $.20 < d > .70$ ), die jedoch nicht eindeutig kausal auf die Intervention zurückzuführen sind. Zwar konnten nach acht Wochen Förderung die Leseprobleme noch nicht vollständig aufgelöst werden, aber Scheren- bzw. Matthäus-Effekte ließen sich deutlich verringern ( $-1.4 > z < -0.9$ ). Damit kann die Studie zeigen, dass die Lernverlaufsdiagnostik im Zuge der Umsetzung von Inklusion großes Potential hat, Diagnostik und Förderung enger miteinander zu verknüpfen (Grosche & Volpe, 2013).

## **Wie das Lehrkraftfeedback die soziale Akzeptanz von Grundschulkindern beeinflussen könnte – Ergebnisse einer Experimentalstudie**

Die ungünstige soziale Integration von Schülern mit sonderpädagogischem Förderbedarf im Gemeinsamen Unterricht ist ein bislang ungelöstes Problem der Integrationspädagogik. In der Vergangenheit zeigten verschiedene Studien, dass Schüler mit sonderpädagogischem Förderbedarf in ihren Klassen signifikant seltener beliebt und häufiger abgelehnt werden als ihre Klassenkameraden ohne besonderen Förderbedarf (Bless & Mohr, 2007; Huber & Wilbert, 2012; Kavale & Forness, 1996).

Internationale Ansätze untersuchten in der Vergangenheit die Hypothese, inwieweit soziale Referenzierungsprozesse in der Schule die soziale Akzeptanz von Schulkindern beeinflussen könnten (Feinman, 1992; Webster & Foschi, 1992). Im Mittelpunkt steht dabei das Feedback, das eine Lehrkraft ihren Schülerinnen und Schülern im Unterricht gibt. Dabei wird angenommen, dass positives Lehrkraftfeedback ursächlich die Chance auf eine positive soziale Akzeptanz in der Klasse erhöht. Im Umkehrschluss wird davon ausgegangen, dass negatives Lehrkraftfeedback die Chance auf eine gute soziale Akzeptanz verringert. Erste Studien im englischen Sprachraum konnten in der Vergangenheit signifikante Effekte des Lehrkraftfeedbacks auf die soziale Akzeptanz von Schulkindern im Allgemeinen und von Schulkindern mit sonderpädagogischem Förderbedarf im Besonderen nachweisen (Chang, 2003; McAuliffe, Hubbard, & Romano, 2009; White & Jones, 2000). Eine kritische Würdigung der bislang vorliegenden Studien zeigen jedoch zum Teil erhebliche Mängel auf. So sind aus den bislang verwendeten Forschungsdesigns keine validen Aussagen zur Richtung von Ursache und Wirkung der Untersuchungsvariablen ableitbar. Das hier dargestellte Forschungsprojekt knüpft an dieser Forschungslücke an.

Im Vortrag sollen erste Befunde eines Projektes zum Einfluss des Lehrkraftfeedbacks auf die soziale Akzeptanz bei Grundschulkindern vorgestellt werden. Es wird angenommen, dass das Lehrkraftfeedback eine zentrale soziale Referenz für Grundschulkindern in der Schule darstellt. Im Rahmen des Projektes wurde ein PC-gestütztes Experimentaldesign entwickelt. Im Mittelpunkt steht dabei ein Pre-Posttest-Design. Mit Hilfe eines computergestützten Experiments wurde bei N=480 Probanden (Klasse 3 und 4) aus Deutschland und Österreich untersucht, inwieweit sich die soziale Akzeptanz von vier fiktiven Grundschulkindern durch eine systematische Manipulation des Lehrkraftfeedbacks beeinflussen lässt. Hierzu wurden den Kindern zunächst vier Schülerinnen (bei weiblichen Probanden) und Schüler (bei männlichen Probanden) mit einem Foto vorgestellt. In einem Pretest wurde die spontane soziale Akzeptanz der vier fiktiven Kinder gemessen. Es folgte eine kurze Vorstellung einer fiktiven Lehrkraft und schließlich die experimentelle Manipulation. Den vier fiktiven Schulkindern wurde dabei zufällig ein Lehrkraftfeedback (positiv / negativ) zugeordnet. Als Kontrollvariable diente eine ebenfalls zufällig variierte Information darüber, wieviel Spaß es macht, mit dem betreffenden fiktiven Schüler / der betreffenden Schülerin das Kartenspiel UNO zu spielen. Die Kontrollvariable wurde ebenfalls zweifach unterteilt (viel Spaß / keinen Spaß). Am Ende des Experiments erfolgte jeweils eine Posttestmessung der sozialen Akzeptanz. Die Reihenfolge der präsentierten Schülerinformationen wurde zufällig variiert. Alle Informationen wurden den Schülern über Kopfhörer vorgelesen und konnten gleichzeitig mitgelesen werden. Die Effekte des Lehrkraftfeedbacks wurden mit der Kontrollvariable verglichen. Das Experiment entspricht damit einem 2 x 2 Versuchsplan. Die Datenanalyse erfolgte mit einer 2x2 MANOVA mit Meßwiederholung. Als Innensubjektfaktoren fließen die vier Schülertypen in die Datenanalyse ein. Geschlecht und Alter der Schüler wurden als Zwischensubjektfaktoren verwendet.

Die Ergebnisse zeigen signifikante Haupteffekte für das Lehrkraftfeedback und die Kontrollvariable, wobei für das Lehrkraftfeedback ein signifikant größerer Einfluss festgestellt wurde. Darüber hinaus wurden signifikante Interaktionseffekte für das Alter und das Geschlecht der Probanden gefunden. Die Befunde zeigen erstmals das Lehrkraftfeedback einen signifikanten und ursächlichen Einfluss auf soziale Integrationsprozesse bei Grundschulkindern nehmen könnte. Die größeren Effektstärken des Lehrkraftfeedbacks gegenüber der Kontrollvariable zeigen zudem, dass dem Rückmeldeverhalten von Lehrerinnen und Lehrern eine bedeutsamere Rolle zukommen könnte als anderen positive oder negative Informationen über ein Schulkind. Die Befunde der Studie werden vorgestellt und vor dem Hintergrund der aktuellen Inklusionsdebatte diskutiert. Ansatzpunkte für die Förderung sozialer Integration in der schulischen Praxis werden abgeleitet.



## **Validität**

Chair(s): Steffen Brandt (Art of Reduction), Andreas Frey (Friedrich-Schiller-Universität Jena), Johannes Hartig (DIPF Frankfurt)

Diskutant/in: Knut Schwippert (Universität Hamburg)

Außerhalb Deutschlands ist das Thema der Validität in den vergangenen Jahren wieder deutlicher in den Fokus der wissenschaftlichen Diskussion gekommen. Wahrnehmbar ist dies sowohl durch Symposien, die in den letzten Jahren verstärkt zu diesem Thema auf internationalen Tagungen (etwa auf den Tagungen des National Council of Measurement in Education oder der Psychometric Society) veranstaltet wurden als auch durch eine erhöhte Anzahl von Veröffentlichungen, zuletzt zum Beispiel durch das Sonderheft "Validity" des Journal of Educational Measurement.

Ein Hauptanliegen dieses Symposiums ist es daher, diesem aus unserer Sicht eminent wichtigen Thema auch im deutschsprachigen Raum, in dem Validität vielfach immer noch ausschließlich durch Analyse korrelativer Zusammenhänge untersucht wird, eine größere Aufmerksamkeit zuteil werden zu lassen. Die korrelative Analyse von Validität steht im Gegensatz zu einem moderneren, auf Testwertinterpretationen bezogenen und argumentationsbasierten Konzept der Validität, wie es auch in den für Ende dieses Jahres angekündigten Standards for Educational and Psychological Testing vertreten wird, die gemeinsam von international führenden Fachgesellschaften der Erziehungswissenschaft und der Psychologie (AERA, NCME, APA) herausgegeben wird. Die Beiträge des Symposiums greifen den aktuellen Forschungsstand zum Thema Validität auf und widmen sich dieser dabei aus unterschiedlichen Perspektiven.

Der erste Beitrag von Andreas Frey stellt hierbei den aktuell auf internationaler Ebene geteilte Verständnis von Validität dar und geht auf empirische Weise der Frage nach inwieweit dieses Verständnis derzeit in Deutschland bei der Rezension von Testverfahren aufgegriffen wird.

Der zweite Beitrag von Steffen Brandt betrachtet psychometrische Aspekte der Validität. Hier wird anhand simulierter Daten diskutiert, inwieweit eine eindimensionale Testkonstruktion auf Basis der Item-Response-Theorie eine valide Basis für eine mehrdimensionale Interpretation des Tests sein kann.

Der dritte Beitrag von Jakob Mainert und Samuel Greiff gibt ein Beispiel für die Untersuchung der Validität der Interpretation von Kompetenztestwerten zum komplexen Problemlösen unter Rückgriff auf die klassischen Konzepte der konvergenten, divergenten und prädiktiven Validität.

Der vierte Beitrag von Felix Walker, Stephan Abele, Nico Link, Tobias Gschwendtner und Reinhold Nickolaus gibt ein weiteres Beispiel zur empirischen Untersuchung der Validität von Kompetenztestwert-Interpretationen. Hier jedoch werden die Ergebnisse zur Konstruktvalidität zweier Kompetenztests für KFZ-Mechatroniker/-innen bzw. Elektroniker/-innen dargestellt.

Nach der Präsentation der geschilderten Beiträge, die sich dem Thema Validität aus den Disziplinen Erziehungswissenschaft, Psychologie und Berufs- und Wirtschaftspädagogik nähern, werden diese von Knut Schwippert zusammenfassend diskutiert.

Dienstag,

04.03.2014,

14:15 - 16:00,

PEG-Gebäude

Raum 1.G168

## Validität: Internationaler Forschungsstand und Umsetzung in Deutschland

### Theoretischer Hintergrund

Zum Ende des Jahres 2013 ist die sechste Auflage der gemeinsam von international führenden Fachgesellschaften der Erziehungswissenschaft und der Psychologie (AERA, NCME, APA) herausgegebenen Standards for Educational and Psychological Testing (nachfolgend: „Test-Standards“) angekündigt. Aus diesem Anlass wird hinterfragt, inwieweit die Empfehlungen der vorherigen Auflage in Deutschland umgesetzt werden. Dabei wird auf das Konzept der Validität fokussiert, da die gültige Interpretation von Testergebnissen für alle in der empirischen Bildungsforschung aktiven Fachdisziplinen von hoher Relevanz ist.

Seit den ersten Konzeptualisierungen zu Beginn des 20. Jahrhunderts hat das Verständnis der Validität eine Metamorphose durchlaufen. Während zunächst die Vorhersage von Verhaltenskriterien im Vordergrund stand, wurden bald darauf mehrere Validitätsarten getrennt betrachtet. Validität wurde dabei als Eigenschaft eines Testverfahrens angesehen. Als Meilenstein der Konzeptualisierung der Validität ist der von Cronbach und Meehl (1955) beschriebene Ansatz der Konstruktvalidität anzusehen, der Validitätsuntersuchungen in den hypothetisch-deduktiven Ansatz der Erkenntnisgewinnung einordnet. Die von dieser Arbeit angestoßenen Diskussionen mündeten in dem Validitätskonzept der vierten Auflage der Test-Standards. Validität wird nun nicht mehr als Testeigenschaft angesehen, sondern als die Angemessenheit, Sinnhaftigkeit und Nützlichkeit von Testwertinterpretationen. Zudem wird Konstruktvalidität als übergeordnetes Konzept verstanden, dem die verschiedenen Validitätsarten untergeordnet sind. Diese Sichtweise erwies sich in den Folgejahren als stabil und ist auch in der fünften Auflage der Test-Standards zu finden und für die sechste Auflage zu erwarten. Eine kompatible Sichtweise wird im Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologievereinigungen (Testkuratorium, 2010) formuliert. Jonson und Plake (1998) untersuchten inwieweit die Änderungen des Validitätskonzepts praktisch umgesetzt werden. Bei ihrer Analyse publizierter Rezensionen englischsprachiger Tests fanden sie, dass die Änderungen zwar in die Praxis übergehen, dass dies aber mit zeitlicher Verzögerung geschieht.

Fragestellung: Es wird untersucht, in welchem Ausmaß das seitens der aktuellen Test-Standards (AERA, APA & NCME, 1999) formulierte Verständnis von Validität im deutschsprachigen Bereich genutzt wird.

Methode: Zur Beantwortung der Fragestellung wurde auf das vom Leibniz-Zentrum für Psychologische Information und Dokumentation (2013) zusammengestellte Verzeichnis von Testrezensionen zurückgegriffen. Dieses umfasst Tests, Skalen und Fragebögen aus der Psychologie, pädagogische Verfahren einschließlich Schulleistungstests sowie psychiatrische Verfahren. Es wurden alle Rezensionen deutschsprachiger Instrumente aus den für die empirische Bildungsforschung primär relevanten Inhaltsbereichen Entwicklungstests, Intelligenztests, Leistungs-, Fähigkeits- und Eignungstest, Schulleistungstests, Interessentests und Persönlichkeitstests ab dem Jahr 1980 analysiert. Die Passung zwischen dem in den Rezensionen verwendeten Validitätskonzept und dem der Test-Standards von 1999 wurde mit einem standardisierten Bewertungsschema eingestuft.

Ergebnisse: Die Analysen sind momentan noch nicht ganz abgeschlossen, werden dies aber bis zur Konferenz sein. Vorläufige Befunde weisen darauf hin, dass bei der Besprechung deutschsprachiger Tests und Fragebögen ein Validitätsverständnis vorherrscht, das dem internationalen Stand vor 1985 entspricht. Insbesondere wird Validität zumeist als Testeigenschaft angesehen und nicht als Gültigkeit von Testwertinterpretationen. Über die Zeitspanne von 1980 bis 2013 zeigen sich nur geringfügige Veränderungen. Auf der Konferenz werden die abschließenden Analysen zum Zeitverlauf und zum Vergleich der genutzten Validitätskonzepte zwischen Inhaltsbereichen berichtet.

Diskussion: Die Befunde weisen auf Nachholbedarf im Bereich der Validität hin. Das in Deutschland vorherrschend genutzte Validitätsverständnis hinkt dem internationalen Stand um circa 30 Jahre hinterher.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281–302.

Jonson, J. L., & Plake, B. S. (1998). A Historical Comparison of Validity Standards and Validity Practices. Educational and Psychological Measurement, 58, 736-753.

Testkuratorium (2010). TBS-TK - Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologener Vereinigungen. Revidierte Fassung vom 09. September 2009. Psychologische Rundschau, 61, 52-56.

Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). (Hrsg.). (2013). Verzeichnis Testverfahren. Kurznamen. Langnamen. Autoren. Testrezensionen (20., aktualisierte Auflage). Trier: ZPID.

## Validität von ein- und mehrdimensionalen Auswertungen bei PISA

### Theoretischer Hintergrund

Die Entscheidung, ob ein Datensatz ein- oder mehrdimensional auszuwerten ist, wird in der Regel anhand von Kennwerten zur Modellpassung gefällt. Im Rahmen der Item Response Theorie betrachtet man hierzu Kennwerte wie den AIC, BIC, oder CAIC (Akaike, 1974; Schwarz, 1978; Bozdogan, 1987), die Modellpassung und Stichprobengröße in unterschiedlichem Ausmaß miteinander verrechnen und gewichten. Nach der Testentwicklung und der Durchführung des Haupttests kann dann anhand dieser Indizes analysiert werden, ob der Test die angenommene Dimensionalität aufweist. Insbesondere im Rahmen von groß angelegten Studien wie etwa dem Programme for International Student Assessment (PISA) führt dieses Standardvorgehen jedoch zu einem Konflikt, der in der Regel wenig beachtet wird, da er als eher theoretischer Natur betrachtet wird. Es ist bekannt, dass Datensätze im Rahmen von PISA nicht nur ein- oder mehrdimensional ausgewertet werden, sondern auch ein- und mehrdimensional. So wird etwa für Mathematik nicht nur ein eindimensionaler Kompetenzwert berichtet, sondern basierend auf dem gleichen Datensatz werden auch mehrdimensionale Kompetenzwerte für die mathematischen Inhaltsbereiche "Quantität", "Veränderung und Beziehungen", "Raum und Form" und "Unsicherheit" berichtet (OECD, 2013). Aus analytischer Sicht führt diese pragmatische Herangehensweise, die es ermöglicht die benötigten Leistungswerte zu berechnen, zu der Konsequenz, dass sowohl Ergebnisse der oben erwähnten Kennwerte zum Modellfit als auch existierende, oder zumindest implizit angenommene lokale Abhängigkeiten von Items innerhalb eines Inhaltsbereichs außer Acht gelassen werden. In den vergangenen Jahren wurde daher vermehrt versucht, hierarchische oder auch Testlet-Modelle anzuwenden, um solche Abhängigkeiten zu berücksichtigen (Wainer, Bradlow, & Wang, 2007; Gibbons et al., 2007; Brandt & Ducker, 2013).

Diese bekannten Probleme bei der gleichzeitigen ein- und mehrdimensionalen Auswertung eines Datensatzes beziehen sich letztendlich auf die Reliabilitäten der Ergebnisse. Ein bisher unberücksichtigtes Problem bei solchen mehrdimensionalen Analysen besteht jedoch in der Validität dieser Ergebnisse.

Fragestellung: Im Rahmen der Testentwicklung groß angelegter Studien werden die Aufgaben von den ersten Pre-Pilotierungen und Pilotierungen an, über den Feldtest, hin bis zur Hauptstudie so verändert bzw. ausgewählt, dass diese keine Mehrdimensionalität aufzeigen, sondern vielmehr Eindimensionalität bezüglich der übergeordneten Mathematikkompetenz. Das heißt, eine Aufgabe die gemäß der eindimensionalen Auswertung keine ausreichenden Fit-Werte zeigt, unter Umständen, weil sie etwa zu speziell den Inhaltsbereich "Daten und Zufall" misst, wird bereits lange vor der eigentlichen Hauptstudie aus dem Test ausgeschlossen. Ziel von mehrdimensionalen Auswertungen ist in der Regel die Entwicklung von Skalen, die Unterschiede aufweisen. Eine eindimensionale Testentwicklung hingegen hat das Gegenteil zur Folge, nämlich dass eventuelle Unterschiede zwischen den Dimensionen möglichst ausgeblendet werden. Bezüglich der durchgeführten mehrdimensionalen Auswertungen stellt sich hierbei weniger die Frage, inwieweit die eindimensionale Testentwicklung noch eine reliable Unterscheidung der mehrdimensionalen Skalen zulässt, sondern vielmehr inwieweit die Ergebnisse noch mit den tatsächlich vorhandenen speziellen Fähigkeiten in den definierten Inhaltsbereichen übereinstimmen; das heißt, inwieweit die Ergebnisse valide für die jeweiligen Inhaltsbereiche sind.

Methode: Um die beschriebene Problematik zu verdeutlichen und den möglichen Einfluss der ein- bzw. mehrdimensionalen Testentwicklung zu untersuchen, wird eine Simulationsstudie durchgeführt. Ausgehend von mehrdimensional generierten Datensätzen mit großer Aufgabenanzahl, wird dabei zunächst die Selektion einer Teilmenge von Aufgaben hin zu einem möglichst reliablen (fiktiven) eindimensionalen Konstrukt durchgeführt als auch die hin zu möglichst reliablen mehrdimensionalen Konstrukten. Im Anschluss werden dann beide Tests mehrdimensional ausgewertet und die Ergebnisse werden auf Unterschiede hin untersucht.

Ergebnisse und Diskussion: Die Simulationsstudie ist derzeit noch nicht vollständig abgeschlossen, die ersten Ergebnisse zeigen jedoch, dass die eindimensionale Auswertung der Daten deutlich andere Itemfit-Werte als die mehrdimensional Auswertung liefert und dadurch die Itemselektion entsprechend beeinflusst ist. Die abschließenden Analysen mit allen Ergebnisse werden auf der Konferenz präsentiert. Abschließend wird diskutiert inwieweit die berichteten Ergebnisse die Validität mehrdimensionaler Analysen auf Basis eindimensional entwickelter Tests in Frage stellen.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52, 345-370.

Brandt, S. & Ducker, B. (2013). Increasing unidimensional measurement precision using a multidimensional item response model approach. *Psychological Test and Assessment Modeling*, 55, 148-161.

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., ... & Stover, A. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement*, 31, 4-19.

OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris: OECD Publishing.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.

Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. Cambridge: University Press.



## **Validitätsforschung in Large Scale Assessments: Innovationen ohne Validitätsdilemma am Beispiel Komplexes Problemlösen**

Large Scale Assessments (LSA) berücksichtigen zunehmend fächerübergreifende Fähigkeiten (OECD, 2010; LLLight´In´Europe, 2010), um ein genaueres Bild zu zeichnen, wie gut vorbereitet Schüler, Studenten und Arbeitnehmer auf die Lernanforderungen und Arbeitsplätze des 21. Jahrhunderts sind. Als eine dieser Fähigkeiten gilt Komplexes Problemlösen (KPL), das dynamische, komplexe und nicht-routinemäßige Probleme, wie sie beim Lernen und am Arbeitsplatz auftauchen können, simuliert (Greiff & Fischer, 2013). Nach bedeutsamen psychometrischen Verbesserungen (Greiff et al., 2013) wurde KPL in die Testbatterien von PISA 2012 (OECD, 2010) und LLLight´In´Europe (2010), einem Assessment an 4000 Mitarbeitern in 60 europäischen Unternehmen, aufgenommen. Wir führen diesen Assessment-Erfolg von KPL zurück auf die Orientierung an internationalen Validitätsstandards von APA, AERA und NCME (1999). Diese Standards für schulische und psychologische Tests betrachten Validität als Angemessenheit, Aussagefähigkeit und Nützlichkeit bestimmter Interpretationen von Testwerten auf der Grundlage theoretischer Konstrukte (Cronbach, 1988; Messick 1989). Während die Validitätstheorie also weit entwickelt ist, hinkt die Anwendung zumindest in Europa häufig Jahrzehnte hinterher und betrachtet Validität noch als Testmerkmal, anstatt, wie von der APA (1999) empfohlen, als Ausmaß, mit dem intendierte Testwertinterpretationen zutreffend sind. Jedoch ist gerade im Kontext LSAs ein aktuelles Validitätsverständnis besonders wichtig, da hier häufig konkrete Implikationen für die Bildungspraxis und politische Entscheidungen abgeleitet werden, während gleichzeitig Innovationen Einzug halten. Werden Neuzugänge wie KPL nicht rechtzeitig angemessen validiert, ergibt sich aus dem Bedarf an Innovationen im LSA und der unklaren Interpretierbarkeit ihrer Messwerte ein Dilemma, das durch konsequente Validitätsforschung aufgefangen werden muss. Als Beispiel dafür beschreiben wir KPL und unsere ersten Versuche der Validierung.

Für den Versuch die Validitätsstandards auf KPL anzuwenden, um dessen Anwendbarkeit in LSA empirisch zu prüfen, nutzen wir drei Validitätskriterien. Mit Ergebnissen aus einer aktuellen Studie (Greiff et al., 2013) zu konvergenter, divergenter und prädiktiver Validität dreier Messinstrumente für KPL wollen wir das Ausmaß, mit dem die intendierten Testwertinterpretationen zutreffend sind, bestimmen. Wir fragen deshalb, ob (1) drei Instrumente für KPL dasselbe messen, (2) KPL unterschiedlich vom schlussfolgernden Denken ist und (3) KPL auf Konstruktebene Schulerfolg vorhersagt?

Greiff und Kollegen (2013) untersuchten eine Stichprobe (n=339) deutscher Universitätsstudenten, an denen drei KPL Assessments namens „MicroDYN“, „the Genetics Lab“ und „MicroFIN“ auf der Basis multipler komplexer Systeme (Funke, 2010) und der Matrizenubtest des „Intelligenz Struktur Test-Revised“ (IST-R; Beauducel, Liepmann, Horn, & Brocke, 2010) als Maß für schlussfolgerndes Denken erfasst wurden. Die Studenten haben darüber hinaus ihre Schulnoten als Maß für Schulerfolg berichtet. Für Frage (1) zur konvergenten Validität kam ein Multitrait–Multimethod-Ansatz nach Eid (2006) mit Trade- und Methodenfaktoren zum Einsatz, der die Generalisierbarkeit von Ergebnissen einzelner Messmethoden auf die Konstruktebene empirisch prüft. Für die Fragen (2) und (3) zur divergenten und prädikativen Validität, die Zusammenhänge zu schlussfolgerndem Denken und Schulerfolg betrachten, verwendeten Greiff und Kollegen (2013) Strukturgleichungsmodelle nach Kline (2011). Die Ergebnisse auf diese empirischen Fragen erlauben eine erste Abschätzung über die Einhaltung der Validitätsstandards in KPL gemäß APA, AERA und NCME (1999).

Die Ergebnisse geben Hinweise darauf, dass die drei Instrumente sich ausreichend ähnlich sind in dem, was sie messen, mit Konsistenzmaßen zwischen .50 und .60. Damit sind die KPL-Werte relativ unabhängig von ihren Instrumenten. Außerdem konnten wir belegen, dass KPL verschieden ist von schlussfolgerndem Denken, indem dies zwei KPL Dimensionen, Wissenserwerb und Wissensanwendung, nur mit ERWERB=.49 und ANWENDUNG=.53 vorhersagt. Bei der Vorhersage von Schulerfolg klärt KPL unabhängig vom schlussfolgernden Denken Varianz für naturwissenschaftliche Fächer auf (KPL=.22). Die Ergebnisse geben Hinweise auf konvergente, divergente und prädikative Validität für KPL und weisen auf Entwicklungspotenziale für die Validierung von innovativen Assessments vor allem in sehr großen Erhebungen hin, um eine gute Interpretierbarkeit der Ergebnisse zu gewährleisten und damit die Tragfähigkeit von Entscheidungen in der Praxis, die auf solche Ergebnisse zurückgreifen, zu stärken.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Beauducel, A., Liepmann, D., Horn, S., & Brocke, B. (2010). Intelligence Structure Test. Oxford: Hogrefe.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer & H. Braun (Hrsg.), Test validity. Hillsdale, NJ: Lawrence Erlbaum Associates.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7, 69–89.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, 11, 133–142.

Greiff, S., & Fischer, A. (2013). Der nutzen einer komplexen problemlöse- kompetenz: Theoretische überlegungen und empirische befunde. *Zeitschrift für Pädagogische Psychologie*, 27, 1–13.

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence*, 41, 579–596

Greiff, S., Holt, D. V., & Funke, J. (2013). Perspectives on problem solving in cognitive research and educational assessment: analytical, inter- active, and collaborative problem solving. *Journal of Problem Solving*, 5, 71–91.

Greiff, S., Wüstenberg, S., Molnar, G., Fischer, A., Funke, J., & Csapo, B. (2013). Complex Problem Solving in educational settings — something beyond g: concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, 105, 364–379.

Kline, R. B. (2011). Principles and practice of structural equation modeling (3. Ausg.). New York, NY: Guilford.

LLLight´in´Europe (2010). Concept and objectives, progress beyond state-of-the-art, S/T methodology and work plan.

Messick, S. (1989). Validity. In R. L. Linn (Hrsg.), Educational measurement. Washington, DC: American Council on Education and National Council on Measurement in Education.

OECD (2010). PISA 2012 field trial problem solving framework. Online in Internet: URL: <http://www.oecd.org/dataoecd/8/42/46962005.pdf>

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, 30, 463–480.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving. More than reasoning? *Intelligence*, 40, 1–14.



Felix Walker (Universität Stuttgart), Stephan Abele (Universität Stuttgart), Nico Link (Universität Stuttgart), Tobias Gschwendtner (Universität Stuttgart), Reinhold Nickolaus (Universität Stuttgart)

## **Validitätsansprüche an die Kompetenzmessung in der Berufsbildung und ihre Einlösung in gewerblich-technischen Anforderungskontexten**

### Theoretischer Hintergrund

Für die Akzeptanz von wissenschaftlichen Testverfahren in der Bildungspraxis ist eine überzeugende Validitätssicherung grundlegend. Dies erfordert im Bereich der beruflichen Bildung neben einer hohen Augenscheinvalidität, dass die tatsächlich in der beruflichen Praxis erbrachten Leistungen zum Kriterium werden, an dem sich die Testverfahren zu bewähren haben (Gschwendtner/Abele/Nickolaus, 2009). Besonders nachdrücklich wird dieser Anspruch im Kontext der Vorbereitung eines Large Scale Assessments eingebracht (Baethge & Arends, 2009). Die wissenschaftlichen Herausforderungen, die sich für das Einlösen der kriterienbezogenen Validität bzw. der Konstruktvalidität (Hartig & Frey, 2013; Hartig & Frey, 2012; Herzberg & Frey, 2011; Hartig, Frey, & Jude, 2008) in diesem Kontext ergeben, sind erheblich und werden in einigen Projekten des ASCOT-Programms im Rahmen von Validierungsstudien bearbeitet (Abele, 2013a; Walker, 2013).

Das Konstrukt berufsfachlicher Kompetenz lässt sich nach den vorliegenden Befunden domänenübergreifend mit dem Fachwissen und der Fähigkeit dieses Fachwissen in variierenden, auch problemhaltigen Anforderungskontexten erfolgreich einzusetzen, in zumindest zwei übergeordnete Subdimensionen ausdifferenziert, die je nach beruflichen Anforderungskontexten durch (manuelle) Fertigkeiten ergänzt werden (Nickolaus & Seeber, 2013). Während es im Bereich allgemeiner Bildung oder auch bezogen auf das berufsfachliche Wissen schwer fällt, scharfe Außenkriterien auszuweisen, da die Verwertbarkeit dieser Wissensformen vielfältig und schwer abgrenzbar ist, bietet sich im Bereich der Wissensanwendung in spezifischen beruflichen Kontexten die Möglichkeit solch scharfe Kriterien auszuweisen und auch zu prüfen. Das gilt z.B. für die Fehlerdiagnosekompetenz (analytische fachspezifische Problemlösefähigkeit) von Kfz-Mechatronikern, für die in einer Validitätsstudie gezeigt werden konnte, dass die Testleistungen in einer simulierten Arbeitsumgebung jenen in der realen Arbeitsumgebung entsprachen (Gschwendtner, Abele, & Nickolaus, 2009). Abele (2013b) zeigt darüber hinaus, dass die mit simulationsbasierten Tests gemessenen Testwerte konvergent valide im Hinblick auf strukturell einschlägigen Prüfungsanforderungen interpretiert werden können.

Fragestellung: Offen ist die Frage, inwieweit dies auch in anderen Domänen repliziert werden kann, welchen Anforderungen die Simulationsumgebungen zur Erfassung berufsfachlicher Kompetenzen genügen müssen, um diesem Kriterium standzuhalten und ob für andere Anforderungskontexte, wie z.B. die konstruktive Problemlösekompetenz, wie sie beispielsweise bei Elektronikern bei der Entwicklung von Steuerungsprogrammen benötigt wird, auch durch Papier- und Bleistifttests eingelöst werden kann. Kriterium wäre in jedem Falle die in realen Arbeitsumgebungen erbrachten Leistungen, die im Falle von Programmierarbeiten allerdings in wesentlichen Teilen auch am Arbeitsplatz in Simulationsumgebungen erbracht werden.

Methode: Im Rahmen des Beitrags wird zunächst das zu Grunde gelegte Validierungsdesign (und damit verbunden die Validierungslogik) vorgestellt. Anschließend werden Ergebnisse aus zwei Studien referiert, die in der gewerblich-technischen Berufsbildung angesiedelt sind. Geprüft werden im Kern zwei Hypothesen: (1) Testwerte auf Basis simulationstechnisch gestützter Verfahren erweisen sich zur Abschätzung der Fehleranalysefähigkeit (analytische berufsspezifische Problemlösekompetenz) gemessen an den in realen Systemen erbrachten Leistungen in unterschiedlichen Domänen als valide. (2) Bei Elektronikern für Automatisierungstechnik ist es möglich, mit Papier- und Bleistifttests auch die konstruktiven Problemlösekompetenzen valide abzuschätzen. Die Hypothesenprüfung erfolgt auf der Basis von Datensätzen der Ausbildungsberufe KFZ-Mechatroniker/-innen (N= 257) und Elektroniker/-innen (N=123) am Ende ihrer Ausbildung.

Ergebnisse: Der gegenwärtige Stand der Analysen legt nahe, dass mit simulationstechnisch gestützten Verfahren ermittelte Testwerte valide im Sinne des Kriteriums der analytischer Problemlösekompetenz im Ausbildungsberuf des KFZ-Mechatronikers/Mechatronikerin interpretiert werden können. Gleiches gilt im Bereich des Elektronikers/Elektronikerin für Automatisierungstechnik für Papier- und Bleistift-basierte Verfahren zur Erfassung der konstruktiven Problemlösekompetenz.

Diskussion: Die Ergebnisse der Validierungsstudien zeigen einerseits, dass mit den entwickelten Verfahren valide und (test)ökonomisch vorteilhafte Abschätzungen zentrale Facetten beruflicher Fachkompetenz möglich sind und legen in Verbindung mit den Befunden Abeles den Gedanken nahe, diese auch in Prüfungskontexten einzusetzen. Zu vermuten ist, dass dies in hohem Grade an die strukturellen Affinitäten der test- und realitätsbezogenen Anforderungskontexte gebunden ist.

Abele, S. (2013a). Diagnostik beruflicher Kompetenzen bei beschränkter Testzeit: Leistungsfähigkeit eines Segmentierungs-Ansatzes. Vortrag auf der 4. ASCOT-Tagung am FBB. Nürnberg, 19.09.2013.

Abele, S. (2013b). Modellierung, Entwicklung und Determinanten berufsfachlicher Kompetenz in gewerblich-technischen Ausbildungsberufen. Analysen auf Basis von testbasierten, berufsschulischen und betrieblichen Leistungsdaten sowie Prüfungsergebnissen. Aachen: Shaker, Universität Stuttgart, Dissertation.

Baethge, M. & Arends, L. (2009). Feasibility Study Vet-LSA. Bonn, Berlin

Frey, A. & Hartig, J. (2013). Wann sollten computerbasierte Verfahren zur Messung von Kompetenzen anstelle von papier- und bleistift-basierten Verfahren eingesetzt werden? Zeitschrift für Erziehungswissenschaft, 16, 53-57.

Hartig, J. & Frey, A. (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. Psychologische Rundschau, 63, 43-49.

Hartig, J., Frey, A. & Jude, N. (2008). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), Testtheorie und Fragebogenkonstruktion. Berlin: Springer.

Herzberg, P. Y. & Frey, A. (2011). Kriteriumsorientierte Diagnostik. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), Methoden der psychologischen Diagnostik (Psychologische Diagnostik, Bd. 2, [Vollst. Neuausg.], S. 281–325). Göttingen [u.a.]: Hogrefe.

Walker, F. (2013). Minimal komplexe Aufgaben als Lösungsansatz für die Reliabilitätsproblematik beim analytischen Problemlösen in automatisierungstechnischen Systemen. Vortrag auf der 4. ASCOT-Tagung am FBB. Nürnberg, 19.09.2013.





Chair(s): Jürgen Seifried (Universität Mannheim)

Diskutant/in: Klaus-Peter Wild (Universität Regensburg)

Seit einigen Jahrzehnten arbeitet die (berufliche) Lehr-Lern-Forschung mit ökologisch validen Verfahren, mit deren Hilfe man das subjektive Erleben von Individuen in alltäglichen Situationen abbilden kann. Csikszentmihalyi und Larson (1987) etablierten in diesem Zusammenhang die „Experience Sampling Method“ (ESM), die auf die Erfassung von Aktivitäten sowie subjektiven Erlebensqualitäten abzielt. Zudem kann ermittelt werden, inwiefern äußere Umstände mit dem subjektiven Erleben in Verbindung stehen. Die Arbeiten von Csikszentmihalyi und Larson besitzen auch für die Lehr-Lern-Forschung unmittelbar Relevanz (siehe z.B. Götz et al., 2010), und auch in der beruflichen Bildung sind seit dem DFG-Schwerpunktprogramm „Lehr-Lern-Prozesse in der kaufmännischen Erstausbildung“ (Beck & Dubs, 1998; Beck & Krumm, 2001) ESM-Studien ebenfalls gängige Forschungspraxis (siehe z.B. Wild & Krapp, 1996, oder Sembill, Seifried & Dreyer, 2008).

Das vorliegende Symposium bietet einen Überblick über die Anwendung der ESM in der Lehr-Lern-Forschung. Es geht darum, die Potenziale, aber auch die Limitationen der ESM aus verschiedenen Perspektiven zu beleuchten. Einen Schwerpunkt bildet dabei die Unterrichtsforschung. Zudem wird – über dieses gängige Einsatzfeld hinaus – untersucht, inwiefern sich die ESM für die Vorhersage der Berufswahlintention sowie die Kompetenzmessung in beruflichen Settings eignet. Das Symposium vereint Forschungsarbeiten aus der Berufs- und Wirtschaftspädagogik sowie der Pädagogischen Psychologie und umfasst vier Beiträge:

- (1) In der Studie von Keller et al. steht die Analyse des Zusammenhangs zwischen der Unterrichtsgestaltung und dem Erleben von Emotionen im Unterricht. Mehrebenenanalytisch (N=212 Schülerinnen und Schüler aus 41 Gymnasialklassen) konnten signifikante Zusammenhänge zwischen Unterrichtsgestaltungsformen sowie Emotionen gefunden werden, und zwar über verschiedene Fächer hinweg, was auf eine Generalisierbarkeit der Befunde hindeutet.
- (2) Bieg et al. analysieren unter Rückgriff auf die Erwartungs x Wert-Theorie, inwiefern Emotionen (differenziert nach State und Trait-Komponenten) bei der Berufswahl eine Rolle spielen (N=622 SchülerInnen aus den Klassenstufen 9 und 10 an Gymnasien im Fach Mathematik). Die Befunde weisen darauf hin, dass insbesondere Trait-Emotionen i. S. von Überzeugungen (vermutete Emotionen) und weniger tatsächlich erlebte Emotionen für die Entscheidung für oder gegen einen mit Inhalten der Mathematik verbundenen Beruf von Relevanz sind.
- (3) Kristina Kögler nutzt die ESM zur Erfassung der Schülerlangeweile im kaufmännischen Unterricht. Mittels aufwendiger Prozesserhebungen (N=96 aus vier Klassen an kaufmännischen Schulen, acht Unterrichtseinheiten) werden unterrichtsbezogene Ursachen für Schülerlangeweile herausgearbeitet.
- (4) Rausch u.a. berichten über die Möglichkeiten und Grenzen der ESM bei der Erfassung nicht-kognitiver Kompetenzfacetten in der kaufmännischen Berufsausbildung. Im Vortrag werden die Befunde aus einer Pilotstudie (N=120 Auszubildende) berichtet.

Beck, K. & Dubs, R. (Hrsg.) (1998). Kompetenzerwerb in der Berufserziehung. Kognitive, motivationale und moralische Dimensionen kaufmännischer Qualifizierungsprozesse. Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft Nr. 14. Stuttgart: Franz Steiner.

Beck, K. & Krumm, V. (Hrsg.) (2001). Lehren und Lernen in der beruflichen Erstausbildung. Grundlagen einer modernen kaufmännischen Berufsqualifizierung. Opladen: Leske + Budrich.

Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-sampling method. The Journal of nervous and mental disease, 175, 526-536.

Goetz, T., Frenzel, A. C., Stoeger, H., & Hall, N. C. (2010). Antecedents of everyday positive emotions: An experience sampling analysis. Motivation and Emotion, 34, 49-62.

Sembill, D., Seifried, J. & Dreyer, K. (2008). PDAs als Erhebungsinstrument in der beruflichen Lernforschung – Ein neues Wundermittel oder bewährter Standard? Empirische Pädagogik, 22, 64-77.

Wild, K.-P. & Krapp, A. (1996). Die Qualität subjektiven Erlebens in schulischen und betrieblichen Lernumwelten: Untersuchungen mit der Erlebens-Stichproben-Methode. Unterrichtswissenschaft, 24, 195-216.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
PEG-Gebäude  
Raum 1.G191



## Unterrichtsgestaltung und Emotionen von Schülerinnen und Schülern: Eine Experience-Sampling-Studie

Theoretischer Hintergrund: Emotionales Erleben wird zunehmend als wichtiges Ergebnis von Lernprozessen bei Schüler/innen erkannt. Dabei wird angenommen, dass die Unterrichtsgestaltung die Emotionen von Schüler/innen beeinflusst (siehe Pekrun, 2006). Als diesbezüglich zentrale Facetten der Unterrichtsgestaltung haben sich bisher unter anderem Klarheit/Strukturiertheit (Frenzel, Pekrun, & Goetz, 2007), Lob/Feedback (Goetz, Pekrun, Hall, & Haag, 2006), und Enthusiasmus der Lehrperson (Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009) gezeigt.

Emotionen von Schüler/innen sind je nach Fach unterschiedlich stark ausgeprägt (z.B. Goetz, Frenzel, Lüdtke, & Hall, 2011). Bisherige Untersuchungen fokussierten dabei jedoch entweder ausschließlich auf die mittlere Ausprägung von Emotionen in verschiedenen Fächern, oder lediglich auf ein Fach. Das Ausmaß an Fachspezifität bezüglich des Zusammenhängens zwischen Unterrichtsgestaltung und Emotionen bleibt unklar.

In bisherigen Studien berücksichtigte Trait-Emotionen können durch subjektive Überzeugungen verzerrt sein (Robinson & Clore, 2002). Zudem wurden bisherige Untersuchungen meist auf interindividueller Ebene durchgeführt, sodass eine Übertragbarkeit der Befunde auf eine intraindividuelle Ebene für State-Emotionen unklar ist. Fragestellung: Ziel der vorliegenden Studie ist es mittels der Experience-Sampling Methode zu untersuchen, inwieweit Unterrichtsgestaltung und Emotionen (Freude, Stolz, Angst, Ärger, Hilflosigkeit und Langeweile) von Schüler/innen auf intraindividuelle Ebene zusammenhängen und inwieweit dieser Zusammenhang sich zwischen Fächern (Mathematik, Physik, Deutsch und Englisch) unterscheidet.

Methode: Insgesamt N=212 Schüler/innen (50% männlich) aus der 8. und 11. Jahrgangsstufe (mittleres Alter: M8.Jgst=14.5 bzw. M11.Jgst=17.7) wurden zufällig aus 41 Gymnasialklassen ausgewählt (pro Klasse ca. 2-4) um an der Experience-Sampling Studie teilzunehmen.

Über einen Zeitraum von ca. 10 aufeinanderfolgenden Schultagen haben die Schüler/innen zu Beginn einer Mathematik-/Physik/Deutsch-/Englischstunde einen PDA aktiviert, der dann randomisiert während der folgenden Unterrichtsstunde einen Signalton gab und einen Kurzfragebogen präsentierte.

Aus Zeit- und Validitätsgründen wurden alle Konstrukte mittels Einzelitems erfasst. Dabei wurden zunächst die 8 Items zur Unterrichtsgestaltung (Verständlichkeit, Anschaulichkeit, Enthusiasmus, Klassenführung, wenig Klarheit, zu hohe Schwierigkeit, zu hohe Geschwindigkeit und zu hohes Anforderungsniveau der Lehrperson) auf einer fünfstufigen Skala eingeschätzt. Anschließend bewerteten die Schüler/innen ihr emotionales Erleben ("Wie stark erlebst du in dieser Unterrichtsstunde Freude/Stolz/Angst/Ärger/Hilflosigkeit/Langeweile?") von (1) gar nicht bis (5) sehr stark.

Um der genesteten Datenstruktur Rechnung zu tragen, wurden alle Analysen mittels Mehrebenenmodellen durchgeführt. Zunächst wurden die acht Facetten der Unterrichtsgestaltung faktoranalysiert (Reise, Ventura, Nuechterlein, & Kim, 2005) und anschließend in Mehrebenen-Regressionsmodellen die Emotionen auf intraindividuelle Ebene durch die Faktoren der Unterrichtsgestaltung vorhergesagt.

Ergebnisse: Bei der Faktorenanalyse der Unterrichtsfacetten hat sich das Zwei-Faktoren Modell als die beste Lösung erwiesen ( $\chi^2=27.454$ ,  $df=13$ ,  $CFI=.996$ ,  $RMSEA=.000$ ). Die beiden Faktoren wurden als Unterstützende Darbietung (hohe Verständlichkeit und Anschaulichkeit, hohe Klassenführung und hoher Enthusiasmus) und Überhöhte Anforderungen (geringe Klarheit, zu schwierig und zu schnell, zu hohe Anforderungen) interpretiert.

Die Mehrebenen-Regressionsanalysen zeigten zunächst beinahe durchgängig signifikante Zusammenhänge zwischen den Unterrichtsgestaltungsfaktoren und den Emotionen. So zeigte sich, dass je unterstützender die jeweilige Unterrichtsstunde wahrgenommen wird (hohe Ausprägung auf Faktor 1), desto mehr Freude und Stolz, und desto weniger Langeweile berichten die Schüler/innen. Umgekehrt, je überhöhter die Anforderungen in der Unterrichtsstunde (hohe Ausprägung auf Faktor 2), desto weniger Freude und Stolz, und desto mehr Angst, Ärger, Hilflosigkeit und Langeweile empfinden die Schüler/innen. Die Effekte sind dabei als stark zu bezeichnen ( $.25 \leq R^2_{\text{Level1}} \leq .45$ ). Auf Basis dieser Ergebnisse lässt sich also schlussfolgern, dass eine klare Unterrichtsführung, in der die Anforderungen dem Leistungsniveau der Schüler/innen angemessen sind, positive Emotionen fördert, und negative Emotionen vermindert.

Es zeigten sich nur wenige signifikante Unterschiede in der Stärke des Zusammenhangs von Unterrichtsgestaltung und Emotionen bezüglich unterschiedlicher Fächer. Es lässt sich also ableiten, dass die Einflüsse von Unterrichtsgestaltung auf Emotionen über Fächer hinweg generalisierbar sind. Dieser Befund hat auch Auswirkungen auf die Lehrerbildung.

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101(3), 705-716. doi: 10.1037/a0014695 .

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, 5, 478-493.

Goetz, T., Frenzel, A. C., Lüdtke, O., & Hall, N. (2011). Between-domain relations of academic emotions: Does having the same instructor make a difference. *The Journal of Experimental Education*, 79, 84-101.

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology*, 76(2), 289-308. doi: 10.1348/000709905x42860 .

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341.

Reise, S. P., Ventura, J., Nuechterlein, K. H., & Kim, K. H. (2005). An illustration of multilevel factor analysis. *Journal of Personality Assessment*, 84(2), 126-136.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934-960.



## ***Auf die Freude kommt es an! Die Rolle von Trait- und State-Emotionen bei der Vorhersage von Berufswahlintention in Mathematik***

Theoretischer Hintergrund: Eine der wichtigsten Entscheidungen im Leben eines Menschen ist die Wahl des Berufes. Deshalb ist es nicht verwunderlich, dass sich unter anderem die pädagogisch-psychologische Forschung mit der Berufswahl von Menschen und welche Faktoren diese beeinflussen befasst. Basierend auf der Erwartungs × Wert-Theorie leistungsbezogener Entscheidungen (z.B. Eccles, 1985) spielen sowohl Erwartungs- als auch Valenzaspekte bei der Vorhersage von leistungsbezogenen Entscheidungen, wie der Berufswahl, eine bedeutsame Rolle. Zur Prüfung der Modellvorhersagen wurden Emotionen, die der Valenzkomponente zuzurechnen sind, in der bisherigen Forschung kaum berücksichtigt. Neben der Untersuchung von Emotionen überhaupt, steht die differenzierte Betrachtung der prädiktiven Validität von State-Emotionen im Vergleich zu Trait-Emotionen (habituelle Emotionen) noch aus (z.B. Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). Erste Forschungsergebnisse lassen vermuten, dass vor allem Trait-Emotionen im Sinne von Überzeugungen hinsichtlich eigener Emotionen und weniger tatsächlich erlebte Emotionen für Wahlentscheidungen von Bedeutung sind (Wirtz, Kruger, Napa Scollon, & Diener, 2003).

Fragestellung / Hypothesen

Ausgehend von der Erwartungs × Wert-Theorie waren die Beantwortung folgender Forschungsfragen Ziel der vorliegenden Studie: Inwieweit sagen Emotionen (Wertkomponente) und das akademische Selbstkonzept (Erwartungskomponente) die Berufswahlintention vorher? Haben Trait-Emotionen oder State-Emotionen stärkere Vorhersagekraft?

Hypothese 1: Die Emotionen Freude und Angst sagen die Berufswahlintention vorher. Dabei ist Freude ein positiver und Angst ein negativer Prädiktor.

Hypothese 2: Die Emotionen Freude und Angst haben zusätzlich zum akademischen Selbstkonzept prädiktive Validität.

Hypothese 3: Die Emotionen und das akademische Selbstkonzept interagieren in der Vorhersage der Berufswahlintention.

Hypothese 4: Trait-Emotionen haben stärkere prädiktive Validität als State-Emotionen.

Methode: Zur Untersuchung der Hypothesen wurden Daten im Fach Mathematik von 622 Schüler/innen und Schülern der Klassenstufen 9 und 10 aus 31 Schulklassen an Gymnasien in Baden-Württemberg erhoben. Hierzu beantworteten die Studienteilnehmenden sowohl zu Studienbeginn einen Trait-Fragebogen zur Erfassung der relevanten Variablen (Trait-Emotionen, Selbstkonzept, Berufswahlintention) als auch im weiteren Studienverlauf je einen kurzen Fragebogen zu ihren State-Emotionen nach jeder Mathematikstunde über einen Zeitraum von drei Wochen (Tagebuchphase). Als Indikator für die Erwartungskomponente wurde das akademische Selbstkonzept in Mathematik mit drei Items erfasst (Marsh, 1990;  $\alpha = .88$ ). Die Wertkomponente wurde durch die Emotionen Freude und Angst operationalisiert (Pekrun et al., 2007;  $\alpha = .84$  für Freude und  $\alpha = .71$  für Angst), wobei parallel formulierte Items zur Erfassung von Trait- und State-Emotionen verwendet wurden. Die Berufswahlintention im Sinne der Intention einen mathematikintensiven Beruf zu ergreifen, wurde mit vier Items erfasst (Schuster & Martiny, 2012 und Eigenentwicklung;  $\alpha = .94$ ). Erste Analysen mittels basaler multipler Regression wurden bereits durchgeführt. Zur weiteren Analyse der Daten werden latente Strukturgleichungsmodelle unter Berücksichtigung der Mehrebenenstruktur der Daten (Schüler/innen in Klassen und in Schulen) berechnet.

Ergebnisse und Diskussion: Sowohl die Erwartungskomponente (akademisches Selbstkonzept) als auch die Wertkomponente (Emotionen Freude und Angst) waren sowohl getrennt als auch gemeinsam signifikante Prädiktoren in der Vorhersage der Berufswahlintention. Es zeigt sich, dass das akademische Selbstkonzept 34% der Varianz aufklärt, während Trait-Freude und Trait-Angst gemeinsam 20 % der Varianz in der Berufswahlintention erklären. Dabei ist Freude ein stärkerer Prädiktor als Angst ( $b = .34$  vs.  $b = -.17$ ). Die Interaktion von Selbstkonzept und Trait-Freude wurde zusätzlich zu den Haupteffekten signifikant. Als weiteres Ergebnis lässt sich feststellen, dass Trait-Emotionen höhere prädiktive Validität haben als State-Emotionen. Es sind also weniger die tatsächlich erlebten Emotionen als die vermuteten Emotionen bei der Vorhersage der Berufswahlintention bedeutsam. Die Ergebnisse zeigen, dass neben dem akademischen Selbstkonzept als bedeutendstem Prädiktor nicht allein die Abwesenheit von Angst für die Berufswahlintention entscheidend ist, sondern vielmehr auch das (vermutete) Erleben positiver Emotionen von großer Wichtigkeit ist. Dies beinhaltet deutliche Implikationen im Hinblick auf die Förderung von Selbstkonzept und positiven Emotionsüberzeugungen bei Schülerinnen und Schülern. Die alleinige Reduktion von Angst scheint nicht ausreichend um die Berufswahlintention von Lernenden zu beeinflussen.

Eccles, J. S. (1985). Why doesn't Jane run? Sex differences in educational and occupational patterns. In F. D. Horowitz & M. O'Brien (Eds.), *The gifted and talented: Developmental perspectives*. (pp. 251-295). Washington, DC US: American Psychological Association.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science*, 24, 2079-2087. doi: 10.1177/0956797613486989

Marsh, H. W. (1990). Self-Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph. Macarthur, Australia: University of Western Sydney.

Pekrun, R., vom Hofe, R., Blum, W., Frenzel, A. C., Götz, T., & Wartha, S. (2007). Development of mathematical competencies in adolescence: The PALMA longitudinal study. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report of the DFG Priority Programme*. (pp. 17-37). Münster, Germany: Waxmann.

Schuster, C., Martiny, S.E. (2012, September). The influence of stereotype threat on thoughts, emotions, and motivation. A scenario study. Vortrag auf dem 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPS) in Bielefeld (Germany).

Wirtz, D., Kruger, J., Napa Scollon, C., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science*, 14, 520-524. doi: 10.1111/1467-9280.03455



## **Continuous-State-Sampling zur Erfassung von Schülerlangeweile in kaufmännischen Unterrichtsprozessen**

Empirische Befunde älteren Datums zeigen, dass sich Lernende in mindestens einem Drittel der Unterrichtszeit langweilen (z.B. Larson & Richards 1991). Als bislang erforschte Ursachen für die Langeweile von Lernenden gelten verschiedene personenbezogene Faktoren wie etwa ungünstige emotional-motivationale Persönlichkeitskonstellationen oder Vorwissensbestände, die in eine Über- oder Unterforderung münden können, aber besonders auch situative Aspekte wie beispielsweise inhaltliche oder methodische Monotonie und Engführung des Unterrichts, mangelnder Sinnbezug der Unterrichtsinhalte, mangelnde Aktivierung der Lernenden sowie die Unausweichlichkeit der Unterrichtssituation (in der Gesamtschau s. z. B. Daschmann, Goetz, & Stupnisky 2011). Die Befundlage resultiert dabei schwerpunktmäßig aus retrospektiven Fragebogenerhebungen im Sinne von Selbstauskünften der Lernenden zu Ausmaß, Erscheinungsformen und Bedingungsfaktoren in allgemeinbildenden Unterrichtskontexten. Prozessorientierte Untersuchungen, die an das Erleben der Lernenden in der konkreten Unterrichtssituation anknüpfen, sind vergleichsweise unterrepräsentiert, und auch der berufsbildende Bereich bleibt weitgehend ausgeblendet.

Doch auch in der beruflichen Bildung ist angesichts heterogener Klassenverbände und bisweilen methodischer Monokulturen von einer Existenz des Phänomens Langeweile auszugehen. Im kaufmännischen Bereich ist dabei besonders das Fach Rechnungswesen von Interesse, das in didaktischer und methodischer Hinsicht seit längerem als reformbedürftig gilt (z.B. Preiss 2001; Seifried 2004) und damit im Hinblick auf die Untersuchung emotional-motivationaler Erlebensqualitäten von Schülerinnen und Schülern als lohnenswertes Explorationsfeld erscheint. Pädagogische und didaktische Relevanz erhält das Problem der Langeweile zudem insofern, als es mit negativen Konsequenzen wie Motivations- oder Leistungsabfall, Schulmüdigkeit, Absentismus oder Schulabbruch in Verbindung gebracht wird und Lehrkräfte die Ursachen vornehmlich auf Seiten der Lernenden vermuten (Wuttke, 2010). Der vorliegende Beitrag nimmt diese Ausgangslage zum Anlass, die Frage der Schülerlangeweile und ihr Ursachenspektrum im Rahmen eines prozessanalytischen Zugangs im Unterricht des Faches Rechnungswesen in den Blick zu nehmen.

Einen geeigneten methodischen Zugang zur situationsnahen und engmaschigen Abbildung des Langeweileerlebens im Unterrichtsverlauf bietet das bewährte Continuous-State-Sampling-Verfahren (im Überblick s. Sembill, Seifried & Dreyer 2008). Mittels dieser Methodik werden über äquidistante Messzeitpunkte im Unterrichtsprozess kontinuierlich verschiedene Erlebensqualitäten der Lernenden wie etwa deren Langeweile erhoben, die sich im Nachgang empirisch mit den über strukturierte Beobachtungsverfahren gewonnenen Erkenntnissen zur Unterrichtsgestaltung verknüpfen lassen (vgl. etwa die Arbeiten zum Selbstorganisierten Lernen, im Überblick Sembill et al., 2007 und die Beiträge der Forschergruppe um Wild, 2001). Auf diesem Wege wird ein tiefgehender Einblick in die situativen Ursachenkonstellationen von Schülerlangeweile möglich. Mehrebenenanalytische Auswertungsverfahren ermöglichen darüber hinaus auch die Berücksichtigung personenbezogener Bedingungsfaktoren der Langeweile auf der Individualebene. Mit diesem Vorgehen wird aufgrund der simultanen Berücksichtigung unterschiedlicher Prädiktorengruppen auf mehreren Ebenen die potentielle Deutungsverzerrung separater Analysen vermieden (Hochweber & Hartig 2012).

Der vorliegende Beitrag berichtet aus einer Prozessuntersuchung im Fach Rechnungswesen, in der neben den im Rahmen der Videographie beobachteten methodischen Parameter der Unterrichtsgestaltung die Schülerlangeweile mittels der Continuous-State-Sampling-Methodik erhoben wurde. Es nahmen 96 Lernende in vier Klassen an der Erhebung in einer kaufmännischen Schule in Nordbayern teil, in jeder Klasse gingen 8 Unterrichtseinheiten mit jeweils siebenminütig getakteten Messzeitpunkten (n=193) in die Datenauswertung ein.

Im Zuge der Analyse werden situative Prädiktoren des Langeweileerlebens wie etwa die methodische und inhaltliche Unterrichtsgestaltung, Monotonie und das Ausmaß der Schüleraktivierung, aber auch personenbezogene Faktoren wie Vorwissen, Selbstwirksamkeit und habitualisierte Lernemotionen auf ihre Bedeutsamkeit für die Langeweileentstehung getestet. Die Ergebnisse deuten unter Kontrolle der individuellen Unterschiede auf eine übergeordnete Bedeutung von unterrichtlichen Leerläufen und monotonen Unterrichtsphasen hin, der Einfluss methodischer Parameter divergiert indes in Abhängigkeit der inhaltlichen Ausgestaltung des Unterrichts.

Daschmann, E. C., Goetz, T. & Stupnisky, R. H. (2011). Testing the predictors of boredom at school: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology*, 81: no. doi: 10.1348/000709910X526038 .

Hochweber, J. & Hartig, J. (2012). Mehrebenenanalyse. In S. Maschke & L. Stecher (Hrsg.), *Enzyklopädie Erziehungswissenschaft Online*. doi: 10.3262/EE007120217 .

Larson, R. W. & Richards, M. H. (1991). Boredom in the Middle School Years: Blaming Schools versus Blaming Students. *American Journal of Education*, 99(4), 418-443.

Preiss, P. (2001). Curriculare und methodische Neuorientierung des Rechnungswesensunterrichts: Vom Buchhalter zum kaufmännischen Sachbearbeiter – von der Bilanz zu Wertströmen im Unternehmensmodell. In H. Reinisch, R. Bader & G. A. Straka (Hrsg.), *Modernisierung der Berufsausbildung in Europa* (S. 135-144). Neue Befunde der berufs- und wirtschaftspädagogischen Forschung. Opladen: Leske + Budrich.

Seifried, J. (2004). Fachdidaktische Variationen in einer selbstorganisationsoffenen Lernumgebung – Eine empirische Untersuchung des Rechnungswesensunterrichts. Wiesbaden: DUV.

Sembill, D., Seifried, J. & Dreyer, K. (2008). PDAs als Erhebungsinstrument in der beruflichen Lernforschung – Ein neues Wundermittel oder bewährter Standard? *Empirische Pädagogik* 22 (1), 64-77.

Sembill, D., Wuttke, E., Seifried, J., Egloffstein, M. & Rausch, A. (2007). Selbstorganisiertes Lernen in der beruflichen Bildung – Abgrenzungen, Befunde und Konsequenzen. *bwpat*, 13. Verfügbar unter [http://www.bwpat.de/ausgabe13/sem\\_bill\\_etal\\_bwpat13.shtml](http://www.bwpat.de/ausgabe13/sem_bill_etal_bwpat13.shtml)

Wild, K.-P. (2001). Die Optimierung von Videoanalysen durch zeitsynchrone Befragungsdaten aus dem Experience Sampling. In S. von Aufschnaiter & M. Welzel (Hrsg.), *Nutzung von Videodaten zur Untersuchung von Lehr-Lernprozessen* (S.61-74). Münster: Waxmann

Wuttke, E. (2010). Was wissen Lehrerinnen und Lehrer über Langeweile? Eine explorative Studie. *Die berufsbildende Schule* 62, 312-318.



Andreas Rausch (Otto-Friedrich-Universität Bamberg), Jürgen Seifried (Universität Mannheim), Detlef Sembill (Otto-Friedrich-Universität Bamberg), Julia Warwas (Otto-Friedrich-Universität Bamberg), Steffen Brandt (Wissenschaftszentrum Kiel)

## ***Embedded Experience Sampling (EES) – In Testsituationen eingebettete Erlebensstichproben zur Erfassung nicht-kognitiver Kompetenzfacetten***

Ansätze der Kompetenzmodellierungen umfassen üblicherweise neben kognitiven Facetten wie Wissen, Fertigkeiten, Fähigkeiten auch nicht-kognitive Facetten wie Interessen, Einstellungen, Motive, Selbstkonzepte (bspw. Weinert, 2001; Erpenbeck & von Rosenstiel, 2003), wenngleich diesen in der empirischen Modellierung und Operationalisierung zumeist eine untergeordnete Rolle zukommt. Einer elaborierten Testentwicklung im Bereich des deklarativen und prozeduralen Wissens werden häufig lediglich globale Selbstberichtsskalen zu nicht-kognitiven Facetten wie Interessen und Selbstwirksamkeitserwartungen zur Seite gestellt, wobei Letztere dann zumeist Randnotizen bleiben. Angemessener erschien uns dagegen eine integrierte Erhebung von kognitiven und nicht-kognitiven Facetten der Kompetenz (Sembill, Rausch & Kögler, 2013). Hierzu haben wir unter der Bezeichnung „Embedded Experience Sampling“ (EES) eine Methode entwickelt, die dazu dient, nicht-kognitive Erlebensprozesse in komplexen Testsituationen zu erfassen (Rausch & Seifried, in prep.).

Dem Projekt „Domänenspezifische Problemlösekompetenz bei Industriekauf-leuten“ (DomPL-IK) im Rahmen der BMBF-Initiative „Technologie-orientierte Kompetenzmessung in der Berufsbildung“ (ASCOT) liegt ein ganzheitliches Kompetenzkonstrukt zugrunde, das sowohl kognitive und metakognitive als auch nicht-kognitive Facetten umfasst. Die Modellierung ist dabei konsequent am Problemlöseprozess orientiert. Domänenspezifische Problemlösekompetenz umfasst demnach im (a) kognitiven Segment die Facetten (1) Handlungsbedarfe und Informationsquellen identifizieren, (2) Informationsverarbeitung mittels domänenspezifischer Algorithmen und Werkzeuge, (3) Entscheidungen treffen und begründen sowie (4) Entscheidungen angemessen kommunizieren. Für das (b) metakognitive Segment sind (5) Planvolles (strukturiertes) Vorgehen, (6) Persistenz (fokussiertes Vorgehen) und (7) retrospektive Handlungskontrolle zu nennen. Im nicht-kognitiven Bereich wird zum einen das (c) domänenspezifische Selbstkonzept als Bündel aus (8) angemessenem Vertrauen in die eigene Kompetenz, (9) Unsicherheits-/Ungewissheitstoleranz und (10) angemessenem Vertrauen in die eigene Lösung berücksichtigt. Zum anderen wird das (d) domänenspezifische Interesse als Bündel aus (11) inhaltlichem Interesse am Problem, (12) positiver Aktiviertheit bei der Problembearbeitung und (13) Interesse am Fortgang des Problems und an potenziellen Lernchancen modelliert. Zur Erhebung der 13 Facetten wurden komplexe Problemszenarien aus der kaufmännischen Domäne entwickelt, deren Bearbeitung ca. 30 Minuten in Anspruch nimmt.

Zur Messung der im nicht-kognitiven Bereich genannten Facetten (Ziffern 8 bis 13) wurde das so genannte „Embedded Experience Sampling“ (EES) entwickelt. EES baut auf der von der Forschergruppe um Csikszentmihalyi entwickelten Methode des „Experience Sampling“ (ESM) auf und zielt wie diese auf die Sammlung von Daten nahe am Handlungsprozess bzw. „in situ“ ab (Hektner, Schmidt, & Csikszentmihalyi, 2006), wobei es sich hier um Problemlöseprozesse in Testsituationen handelt. Im Rahmen von EES werden die Probandinnen und Probanden komplexer Problemlösetests an gewissen Stellen im Prozess unterbrochen und zur Beantwortung kurzer Fragen aufgefordert, die ihr aktuelles Erleben in Bezug auf die Problemsituation betreffen. Die EES-Eingabeaufforderungen können wie auch ESM-Eingaben ereignisbasiert oder zeitgesteuert ausgelöst werden und prinzipiell sowohl offene als auch geschlossene Items enthalten. Der wichtige Unterschied zum gängigen Experience Sampling besteht darin, dass die EES-Eingaben möglichst authentisch in den Kontext des Problemlöseszenarios eingelassen sind. In unserem Fall ähneln sie alltäglichen Interaktionen am Arbeitsplatz, in denen bspw. die Vorgesetzte fragt, ob man sich die Aufgabe zutraut, oder der Proband kurz mit einem fiktiven Kollegen, der sich nach dem allgemeinen Befinden erkundigt, kommuniziert. Hierdurch soll die „Künstlichkeit“ sonst üblicher ESM-Erhebungen ein Stück weit überwunden werden. Diese spielerische Einbindung und der Handlungsdruck innerhalb der Testsituation sollten zudem zu recht schnellen, spontanen Angaben, bei denen nicht über sozial erwünschte Antworten nachgedacht wird, führen. Dies äußerten zumindest die Probandinnen und Probanden einer Pilotstudie in Interviews, die nach dem Testdurchlauf stattfanden. Die EES-Erhebungen wurden trotz einer in dieser Pilotstudie sehr umständlichen Paper & Pencil-Version als authentisch, kreativ und „belebend“ wahrgenommen.

Mittels EES möchten wir Konstrukte, die üblicherweise in globalen und abstrakten Selbstberichtsskalen erhoben werden, in konkreten Handlungsprozessen erheben und erhoffen uns dadurch eine höhere Validität. Die Methode, exemplarische Umsetzungen und Befunde aus einer Pilotstudie mit ca. 120 Auszubildenden werden im Rahmen des Beitrags vorgestellt und diskutiert.

Erpenbeck, J. & von Rosenstiel, L. (2003): Einführung. In: Erpenbeck, J. & von Rosenstiel, L. (Hrsg.): Handbuch Kompetenzmessung – erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Stuttgart: Schäffer-Poeschel, IX-XL.

Hektner, J. M., Schmidt, J. A. & Csikszentmihalyi, M. (2007): Experience sampling method – measuring the quality of everyday life. Thousand Oaks, London, New Delhi: Sage Publications.

Rausch, A. & Seifried, J. (in prep.). Embedded Experience Sampling (EES)—An integrated approach of measuring non-cognitive facets of competence.

Sembill, D., Rausch, A. & Kögler, K. (2013). Non-cognitive facets of competence—Theoretical foundations and implications for measurement. In K. Beck & O. Zlatkin-Troitschanskaia (Eds.), From Diagnostics to Learning Success: Proceedings in Vocational Education and Training (pp. 199-212). Rotterdam: Sense.

Weinert, F. E. (2001): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, Basel, Bonn: Beltz, 17-31.



## **Expertise von Lehrkräften in Mathematik und in Naturwissenschaften zwischen professionellem Wissen, Noticing und Kompetenzkonstrukten**

Chair(s): Sebastian Kuntze (Päd. Hochschule Ludwigsburg), Markus Rehm (Päd. Hochschule Heidelberg)

Diskutant/in: Marc Kleinknecht (TU München)

Für das Gestalten kognitiv aktivierender Lernanlässe benötigen Lehrkräfte professionelles Wissen (Shulman, 1986); dies gilt selbstverständlich sowohl für die naturwissenschaftlichen Fächer als auch für das Fach Mathematik. Auf der Grundlage dieses Wissens sollten Lehrkräfte in der Lage sein, Unterrichtssituationen zu analysieren, Beobachtungen kriterienbasiert zu deuten und bezüglich solcher unterrichtsbezogener Wahrnehmungen wissensbasiert zu argumentieren.

Mit unterschiedlicher Schwerpunktsetzung hat sich dies in den Begriffen Noticing (z.B. Sherin, Jacobs, & Philipp, 2011; Berliner, 2001), Professional Vision (z.B. Sherin, 2007), „Usable Knowledge“ (Kersting et al., 2012) sowie dem Konstrukt der Analysekompetenz niedergeschlagen. Lindmeier (2013) bemängelt jedoch den teilweise noch nicht ausreichenden Theoriebezug dieser Ansätze. Nicht zuletzt hängt diese Beobachtung mit dem Spannungsfeld zwischen der Handlungsnähe bei der Erhebung professionellen Wissens einerseits und dessen Theoriebezug andererseits zusammen, wobei nicht zuletzt unterschiedliche Komponenten professionellen Wissens auch innerhalb der Bereiche Pedagogical Knowledge (PK), Content Matter Knowledge (CK) sowie Pedagogical Content Knowledge (PCK) differenziert werden sollten.

Im Symposium wird dieser Problemkomplex aus den Perspektiven unterschiedlicher Domänen (Mathematik, Naturwissenschaften), verschiedener theoretischer Rahmenkonzeptionen und unterschiedlicher untersuchungsmethodischer Ansätze betrachtet. Die gemeinsame Fragestellung lautet: Über welches professionelle Wissen verfügen angehende bzw. praktizierende Lehrkräfte aus der jeweiligen Perspektive und wie kann dies im Spannungsfeld zwischen theoretischen Ansätzen und forschungsmethodischer Operationalisierung beschrieben werden? Den Vorträgen im Symposium ist außerdem gemeinsam, dass innovative Zugänge und Erhebungsformate gewählt wurden, die auch Verknüpfungen zwischen Aspekten professionellen Wissens mit in den Mittelpunkt rücken.

So konzeptualisiert der Beitrag von Markus Rehm und Markus Wilhelm ausgehend von einem vignettenbasierten Befragungsformat Aspekte von CK, PK und PCK bezüglich komplexer Szenarios im naturwissenschaftlichen Unterricht. Bei der Analyse der Daten stehen die Frage nach Zusammenhängen von Sub-Konstrukten und die Ausschärfung des Begriffsumfanges von PCK im Vordergrund.

Auf ein videobasiertes Befragungsformat fokussiert der Beitrag von Marita Friesen et al.: Hier wird untersucht, welche Aspekte einer inhaltspezifisch operationalisierbaren Analysekompetenz bei Lehramtsstudierenden beobachtet werden können. Ein theoriebasiertes hierarchisches Kompetenzmodell wurde hierfür entwickelt, dessen Ausprägungsumfang in einer explorativ orientierten Vorstudie geprüft wird.

Auf fachdidaktisches Kriterienwissen bezogenes Noticing wird im Beitrag von Anika Dreher in den Blick genommen. Bei diesem Ansatz wurden wiederum Textvignetten gewählt. Die Ergebnisse geben nicht nur Aufschluss über Noticing und das jeweils genutzte professionelle Wissen, sondern es zeigen sich u.a. korrelative Zusammenhänge zu weniger unterrichtssituationsbezogenen Sichtweisen der Lehrkräfte.

Der Bereich von CK ist im Beitrag von Sebastian Kuntze zentral: Hier wurde in Form des Wissens zu Verknüpfungen entlang übergreifender fachlicher Konzepte eine bislang kaum erforschte Fachwissenskomponente adressiert, wofür ein spezielles Befragungsformat entwickelt wurde.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35, 463–482.

Kersting, N., Givvin, K., Thompson, B., Santagata, R., & Stigler, J. (2012). Measuring Usable Knowledge: Teachers' Analyses of Mathematics Classroom Videos Predict Teaching Quality and Student Learning. *American Educational Research Journal*, 49(3), 568–589.

Lindmeier, A. (2013). Video-vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel & K. Macha (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken*. Münster: Waxmann.

Sherin, M.G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S.J. Derry (Eds.), *Video research in the learning sciences* (pp. 383-395). Mahwah, NJ: Lawrence Erlbaum.

Sherin, M.G., Jacobs, V., & Philipp, R. (2011). *Mathematics Teacher Noticing*. New York: Routledge.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educ. Researcher*, 15(2), 4-14.



Markus Rehm (Päd. Hochschule Heidelberg), Markus Wilhelm (Päd. Hochschule Luzern), Christian Vollmer (Päd. Hochschule Heidelberg), Katrin Bölsterli (Päd. Hochschule Luzern), Dorothee Brovelli (Päd. Hochschule Luzern), Benjamin Tempel (Päd. Hochschule Heidelberg)

## **Naturwissenschaften unterrichten können - empirische Validierung eines PCK Modells für den naturwissenschaftlichen Unterricht**

### Theoretischer und empirischer Hintergrund

Seit einigen Jahren wird der naturwissenschaftliche Unterricht bspw. in Baden-Württemberg in einem Fach organisiert, das die drei Naturwissenschaften integriert. Es wird spekuliert, ob die disziplinäre Ausbildung (z.B. nur Biologie) der im integrierten Fach unterrichtenden Lehrkräfte ein Grund für den hohen Anteil fachfremden Unterrichts ist. In der Folge der IQB-Vergleichsuntersuchung 2012 wurden plausible Argumente laut, dass fachfremdes Unterrichten mangelnde Schülerleistungen nach sich zieht. Um die Spekulationen in Evidenz zu überführen, benötigt man Testergebnisse zu den fachlichen und fachdidaktischen Fähigkeiten der Lehrkräfte für den integrierten naturwissenschaftlichen Unterricht. Dazu wurde ein vignettengestütztes Testformat entwickelt, das sich einerseits am Professionswissen als Teil des Kompetenzmodells bei Baumert & Kunter (2006) orientiert und somit von den ursprünglich durch Shulman (1986;1987) postulierten Lehrerfähigkeiten (CK, PCK, PK) für einen guten Fachunterricht ausgeht. Andererseits liegt dem Test – in Anlehnung an Kunter und Trautwein (2013) – die Unterscheidung zwischen Sichtstruktur (z.B. Unterrichtsmethoden) und Tiefenstruktur (z.B. kognitive Aktivierung) qualitativ guten Unterrichts zu Grunde.

### Forschungsfragen

Für den naturwissenschaftlichen Unterricht wird das fachdidaktische Wissen (PCK) als Primärkompetenz für einen hochwertigen Fachunterricht angesehen. Ungeklärt bleibt, was PCK wirklich ist. Eine eigenständige Domäne als Schnittmenge von PK und CK (integratives Modell) oder ein Amalgam aus PK und CK (transformatives Modell) (Gess-Newsome, 1999). Für den naturwissenschaftlichen Unterricht stellt sich die Frage: Welche Struktur (integrativ oder transformativ) lässt sich für ein PCK Modell aus den Testdaten des oben genannten Vignettentests empirisch stützen?

### Methode / Design der Studie

Der erwähnte Vignettentest verwendet authentisch-komplexe, schriftliche Unterrichtsvignetten (Rehm & Bölsterli, 2014). Der Test enthält acht Vignetten, wobei jede Vignette mehrere Fähigkeitsaspekte testet. Die Probanden können insgesamt 42 kritisierbare Unterrichtsaspekte (Items) erkennen und sollen diese im Blick auf Verbesserungsmöglichkeiten schriftlich beurteilen (Brovelli et al., 2013). Es wurden Daten von 386 Studierenden erhoben. Davon wurden 152 Datensätze durch vier Rater der zwei an der Untersuchung beteiligten Pädagogischen Hochschulen anhand eines Partial-Credit Modells, das dem mehrstufigen Ratingformat des Kodiermanuals entspricht geratet (Cohens Kappa = .75). Mit den 42 bewerteten Unterrichtsaspekten (Items) wurde eine multidimensionale Skalierung durchgeführt und die erhaltene Struktur mit den in der Forschungsfrage formulierten PCK-Modellen (integrativ oder transformativ) verglichen. Eine explorativ-konfirmatorische Faktorenanalyse mit AMOS-21 (Arbuckle, 2010) generiert ein Strukturgleichungsmodell, das weitere Angaben über die Struktur des zu Grunde liegenden PCK-Modells liefert.

### Ergebnisse

Die Faktorenanalyse führte einerseits zu einer Reduktion der Items (N=25), mit einem für eine Mixed-Methode-Ansatz befriedigenden Cronbach's  $\alpha$  von .617 sowie andererseits zu einem Modell mit gutem Fit der Indizes wie RMSEA  $<.05$ . Das Modell weist 7 Konstrukte auf, wobei eines der Fachkompetenz (CK), und je drei der Pädagogisch-psychologischen Kompetenz (PPK) bzw. der Fachdidaktischen Kompetenz (PCK) zuzuordnen sind. Je eine Subskala des Konstruktes PPK und eine Subskala des Konstruktes PCK entsprechen der Sichtstruktur des Lehrerhandelns, also der Unterrichtsinszenierung. Sie korrelieren gegenseitig hoch ( $r=.77$ ), weisen aber weder signifikante Korrelationen zu den anderen Skalen der gleichen Konstrukte auf, noch zu jenem von CK. Demgegenüber finden sich zahlreiche signifikante Korrelationen innerhalb der Skalen der Tiefenstruktur von PK und PCK ( $.70 < r < .80$ ) als auch zum Konstrukt CK ( $.49 < r < .88$ ).

### Diskussion

Die Ergebnisse der Modellvalidierung geben Hinweise, dass dem Konstrukt PCK ein eher integratives als ein transformatives Modell zu Grunde liegt. Dieses PCK-Modell ist allerdings in sich nicht homogen, da wir aufzeigen können, dass die Sichtstruktur von PCK (z.B. die Subskala fachspezifische Unterrichtsmethode) mit PK korreliert, während Tiefenstruktur mit den Subskalen kognitive Aktivierung und fachspezifische Diagnose mit CK korrelieren. Im Blick auf die Diskrepanzen der Originalmodelle (integrative bzw. transformative PCK-Modell (Gess-Newsome, 1999) und Sicht- und Tiefenstruktur qualitativ guten Unterrichts (Kunter & Trautwein, 2013), muss letztlich diskutiert werden, ob das hier empirisch validierte PCK-Modell die theoretischen Modelle (integrativ bzw. transformativ) vereint bzw. ob es in angemessener Form Sicht- und Tiefenstruktur des Unterrichts abbildet.



## **Aspekte fachdidaktischer Analysekompetenz von Lehramtsstudierenden bezogen auf den Umgang mit Darstellungen im Mathematikunterricht - Das Projekt ANAKONDA - M**

### Theoretischer Hintergrund

Darstellungen mathematischer Objekte spielen für Lernen und Kompetenzaufbau im Mathematikunterricht eine zentrale Rolle (z.B. KMK, 2004; Duval, 2006). So stützt sich etwa ein konzeptionelles Verständnis als Grundlage flexibel einsetzbaren Begriffswissens darauf, dass Lernende auf verschiedene Darstellungen zurückgreifen, zwischen Darstellungen wechseln und diese zueinander in Beziehung setzen können (Duval, 2006). Um das Lernpotential von Darstellungen voll ausschöpfen zu können, benötigen Lehrkräfte professionelles Wissen zum Umgang mit Darstellungen (Dreher & Kuntze, 2013) sowie die Kompetenz, Lehr- und Lernsituationen des Mathematikunterrichts bezogen auf den Umgang mit Darstellungen analysieren zu können. Analysieren wird hier in Anlehnung an das Konzept der professionellen Wahrnehmung als Konstrukt aus Noticing und Knowledge-based Reasoning konzipiert (Sherin et al., 2011): Durch den „professionellen Blick“ von Lehrkräften muss zunächst wichtige von unwichtiger Information getrennt werden, um dann aus der wahrgenommenen Situation auf Grundlage des vorhandenen Wissens Schlussfolgerungen ziehen zu können (vgl. Lindmeier, 2013). Videobasierte Befragungsformate (vgl. z.B. Kersting et al., 2012) ermöglichen es, die Analysekompetenz von Lehramtsstudierenden bezüglich des Umgangs mit Darstellungen in Lernsituationen zu beobachten und auch Rückschlüsse auf deren Professionswissen ziehen zu können. Der Einsatz von Videovignetten in der Unterrichtsforschung und Lehrerbildung liegt dabei insgesamt vor allem in der relativ authentischen Darstellung von Unterrichtsrealität und -komplexität begründet (z.B. Oser et al., 2009). So stellen Oser und Salzmann (2009) Möglichkeiten vor, über die Auswertung von Wahrnehmungen von Lehrkräften zu Unterrichtsvideos Aussagen über deren Kompetenzprofil zu treffen. Im Projekt ANAKONDA-M (Analysekompetenz nutzen in Bezug auf Darstellungen im Mathematikunterricht) werden bestehende Ansätze auf eine bereichsspezifische fachdidaktische Kompetenz bezogen, wodurch eine verstärkte theoretische Verankerung dieses Kompetenzkonstrukts möglich wird. Wir stellen eine teilweise explorativ angelegte Vorstudie dieses Projekts vor, in deren Rahmen dieses Kompetenzkonstrukt mit qualitativen und quantitativen Methoden geschärft werden soll.

### Fragestellungen

Untersucht werden vor diesem Hintergrund die Fragen, über welches professionelle Wissen zum Umgang mit Darstellungen im Mathematikunterricht Lehramtsstudierende und praktizierende Lehrkräfte verfügen, wie fachdidaktische Analysekompetenz als Teil von Lehrer(innen)expertise empirisch operationalisiert und abgebildet werden kann und inwiefern ein hierarchisch aufgebautes Kompetenzmodell für den Bereich fachdidaktische Analysekompetenz in Bezug auf Darstellungen bestätigt bzw. validiert werden kann. Die im Vortrag präsentierte Vorstudie verfolgt im Hinblick auf die Arbeit in ANAKONDA-M die folgende Fragestellung:

- Welchen Stellenwert haben Kriterien im Zusammenhang mit dem Umgang mit Darstellungen für Lehramtsstudierende, wenn sie die Qualität von Lerngelegenheiten in Unterrichtssituationen einschätzen sollen?
- Auf welches professionelle Wissen zum Umgang mit Darstellungen kann zurückgeschlossen werden und inwiefern kann Evidenz zu unterrichtssituationsbasierten Analyseprozessen Aspekten des oben angesprochenen Kompetenzkonstrukts zugeordnet werden?

### Methode und Untersuchungsdesign

Im Rahmen der Vorstudie wurden 31 Lehramtsstudierenden des Faches Mathematik jeweils zwei videografierte authentische Unterrichtssituationen gezeigt, in denen der Umgang mit Darstellungen und insbesondere der Wechsel zwischen verschiedenen Darstellungen eine zentrale Rolle spielte. Im Anschluss erhielten die Teilnehmer der Vorstudie einen Fragebogen, der sowohl offene Fragestellungen als auch Multiple-Choice-Items enthielt. Dabei wurden zunächst unspezifische Fragen zur Qualität der beobachteten Lernanlässe gestellt und anschließend gezielt nach dem Umgang mit Darstellungen gefragt.

Bei der Auswertung der Daten kamen Top-Down-Kodierungsverfahren bezüglich bestimmter Merkmale der Antworten zu den offenen Fragestellungen zum Einsatz; die geschlossenen Fragebogenteile wurden quantitativ ausgewertet. Aufgrund der teilweise explorativen Fokussierung dieser Vorstudie kamen daneben im Rahmen einer Bottom-up-Auswertung auch interpretative Methoden zum Einsatz.

### Ergebnisse

Die Auswertung der Antworten der Studierenden ermöglicht einerseits einen Einblick in deren fachdidaktisches Kriterienwissen, darüber hinaus aber auch in Ergebnisse von Analyseprozessen. Es zeigt sich, dass die Studierenden von sich aus kaum Kriterienwissen zum Umgang mit Darstellungen heranzogen und daher ihre Deutungen der Unterrichtssituationen auch kaum mit derartigem Wissen argumentativ stützten. Im Vortrag wird dies auch anhand von Bottom-up-Analysen von Antwortbeispielen nachvollzogen. Die Auswertung der geschlossenen Multiple-Choice-Fragebogenteile gibt Einblicke, inwiefern die Befragten dennoch über Beobachtungskategorienwissen zum Umgang mit Darstellungen verfügen.



## ***Erhebung von fachdidaktisch-kriterienbasiertem Noticing mit Hilfe von Textvignetten und Zusammenhänge mit Sichtweisen von Mathematiklehrkräften***

### Theoretischer Hintergrund

Noticing (vgl. van Es & Sherin, 2002) umfasst neben dem Bemerkten von bestimmten Geschehnissen im Mathematikunterricht auch das Beurteilen und Reflektieren dieser unter Rückgriff auf entsprechende professionelle Wissenskomponenten und Sichtweisen. Allan Schoenfeld (2011) hat in seinem Meta-Artikel über Studien zu „Mathematics Teacher Noticing“ deshalb dazu aufgefordert, Zusammenhänge zwischen Noticing einerseits und Sichtweisen und Wissen von Mathematiklehrkräften andererseits zu untersuchen. Diese Studie fokussiert dementsprechend auf fachdidaktisch-kriterienbasiertem Noticing und untersucht außerdem Zusammenhänge mit entsprechenden Sichtweisen von Lehrkräften. Als fachdidaktischer Bezugsrahmen für dieses Noticing wurde der Umgang mit vielfältigen Darstellungen gewählt, denn diese sind nicht nur zentral für das Gestalten von Lerngelegenheiten (vgl. Kuntze et al., 2011), sondern sie spielen eine Doppelrolle für das Lernen von Mathematik: Einerseits ist der flexible Umgang mit verschiedenen Darstellungen für ein mathematisches Objekt entscheidend für mathematischen Kompetenzaufbau (z.B. Lesh, Post, & Behr, 1987; Acevedo Nistal et al.), andererseits können aber Repräsentationswechsel auch das Verständnis von Lernenden behindern, vor allem, wenn sie nicht dabei unterstützt werden, Zusammenhänge zwischen diesen Darstellungen herzustellen (Ainsworth, 2006; Renkl et al., 2013; Duval, 2006). Am Beispiel des Inhaltsbereichs „Brüche“, auf den sich die inhalts- und situationsspezifischen Erhebungsteile der Studie konzentrierten, wurde vor diesem Hintergrund untersucht, inwiefern Lehrkräfte in situationsbezogenem Noticing fachdidaktische Analyse Kriterien bezüglich des Umgangs mit Repräsentationen zeigten.

Indem diese Studie sowohl angehende als auch praktizierende Lehrkräfte in den Blick nimmt, können fachdidaktisch-kriterienbasiertes Noticing und Zusammenhänge mit Sichtweisen auch in Bezug auf spezifische Unterschiede zwischen Experten und Novizen untersucht werden.

### Fragestellung

Die Studie konzentriert sich daher auf die folgenden Forschungsfragen:

- Erkennen angehende und praktizierende Lehrkräfte die potentiell hinderliche Rolle von inhaltlich nicht notwendigen Repräsentationswechseln für das Verständnis von Lernenden in spezifischen Unterrichtssituationen?
- Können Zusammenhänge zwischen solchem fachdidaktisch-kriterienbasierten Noticing und entsprechenden Sichtweisen bezüglich des Nutzens von Darstellungen festgestellt werden?

### Methode und Untersuchungsdesign

Das fachdidaktisch-kriterienbasierte Noticing wurde mit Hilfe eines vignettenbasierten Designs erhoben. Den teilnehmenden Mathematiklehrkräften (67 angehende und 77 praktizierende Gymnasiallehrkräfte) wurden Transkripte von fiktiven Unterrichtssituationen zum Thema Brüche vorgelegt, in denen jeweils ein aus fachdidaktischer Sicht für das Verständnis der Lernenden potentiell hinderlicher und inhaltlich nicht notwendiger Darstellungswechsel durchgeführt wird. Es wurde jeweils dazu aufgefordert den Umgang mit Darstellungen mit Blick auf das Verständnis der Lernenden begründet zu beurteilen. Die Antworten wurden von zwei Ratern mit hoher Inter-Rater-Reliabilität im Hinblick auf fachdidaktisch-kriterienbasiertes Noticing nach einem Top-Down-Verfahren kodiert: Dabei wurde zum einen kodiert, ob der Darstellungswechsel als solcher erkannt wurde und zum anderen, ob dieser kritisch reflektiert wurde. Traf beides zu, so wurde in Bezug auf die jeweilige Vignette spezifisches Noticing bescheinigt. Auf dieser Basis konnte über den gesamten Testteil hinweg ein entsprechender Score berechnet werden. Sichtweisen zur Rolle von vielfältigen Darstellungen für mathematisches Lernen wurden sowohl global im Sinne von Gründen für das Nutzen verschiedener Repräsentationen als auch inhaltsbereichsspezifisch im Hinblick auf den Bruchrechnenunterricht mithilfe eines bereits bestehenden Fragebogeninstruments im Multiple-Choice-Format erhoben (vgl. Dreher, Nowinska, & Kuntze, 2013; Dreher & Kuntze, 2013). Zusammenhänge entsprechend der zweiten Forschungsfrage wurden zum einen mit Blick auf Korrelationen zwischen Skalen zu Sichtweisen und dem Score für kriterienbasiertes Noticing untersucht, zum anderen aber auch durch qualitative Analysen aus Antworten der Lehrkräfte extrahiert, um explorativ zu analysieren, auf welches professionelle Wissen bzw. Sichtweisen die Lehrkräfte zurückgriffen.

### Ergebnisse

Die Ergebnisse zeigen einerseits, dass die potentiell hinderliche Rolle von Repräsentationswechseln in spezifischen Unterrichtssituationen relativ selten erkannt wurde. Andererseits konnte solches fachdidaktisch-kriterienbasiertes Noticing signifikant häufiger bei praktizierenden als bei angehenden Lehrkräften festgestellt werden ( $t(142)=4.43$ ,  $p<.001$ ,  $d=0.72$ ). Stärkere Zusammenhänge zwischen spezifischem Noticing und Sichtweisen bei den untersuchten praktizierenden Lehrkräften sprechen für eine höhere Kohärenz in deren professionellem Wissen.

## **Professionelles Wissen von Mathematiklehrkräften zum Analysieren von Unterrichtsinhalten im Hinblick auf übergreifende Konzepte**

### Theoretischer Hintergrund

Diese Studie fokussiert auf professionelles Wissen von Lehrkräften, das notwendig ist, um mathematische Inhalte mittels übergreifender Konzepte zu verknüpfen. Da Fachwissen in diesem Bereich empirisch bislang kaum mit quantitativen Mitteln untersucht wurde, ergänzen die Ergebnisse dieser Studie die Befundlage früherer Studien und können zur Weiterentwicklung theoretischer Modelle von Lehrerexpertise beitragen.

Mathematikbezogenes Metawissen ist nicht nur bedeutsam für verständnisvolles Lernen (Baumert & Köller, 2000), sondern es kann auch Lehrkräfte unterstützen, in kognitiv aktivierenden Lernanlässen den Aufbau kohärenten und strukturierten mathematischen Wissens zu fördern. Solches mathematikbezogenes Metawissen umfasst Wissen zu übergreifenden Konzepten sowie zum Verknüpfen dieser Konzepte mit curricularen (und auch extracurricularen) Einzelinhalten. Nach Shulman (1986a, 1986b, 1987) gehört dieses Wissen primär zum Fachwissen – von einer hohen Relevanz für Kompetenzen des Gestaltens kognitiv aktivierender Lernanlässe (Clausen, Reusser, & Klieme, 2003) ist jedoch auszugehen. In dem verfeinerten Modell von Ball, Thames und Phelps (2008) erscheint zunächst der Unterbereich des "Horizon Knowledge" als einschlägig für Wissen zum Analysieren und Verknüpfen mathematischer Einzelinhalte anhand übergreifender Konzepte. Da das Konstrukt des "Horizon Knowledge" nach Ball und Bass (2009) konzeptuell jedoch noch nicht voll entwickelt ist, besteht ein großer Bedarf nach empirisch handhabbaren Indikatoren in diesem Fachwissensbereich.

Diese Studie fokussiert daher auf professionelles Wissen zum Analysieren und Verknüpfen von Inhalten anhand übergreifender Konzepte, die in der Fachdidaktik als bedeutsam akzeptiert sind, wie beispielsweise „vielfältige Darstellungen Nutzen“, „mit Unendlichkeit Umgehen“, oder „Argumentieren und Beweisen“ (z.B. Kuntze et al., 2011a, 2011b; cf. NCTM, 2000; Schweiger, 2006; Schreiber, 1983). Solches Fachwissen ergänzt Analysen wie COACTIV (Baumert et al., 2010; Krauss et al., 2008; Kunter et al., 2011), die Fachwissen eher bezüglich spezifischer, nicht inhaltsübergreifender Inhalte in den Blick nahmen und bei den befragten Lehrkräften unter anderem starke schultypspezifische Unterschiede aufgezeigt hatten.

### Fragestellung

Die teilweise explorative Studie konzentriert sich daher auf die folgenden Forschungsfragen:

Über welches professionelle Wissen zum Analysieren und Verknüpfen von Einzelinhalten anhand übergreifender mathematikbezogener Konzepte verfügen Mathematiklehrkräfte?

Sind Unterschiede zwischen Lehrkräften verschiedener Schultypen zu erwarten?

Wie groß sind diese Unterschiede im Vergleich zum Wissen von Lehramtsstudierenden?

**Methode und Untersuchungsdesign:** Die Daten wurden im Rahmen des EU-Projekts ABCmaths erhoben, in dem ein neues Testinstrument für die Ziele dieser Studie entwickelt wurde (cf. Kuntze et al., 2011a). Befragt wurden 34 praktizierende Gymnasiallehrkräfte und 24 Lehrkräfte an Haupt- und Werkrealschulen (Baden-Württemberg). Für explorative Vergleiche wurden ferner zwei Stichproben von Lehramtsstudierenden verschiedener Schultypen einbezogen (N1=117, N2=42). Den Lehrkräften wurde ein Testinstrument mit sechs offenen Items vorgelegt, die in einem spezifischen Format gehalten waren. Dieses Format berücksichtigte, dass die Lehrkräfte mit den übergreifenden Konzepten nicht vertraut sein könnten: Um die Validität zu unterstützen, war daher jeweils ein Inhaltsbeispiel zusammen mit dem in dem Item angesprochenen übergreifenden Konzept gegeben und die Lehrkräfte wurden gebeten, möglichst viele Beispiele zum jeweiligen übergreifenden Konzept zu geben. Die Antworten wurden von zwei Ratern nach einem Top-down Kodierverfahren vordefinierten Kategorien zugeordnet, wobei die Unterscheidung zwischen Beispielen in inhaltlicher Nähe zum gegebenen Beispiel und adäquaten Beispielen in weiteren Inhaltsbereichen berücksichtigt wurde.

**Ergebnisse:** Die Ergebnisse zeigen, dass die meisten Gymnasiallehrkräfte nicht in der Lage waren, zumindest ein korrektes Beispiel pro Item anzugeben, d.h. das jeweilige übergreifende Konzept jeweils mit zumindest einem mathematischen Inhalt zu verknüpfen. Die Haupt- und Werkrealschullehrkräfte zeigten jedoch ein signifikant geringeres Fachwissen in diesem Bereich, was als Erweiterung entsprechender Ergebnisse aus COACTIV gedeutet werden kann.

Die angehenden Lehrkräfte waren oft nicht in der Lage, übergreifende Konzepte hinter mathematischen Inhalten zu erkennen und Inhalte entlang dieser übergreifenden Konzepte zu verknüpfen.

**Diskussion:** Wenn die Befunde auch nahelegen, dass sich Wissen zum Analysieren und Verknüpfen von Inhalten anhand übergreifender Konzepte im Laufe des Professionalisierungsprozesses verbessert, zeigen die Ergebnisse Weiterentwicklungspotentiale im professionellen Wissen von Mathematiklehrkräften auf.



## „Welche soll ich nehmen?“ – Zur relativen Wirksamkeit von Strategien zur Motivationsregulation

Chair(s): Malte Schwinger (Universität Marburg)

Theoretischer Hintergrund: Die Nutzung von Strategien zur Veränderung der eigenen Motivation wird allgemein als Vorgang der Motivationsregulation bezeichnet, worunter die mehr oder weniger bewusste Steuerung der eigenen Motivation – in der Regel mit dem Ziel der Anstrengungssteigerung – verstanden wird (Schwinger, von der Laden & Spinath, 2007; Wolters, 2003). Lernende können zahlreiche verschiedene Strategien zur Motivationsregulation einsetzen. Die Strategie Selbstbelohnung beinhaltet etwa eine selbst verabreichte Belohnung für die Erreichung eines bestimmten Ziels (z.B. nach getaner Arbeit abends ins Kino gehen). Umweltkontrolle beschreibt eine Strategie, bei der die Lernumgebung motivationsförderlich gestaltet wird (z. B. indem in der Bibliothek gelernt wird). Lernzielbezogene Selbstinstruktion bedeutet, sich durch die Erinnerung an das langfristige Ziel, seine Kompetenzen erweitern zu wollen, neu zum Lernen zu motivieren. In neueren Studien werden häufig acht motivationale Strategien voneinander unterschieden. Während die generelle Wirksamkeit der meisten dieser Strategien mehrfach belegt werden konnte (Schwinger & Stiensmeier-Pelster, 2012; Wolters & Benzon, 2013), wurde bislang jedoch noch nicht untersucht, welche der acht Strategien den größten Anteil an Varianz in motivationalen Kriterien aufklärt. Da die acht Strategien moderat bis hoch interkorrelieren, muss bei der Berechnung der relativen Effektivität jeder Strategie das Problem der Multikollinearität gelöst werden. Hierzu bietet sich das Verfahren der Relative Weights Analyse (Tonidandel & LeBreton, 2011) an, bei dem mittels Variablentransformation ein neues Set an orthogonal zueinander stehenden Prädiktoren erzeugt wird, welche dann keine Multikollinearität mehr aufweisen. Als Resultat erhält man für jeden Prädiktor den relativen Anteil der Kriteriumsvarianz, der ausschließlich auf diesen spezifischen Prädiktor zurückgeführt werden kann.

Fragestellung: In drei Studien soll mittels Relative Weights Analysen der Frage nachgegangen werden, welche von acht Strategien zur Motivationsregulation die größte relative Wirksamkeit für die Erhöhung der eigenen Motivation aufweist.

Methode: In Studie 1 (N = 531 Studierende, 362 weiblich, Alter: M = 24.19, SD = 3.87) wurden die anfängliche Anstrengungsbereitschaft mit der Subskala des LIST (Wild & Schiefele, 1994) sowie die acht motivationalen Strategien mit dem Fragebogen von Schwinger et al. (2007) querschnittlich erfasst. Die Studien 2 (N = 613 Studierende, 487 weiblich, Alter: M = 23.88, SD = 5.04) und 3 (N = 301 SchülerInnen, 169 weiblich, Alter: M = 17.55, SD = 0.69) wiesen zwei Messzeitpunkte auf. Zu t1 wurden die anfängliche Anstrengungsbereitschaft und die motivationalen Strategien erhoben, zu t2 wurde nochmals die Anstrengungsbereitschaft als zu prädizierendes Motivationskriterium erfasst. Im Gegensatz zu den ersten beiden Studien wurde in Studie 3 die Motivationsregulation spezifischer in Bezug auf einen konkreten Lernprozess untersucht. Alle Skalen erwiesen sich als zufriedenstellend reliabel ( $.68 \leq \alpha \leq .86$ ).

Ergebnisse: In allen drei Studien ergab sich die gleiche Rangreihe der acht motivationalen Strategien, wobei sich die Lernzielbezogene Selbstinstruktion mit Abstand als effektivste Strategie erwies. Ohne Kontrolle der vorherigen Anstrengungsbereitschaft erklärte diese Strategie zwischen 35 und 40% der Kriteriumsvarianz. Bei Berücksichtigung der vorherigen Anstrengungsbereitschaft als weiterem Prädiktor (Studien 2 und 3) klärte die Lernzielbezogene Selbstinstruktion immer noch 10 bzw. 13% zusätzliche Kriteriumsvarianz auf. Teilziele setzen erwies sich in allen Studien als die zweiteffektivste Strategie (16-22%, bis zu 9% bei Kontrolle vorheriger Anstrengungsbereitschaft), gefolgt von Annäherungs-Leistungszielbezogener Selbstinstruktion. Mit einem relativen Anteil an Kriteriumsvarianz zwischen 0.5 und 6% konnten hingegen die drei Strategien Steigerung des situationalen Interesses, Steigerung der persönlichen Bedeutsamkeit und Vermeidungs-Leistungszielbezogene Selbstinstruktion als relativ ineffektiv beurteilt werden. Insgesamt legen die Befunde eine relativ höhere Wirksamkeit (annäherungs-)zielbezogener Motivationsregulationsstrategien nahe, während Strategien, die auf eine Steigerung des Aufgabenwerts abzielen, weniger effektiv zu sein scheinen. Implikationen der Befunde für die weitere Forschung zum Thema Motivationsregulation werden diskutiert.

Schwinger, M., von der Laden, T. & Spinath, B. (2007). Strategien zur Motivationsregulation und ihre Erfassung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39, 57–69.

Schwinger, M. & Stiensmeier-Pelster, J. (2012). Effects of motivational regulation on effort and achievement: A mediation model. International Journal of Educational Research, 56, 35–47.

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. Journal of Business and Psychology, 26, 1-9.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. Educational Psychologist, 38, 189–205.

Wolters, C. A., & Benzon, M. B. (2013). Assessing and Predicting College Students' Use of Strategies for the Self-Regulation of Motivation. The Journal of Experimental Education, 81, 199-221.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
PEG-Gebäude  
Raum 1.G165

## **Unterrichtsinintegrierte Förderung von selbstreguliertem Lernen und Zahl-Größen-Kompetenzen im Anfangsunterricht**

Theoretischer Hintergrund und Fragestellung: Aus verschiedenen empirischen Studien ist bekannt, dass die Zahl-Größen-Kompetenz eine maßgebliche Vorläuferfähigkeit darstellt, wenn die späteren mathematischen Leistungen vorhergesagt werden sollen (z.B. Krajewski & Schneider 2006; Weißhaupt, Peucker, & Wirtz, 2006; Aunola et al., 2004; Krajewski, 2003). Aus diesem Grund ist es sinnvoll, numerische Basiskompetenzen bereits zu Beginn der Schulzeit gezielt zu fördern, um mathematischen Schwierigkeiten vorzubeugen (z.B. Van Luit & Van de Rijt, 1997). Ebenfalls hat sich gezeigt, dass eine kombinierte Förderung mathematischer Inhalte mit Inhalten zur Selbstregulation üblicherweise zu den besten Ergebnissen führt (z.B. Fuchs et al., 2003; Perels, Gürtler, & Schmitz, 2005). Daher soll in der vorliegenden Studie die Effektivität einer Intervention überprüft werden, die darauf abzielt, mathematische Basiskompetenzen sowie selbstregulative Fähigkeiten kombiniert im Anfangsunterricht zu fördern.

Da fächerübergreifende selbstregulative Inhalte als Transfervehikel für domänenspezifische Inhalte gelten (s. auch Hasselhorn & Hager, 2001), sollte die Reihenfolge der Förderkomponenten eine Rolle spielen. Dementsprechend soll in der vorliegenden Studie überprüft werden, inwiefern sich die Effekte der Intervention in Abhängigkeit von der Reihenfolge der Förderkomponenten unterscheiden. Es wurde davon ausgegangen, dass bessere Effekte erzielt werden, wenn zuerst die fächerübergreifenden Fähigkeiten und erst danach die mathematischen Inhalte trainiert werden (SRL + MZZ) im Vergleich zu der umgekehrten Reihenfolge (MZZ + SRL).

Methode: Zur Überprüfung der Fragestellung wurde ein längsschnittliches Interventionsdesign mit Prä- (T1) und Posttest (T2) sowie Follow-up (T3) realisiert. Die Intervention bestand aus der Kombination des Trainingsprogramms „Mengen, Zählen, Zahlen“ (MZZ; Krajewski, Nieding, & Schneider, 2007) zur Förderung von mathematischen Basiskompetenzen mit einer Intervention zur Förderung des selbstregulierten Lernens (SRL; in Anlehnung an Otto, 2007). Beide Trainingskomponenten wurden so adaptiert und konzipiert, dass sie in jeweils vier Doppelstunden durchgeführt werden konnten. Insgesamt dauerte die Intervention daher acht Wochen (eine Doppelstunde pro Woche) und wurde nach vorheriger Schulung der Lehrkräfte durch diese in ihren Klassen durchgeführt.

Insgesamt nahmen an der Studie 485 Erstklässler aus 30 hessischen Schulklassen teil. Die Schulklassen wurden eine der drei Bedingungen zugeordnet: (1) Kombiniertes Training (SRL + MZZ), (2) Kombiniertes Training mit umgekehrter Reihenfolge (MZZ + SRL) und (3) Kontrollgruppe ohne Intervention (KG). Sowohl beim Prä- und Posttest als auch bei dem Follow-up wurden neben den mathematischen Basiskompetenzen auch die selbstregulativen Fähigkeiten erfasst. Zur Erfassung der mathematischen Basiskompetenzen wurden der MBK-1 (Ennemoser, Krajewski & Sinner, in Vorbereitung), die Rechentreppe als Transferaufgabe sowie der DEMAT 1+ (Krajewski, Kuespert, & Schneider, 2002) durchgeführt. Die selbstregulativen Fähigkeiten wurden zu allen drei Messzeitpunkten mittels eines Schülerfragebogens (27 Items,  $\alpha=.82$ ) sowie einer Lehrereinschätzung für jeden Schüler (13 Items  $\alpha=.93$ ) erhoben. Zudem wurde zum ersten Messzeitpunkt ebenfalls die Intelligenz als Kontrollvariable durch den Culture Fair Intelligence Test (CFT, Scale 1; Catell, Weiß, & Osterland, 1997) erfasst.

Ergebnisse: Die gesammelten Daten wurden mit einer mehrbenenanalytischen Regression ausgewertet. Es zeigte sich auf Level 1, dass für die Vorhersage der mathematikspezifischen Kriterien jeweils die Intelligenz sowie der Vortestwert des MBK-1 signifikant waren, so dass diese bei den weiteren Regressionsanalysen als Kovariate auf Individualebene kontrolliert wurde. Auf Ebene 2 wurde die Gruppenzugehörigkeit als unabhängige Variable eingeführt. Für mehrere mathematikspezifische Kriterien zeigten sich positive Effekte zugunsten der Experimentalgruppe 1 (SRL + MZZ) im Vergleich zur Kontrollgruppe, diese ließen sich jedoch nicht für die Experimentalgruppe 2 (MZZ + SRL) nachweisen.

Im Hinblick auf das selbstregulierte Lernverhalten sind die Ergebnisse weniger eindeutig. Für die Selbstberichte zeigen sich unterschiedliche Muster: Im Posttest ist die Experimentalgruppe 2 (MZZ + SRL) überlegen, während im Follow-up die Experimentalgruppe 1 (SRL + MZZ) besser abschneidet. Bezüglich der Lehrereinschätzungen können hingegen keine signifikanten Effekte nachgewiesen werden.

## **Entwicklung berufsbezogener Selbstregulation über vier Jahre von Lehramtsstudium bis Lehrerberuf: Stabilität oder Veränderung?**

Die Bedeutung der Lehrkraft für den Lernzuwachs von Schülerinnen und Schülern wird vielfach betont (Hattie, 2009; Nye et al., 2004). Es konnte gezeigt werden, dass neben professionellem Wissen und Können auch die berufliche Selbstregulation einen Effekt auf die Qualität der Unterrichtsgestaltung und das Wohlbefinden von Lehrkräften hat (Baumert et al., 2010; Klusmann, Kunter, Trautwein, Lüdtke & Baumert, 2008; Kunter, Baumert et al., 2013; Kunter et al., 2013).

Berufliche Selbstregulation wird beschrieben als die Fähigkeit einer Person, in ihrem beruflichen Kontext effektiv mit den eigenen Ressourcen haushalten zu können (Klusmann, 2011). Basierend auf der Conservation of Resources Theory (COR-Theorie) wird postuliert, dass Personen mit einer adaptiven Selbstregulation sowohl hohes arbeitsbezogenes Engagement aufweisen, zur gleichen Zeit aber auch die Fähigkeit haben, ihre Ressourcen zu erhalten und sich von Arbeitsbelangen zu distanzieren (Widerstandsfähigkeit; Hobfoll, 1989). In Anlehnung an Arbeiten von Schaarschmidt (vgl. Schaarschmidt, Kieschke & Fischer, 1999) wurden vier verschiedene Typen der Selbstregulation vorgeschlagen und empirisch bei Mathematiklehrer/innen bestätigt (Klusmann et al., 2008): Der Gesundheitstyp weist hohes Engagement bei gleichzeitig hoher Widerstandsfähigkeit auf. Der Schontyp zeigt niedriges Engagement bei gleichzeitig hoher Widerstandsfähigkeit. Der Risikotyp A zeigt hohes Engagement bei gleichzeitig niedriger Widerstandsfähigkeit. Der Risikotyp B zeigt sowohl niedriges Engagement als auch niedrige Widerstandsfähigkeit. Die Konzeptualisierung der beruflichen Selbstregulation als generischer Aspekt professioneller Kompetenz von Lehrkräften impliziert ihre prinzipielle Erlern- und Vermittelbarkeit. Gleichzeitig wird das Instrument zur Erfassung der Selbstregulation in grundlegenden Teilen dafür verwendet, Lehramtsstudierende auf ihre Passung für das Lehramt zu prüfen, was eine gewisse Stabilität impliziert (Schaarschmidt & Kieschke, 2007). Die vorliegende Arbeit untersucht erstmals empirisch die Frage, ob berufsbezogene Selbstregulation eine stabile Disposition darstellt oder im Verlauf des Studiums und Vorbereitungsdienstes veränderbar ist.

### Fragestellungen

Das vorliegende Forschungsvorhaben beschäftigt sich mit folgenden Fragestellungen:

1. Können die vier Typen berufsbezogener Selbstregulation in der Phase des Studiums empirisch bestätigt werden?
2. Gibt es Veränderungen in der Zugehörigkeit zu den Typen berufsbezogener Selbstregulation über einen Zeitraum von 4 Jahren?
3. Gibt es Unterschiede in den Typveränderungen zwischen Lehramtsanwärter/innen und Personen, die nicht den Lehrerberuf anstreben?

### Methode

#### Stichprobe

Datengrundlage für die vorliegende Untersuchung bildet die Längsschnittstudie „Transformation des Sekundarschulsystems und akademische Karrieren“ (TOSCA; Köller, Watermann, Trautwein & Lüdtke, 2004), in der eine repräsentative Stichprobe von Abiturient/innen aus Baden-Württemberg seit dem Jahre 2002 in ihrer beruflichen und psychosozialen Entwicklung begleitet wird. Am ersten Messzeitpunkt nahmen insgesamt 4730 Schüler/innen teil. Seitdem wurden die Teilnehmer/innen alle zwei Jahre postalisch erneut befragt. Die vorliegende Studie basiert auf den Erhebungen der Jahre 2006, 2008 und 2010, da in diesen unter anderem die berufsbezogene Selbstregulation erfasst wurde. 1453 Personen, die sich im Jahr 2006 in einem Studium befanden (61.0% weiblich; 23.7% Lehramtsanwärter/innen) bilden die Analysestichprobe für die vorliegende Arbeit.

#### Instrumente

Berufsbezogenes Engagement und Widerstandsfähigkeit wurden mit Hilfe von acht Skalen des AVEM (Arbeitsbezogenes Verhaltens- und Erlebensmuster) erfasst (Schaarschmidt & Fischer, 1997; Cronbachs alpha:  $.68 < \alpha < .84$ ).

#### Ergebnisse

Zur Beantwortung der Fragestellungen wurden mittels Mplus (Muthén & Muthén, 2010) die Verfahren latent class analysis (LCA) und latent transition analysis (LTA) angewendet (Collins & Lanza, 2010). Die Ergebnisse weisen darauf hin, dass sich zu allen drei Messzeitpunkten eine 4-Klassen-Lösung mit den zu erwartenden Profilen bestätigen lässt. Im Mittel gehören 43.3% der Studierenden dem Gesundheitstyp, 12.5% dem Schontyp, 19.3% dem Risikotyp A und 24.9% dem Risikotyp B an. Des Weiteren zeigte sich in durchgeführter LTAs, dass in sehr geringem Maße Wechsel zwischen den vier Typen über einen Zeitraum von vier Jahren stattfinden. Die Wahrscheinlichkeit über einen Zeitraum von vier Jahren den Selbstregulationstyp zu wechseln liegt demnach im Mittel bei ca. 25%. Bezogen auf die Wahrscheinlichkeit für Typwechseln zeigen sich keine Unterschiede zwischen Lehramtsanwärter/innen und Personen, die nicht den Lehrerberuf anstreben.

Wolfgang Wagner (Eberhard Karls Universität Tübingen), Sarah Werth (Eberhard Karls Universität Tübingen), Ulrich Trautwein (Eberhard Karls Universität Tübingen), Oliver Lüdtke (Humboldt-Universität zu Berlin)

## ***Die selbstregulationsbezogene Lehrkräftefortbildung „Lernen mit Plan“: (Mediierte) Effekte auf allgemeine und trainingspezifische Unterrichtsmerkmale und Mathematikleistungen***

Verschiedene Lehrkräftefortbildungen haben sich empirisch als wirksam im Hinblick auf die Selbstregulationsförderung von Schülerinnen und Schülern sowie auf deren Leistungszuwachs erwiesen (Dignath & Büttner, 2008; Hattie, Biggs & Purdie, 1996). Bisher wurde unseres Wissens allerdings nicht untersucht, ob die Selbstregulationsförderung im Unterricht tatsächlich den Effekt des Trainings auf den Leistungszuwachs – wie theoretisch angenommen – mediiert (Spörer & Glaser, 2010). Dabei ist zu berücksichtigen, dass neben den fokalen Trainingsinhalten (selbstregulationsspezifische Unterrichtsqualität) auch Treatmenteffekte auf die allgemeine Unterrichtsqualität möglich sind (Werth et al., 2012), die in Folge ebenfalls zu einem Leistungszuwachs beitragen könnten. Ziel der vorliegenden Studie ist eine Überprüfung der beiden postulierten Mediationseffekte. Dabei wurde zunächst erwartet, dass das Lehrkräftetraining, das auf dem Selbstregulations-Modell von Schmitz und Wiese (2006) basierte, einen leistungsförderlichen Effekt im Hinblick auf die Mathematikkompetenz aufweist (Hypothese 1). Weiterhin sollte sich ein positiver Trainingseffekt sowohl auf die selbstregulationsspezifische (Hypothese 2) als auch auf die allgemeine Unterrichtsqualität (Hypothese 3) zeigen. Dabei wurde erwartet, dass der postulierte Leistungszuwachs über die selbstregulationsspezifische (Hypothese 4) sowie die allgemeine Unterrichtsqualität (Hypothese 5) mediiert werden. Die Analysen basieren auf Daten von 74 Hauptschulklassen (686 Schülerinnen und Schüler) des Projekts „Lernen mit Plan“, das aufgrund des Studiendesigns – ein Cluster-randomisiertes Design mit einer (Warte-)Kontrollgruppe und prä-post-follow up-Erhebungen – als besonders aussagekräftige Datenbasis betrachtet werden kann. Für die Analysen wurden Zweiebenen-Strukturgleichungsmodelle verwendet, die das Cluster-Design adäquat reflektieren: die Variable Treatment vs. Kontrollbedingung variiert designbedingt nur auf der Klassenebene, während die Unterrichtsmerkmale (hier aus Schülersicht) sowie die Mathematikleistung als Merkmal auf der Individual-, d.h. Schülerebene erhoben wurden (2-1-1-Mediationsmodell; Preacher, Zyphur & Zhang, 2011). Die Ergebnisse sind im Einklang mit den Hypothesen. Es konnten sowohl positive Effekte des Lehrkräftetrainings auf die Mathematikkompetenz sowie auf die Mediatoren allgemeine bzw. selbstregulationsspezifische Unterrichtsqualität nachgewiesen werden. Bezüglich der beiden indirekten Effekte ergab sich – bei einer p-Wert Adjustierung für multiples Testen (Schochet, 2008) – lediglich für die selbstregulationsspezifische, nicht aber für die allgemeine Unterrichtsqualität ein statistisch signifikanter positiver Effekt.

- Dignath, C. & Büttner, G. (2008). Components of fostering selfregulated learning among students. A meta-analysis on intervention studies at primary and secondary level. *Metacognition and Learning*, 3, 231–264. doi: 10.1007/s11409-008-9029-x
- Hattie, J. A., Biggs, J. & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66, 99–136. doi: 10.3102/00346543066002099
- Preacher, K. J., Zhang, Z., & Zyphur, M. J. Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(2), 161-182. doi: 10.1080/10705511.2011.557329
- Schmitz, B. & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Timeseries analyses of diary data. *Contemporary Educational Psychology*, 31, 64–96. doi: 10.1016/j.cedpsych.2005.02.002
- Schochet, P. Z. (2008). Guidelines for multiple testing in impact evaluations of educational interventions. Princeton, NJ: Mathematica Policy Research.
- Spörer, N. & Glaser, C. (2010). Förderung selbstregulierten Lernens im schulischen Kontext. *Zeitschrift für Pädagogische Psychologie*, 24 (3-4), 171-175. doi: 10.1024/1010-0652/a000014
- Werth, S., Wagner, W., Ogrin, S., Trautwein, U., Friedrich, A., Keller, S. et al. (2012). Förderung des selbstregulierten Lernens durch die Lehrkräftefortbildung «Lernen mit Plan»: Effekte auf fokale Trainingsinhalte und die allgemeine Unterrichtsqualität. *Zeitschrift für Pädagogische Psychologie*, 26(4), 291-305. doi: 10.1024/1010-0652/a000080

## **Der Einfluss induzierter Handlungsvorsätze auf die Veränderung des Zeitinvestments und der Lernzielerreichung: Eine experimentelle Feldstudie in der beruflichen Bildung**

### Theoretischer Hintergrund

Die aktuelle Studie befasst sich mit der Regulation selbstgesteuerten Lernverhaltens von SchülerInnen in der beruflichen Bildung. Dabei wird der Einfluss volitionaler Handlungssteuerung auf das Zeitinvestment im Lernprozess und auf die Erreichung von Lernzielen fokussiert.

Empirische Befunde der Lernzeitforschung zeigen, dass nicht nur ausreichendes sondern im Besonderen kontinuierliches Zeitinvestment in einem positiven Zusammenhang mit Lern- und Leistungserfolg steht. Lernhandlungen müssen ausreichend oft und regelmäßig ausgeführt werden, um gesetzte Lernziele erreichen zu können. Es ist außerdem bekannt, dass eine ausreichend hohe Motivation, ein gewünschtes Ziel erreichen zu wollen, die Voraussetzung für den Beginn einer (Lern-) Handlung darstellt. Doch auch trotz eines entsprechenden motivationalen Antriebs wird oftmals in problematischen Situationen eine Lernhandlung nicht begonnen oder unterbrochen. Laut aktueller Forschung liegt der Grund in volitionalen Defiziten in der Handlungssteuerung und Selbstregulation: Lernzeiten werden nicht geplant, in geeigneten Situationen die Initiative zu lernen nicht ergriffen, Lernabsichten nicht umgesetzt, nicht genug Anstrengung und Konzentration aufgebracht, oder selbstmotivierend in den eigenen Lernprozess eingegriffen. Die Konsequenz ist ein ineffizientes Lernverhalten, indem einerseits nicht ausreichend und andererseits nicht regelmäßig bzw. kontinuierlich genug Zeit in den Lernprozess investiert wird. Dadurch finden Lernhandlungen keinen erfolgreichen Abschluss und gesetzte Ziele werden nicht erreicht.

Mit Hilfe von spezifischer Zielbildung und zusätzlichen konkreten Handlungsplänen (Durchführungsententionen) kann das Volitionsniveau und dadurch selbstregulierendes Lernverhalten während der Handlungsiniiierung positiv beeinflusst und unterstützt werden. Indem zielgerichtete Handlungsvorsätze konkret die notwendigen Rahmenbedingungen und intendierten Handlungsabläufe festlegen, erhöht sich die Wahrscheinlichkeit des Handlungseintritts. Operational werden Handlungsvorsätze in einer Wenn-Dann-Form gebildet, um durch das „Wenn“ die Handlungssituation festzulegen und durch das „Dann“ die eigentliche Handlungsiniiierung regelhaft auszulösen. Handlungsvorsätze stärken somit die Zielrealisierung indem sie regelmäßiges, an eine wiederkehrende Situation gebundenes Lernen bedingen.

### Methodik

Es wird angenommen, dass mit einer pädagogischen Intervention zur Ziel- und Vorsatzbildung auf das Lernzeitinvestments und die Lernzielerreichung positiv eingewirkt werden kann. Neben der Unterstützung der Entwicklung eines individuell und zeitlich angemessenen Lernziels (Kontrollbedingung), soll eine ergänzende Formulierung und Festlegung von Etappenzielen und Handlungsvorsätzen (Experimentalbedingung) stattfinden. Als Erfolgskriterien der Intervention dienen quantitative Indikatoren des Lernverhaltens, wie regelmäßiges Zeitinvestment, Lernzeiterhöhung und Lernzielerreichung.

Die Stichprobe mit N=204 BerufsschülerInnen der 10. und 11. Jahrgangsstufe nahm an einer insgesamt jeweils siebenwöchigen Intervention und ihren begleitenden Datenerhebungen teil. Unmittelbar vor und nach der jeweiligen Interventionsphase wurden Informationen zur bisherigen Motivation und Selbstregulation der SchülerInnen mit umfassenden Fragebögen erhoben. Daten zum Lernverhalten (Lernzeit; Vorsatzumsetzung) und zur Etappenzielerreichung wurden wöchentlich mit einem online-gestützten Tagebuchverfahren erfasst.

### Ergebnisse

Die Auswertung der Längsschnittdaten erfolgte durch multivariate Verfahren, in dem ein Wachstumsmodell mit Hilfe der hierarchisch linearen Modellierung berechnet wurde. Die auf der Tagung vorgestellten Ergebnisse beziehen sich (a) allgemein auf die Veränderung des Zeitinvestments über die Zeit, (b) die Verläufe des Zeitinvestments und der Zielerreichung in Abhängigkeit des Treatments mit spezifischen Etappenzielen und Handlungsvorsätzen und (c) auf die moderierende Wirkung motivationaler Eingangsvariablen und bisheriger Selbstregulationsfertigkeiten im eigenen Lernprozess. Es zeigte sich, dass durch eine Intervention zur Ziel- und Vorsatzbildung das Lernzeitinvestment quantitativ erhöht und über einen längeren Zeitraum kontinuierlich auf einem hohen Niveau gehalten werden kann. Aufgrund der Vorsatzbildung werden Etappenziele in höherem Maße erreicht und die Aktivität im Lernprozess gesteigert. Besonders bei Probanden mit bisher eher niedrig ausgeprägten Selbstregulationsfähigkeiten konnte eine positive Wirkung des Treatments auf das Lernverhalten festgestellt werden.

Corno, L. (2001). Volitional aspects of self-regulated learning. In S. Zimmermann (Hrsg.), *Self-Regulated Learning and Academic Achievement* (S. 191–225). New Jersey: Lawrence Erlbaum Associates.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in experimental social psychology*, 38, 69–119.

Gollwitzer, P. M., & Oettingen, G. (2011). Planning promotes goal striving. In K. D. Vohs & R. F. Baumeister (Hrsg.), *Handbook of self-regulation: Research, theory, and applications* (S. 162-185). New York: Guilford.

Hofer, M. (2004): Schüler wollen für die Schule lernen, aber auch anderes tun. In: *Zeitschrift für pädagogische Psychologie*, 18, 79–92.

Mischo, C. (2006). Der Saisonarbeiter - genauer betrachtet. *Zeitschrift für pädagogische Psychologie*, 2, 97–110.

Wolters, C. (2003): Understanding procrastination from a self-regulated learning perspective. In: *Journal of Educational Psychology* (95), S. 179–187.

Zimmerman, B. J., Greenberg, D., & Weinstein, C. (1994). Self-regulating academic study time: a strategy approach. In D. H. Schunk & B. J. Zimmerman (Hrsg.), *Self-regulation of learning and performance* (S. 181–199). New Jersey: Lawrence Erlbaum.





## **Statusdynamiken und Bildungserfolg – Zur Dynamisierung der sozialen Herkunft**

Chair(s): Theresa Büchler (Universität Bremen)

Der Beitrag untersucht die Veränderungen der sozio-ökonomischen Lagen von Eltern im Zeitraum von vor der Geburt bis zum Ende der Grundschulzeit und deren Einfluss auf die Übergangentscheidung zur Sekundarstufe I. Analysen zum Einfluss der sozialen Herkunft auf den Bildungserfolg von Kindern unterstellen in der Regel ein statisches Verständnis sozialer Herkunft. In der jüngeren Sozialstrukturforschung hat es jedoch immer wieder Versuche gegeben, Sozialstruktur und soziale Ungleichheit grundsätzlich dynamisch zu denken, wobei diese Perspektive vor allem für das Paradigma der Lebenslaufforschung einflussreich geworden ist, aber auch in der „dynamischen Armutforschung“ (Leibfried et al., 1995). In der bildungssoziologischen Diskussion hat diese Perspektive jedoch bislang keinen nennenswerten Eingang gefunden, wenn es um die soziale Herkunft der Kinder geht. Diese wird fast immer zu einem bestimmten Zeitpunkt gemessen (Ausnahmen sind etwa Lange et al., 2002; Becker & Nietfeld, 1999). Diese empirische Vorgehensweise ist nicht nur vor dem Hintergrund der allgemeinen Diskussion um die Verzeitlichung sozialer Ungleichheit als problematisch zu betrachten, sondern auch vor dem Hintergrund der zentralen Bedeutung, die dem „Statuserhaltungsmotiv“ in der soziologischen Erklärung von Bildungsungleichheiten zukommt (v.a. Breen & Goldthorpe, 1997; Esser, 1999). Damit ist die Annahme gemeint, dass Eltern und/oder Kinder den Nutzen von möglichen Bildungsalternativen danach bemessen, inwiefern die jeweiligen Bildungsabschlüsse es ermöglichen, mindestens den (beruflichen) sozialen Status der Herkunftsfamilie zu erhalten. Damit wird nicht nur unterstellt, dass der soziale Status der Herkunftsfamilie als statisch zu betrachten ist, sondern auch, dass Eltern bzw. jungen Erwachsene vorherzusehen vermögen, welcher Bildungsabschluss nötig ist, um einen dem Herkunftsstatus äquivalenten zukünftigen Berufsstatus zu erlangen – und dies trotz Bildungsexpansion, berufsstrukturellem Wandel und damit einem Wandel in den Bildungsvoraussetzungen bestimmter beruflicher Statuspositionen.

Vor dem Hintergrund dieser Problematisierung untersucht der Beitrag, in welchem Ausmaß der „soziale Status“ von Familien über längere Zeiträume hinweg als stabil zu betrachten ist und inwiefern mögliche „Statusdynamiken“ einen Einfluss auf den Bildungserfolg der Kinder haben. Unter Ausnutzung von Längsschnittdaten des Sozio-oekonomischen Panels (SOEP) betrachten wir dabei einen Zeitraum von vier Jahren vor der Geburt eines Kindes bis zum Übergang des Kindes auf die Sekundarstufe I (als abhängiger Variablen). Wir stellen verschiedene Möglichkeiten vor, „Statusdynamiken“ der Herkunftsfamilie zu operationalisieren. Dabei betrachten wir die Einkommensverläufe, berufliche Karriereverläufe sowie eigene (Weiter-)Bildungsaktivitäten der Eltern. Um zunächst Komplexität zu reduzieren, betrachten wir nur stabile Elternpaare, klammern also Familienereignisse (Trennungen/Scheidungen, neue Partnerschaften, Patchworkfamilien) bewusst aus. Wir gehen zeitlich bis vor die Geburt des Kindes zurück, um mögliche „Nestbaueffekte“ einbeziehen zu können. In Erweiterung des statischen Modells des Statuserhalts gehen wir von einem dynamischen Status-Transmissionsmodell aus. Wir erwarten, dass sich zusätzlich zum durchschnittlichen Niveau des sozialen Status auch für die dynamischen Aspekte von Auf- und Abstiegen sowie Schwankungen Effekte auf den Bildungserfolg zeigen. Im Verhältnis zu Stabilität (als Referenzgruppe) sollten Aufstiege positive, Abstiege sowie generelle Schwankungen negative Effekte für den Bildungserfolg haben.

Die empirischen Analysen können diese Hypothesen jedoch nicht bzw. nur eingeschränkt bestätigen. Zwar finden wir einen negativen Effekt für abwärtsgerichtete Dynamiken, aber keinen positiven Effekt für aufwärtsgerichtete Statusdynamiken, also kontinuierliche Verbesserungen der Einkommensposition oder der Schichtung. Dieser Befund erweist sich – nach unserem bisherigen Analysestand – als robust gegenüber der Operationalisierung von Statusdynamiken wie auch der Wahl des Beobachtungszeitraums. Wir vermuten daher abschließend, dass soziale Aufstiegsdynamiken – im Vergleich zu stabilen Verläufen mit gleichem Durchschnittsniveau – möglicherweise mit negativen Effekten für die Förderung der Kinder (z.B. Stress, Leistungsdruck, Zeitmangel) einhergehen, so dass hier primäre und sekundäre Herkunftseffekte in gegensätzliche Richtungen weisen und sich neutralisieren. Wir wollen diese Vermutung abschließend mithilfe von Indikatoren zu elterlichen Sorgen und Zeitverwendung sowie der im Jugendfragebogen des SOEP erhobenen Angaben zur schulischen Leistung und zum elterlichen Unterstützungsverhalten überprüfen (die Ergebnisse zu diesem letzten Schritt liegen noch nicht vor).

Becker, Rolf & Nietfeld, Markus (1999): Arbeitslosigkeit und Bildungschancen von Kindern im Transformationsprozeß. Eine empirische Studie über die Auswirkungen sozio-ökonomischer Deprivation auf intergenerationale Bildungsvererbung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 51 (1): 55-79.

Breen, Richard & Goldthorpe, John H. (1997): Explaining educational differentials. Towards a formal rational action theory. Rationality and Society, 9 (3): 275-305.

Esser, Hartmut (1999): Soziologie - Spezielle Grundlagen. Band 1: Situationslogik und Handeln. Frankfurt a.M.: Campus.

Lange, Andreas, Lauterbach, Wolfgang & Becker, Rolf (2002): Armut und Bildungschancen. Auswirkungen von Niedrigeinkommen auf den Schulerfolg am Beispiel des Übergangs von der Grundschule auf weiterführende Schulstufen. In: Butterwegge, Christoph/ Klundt, Michael (Hg.): Kinderarmut und Generationengerechtigkeit. Opladen: Leske+Budrich, S. 153-170.

Leibfried, Stephan, Lutz Leisering, Petra Buhr, Monika Ludwig, Eva Mädje, Thomas Olk, Wolfgang Voges & Michael Zwick (1995): Zeit der Armut: Lebensläufe im Sozialstaat. Frankfurt am Main: Suhrkamp.

## ***Stärkere Selektion beim Übertritt von Grundschule auf das Gymnasium durch die Einführung des G8?***

Kaum eine Bildungsreform hat derartige Diskussionen ausgelöst wie die Einführung des G8 (Kühn, Ackeren, Bellenberg, Reintjes, & Brahm, 2013). In vielen Bundesländern ist die Reform bereits gerade erst umgesetzt und schon werden die Rufe nach einer Wiedereinführung des G9 laut (BLLV, 2013). Dies ist umso erstaunlicher, als es noch kaum Studien zu den Auswirkungen der Reform. Bisherige Studien zum Thema G8 beziehen sich vor allem auf Evaluationen von vereinzelten regional begrenzten Modellversuchen für (Hoch-) begabte Gymnasiast/innen, die kaum geeignet sind, um die Auswirkungen der Reform einzuschätzen (Homuth, 2012).

Das Ziel dieses Beitrags ist die Evaluation der Einführung des achtjährigen Gymnasiums (G8) hinsichtlich ihrer Auswirkungen auf Schullaufbahneinschätzung der Grundschullehrer. Am Ende der vierten bzw. der sechsten Klasse müssen Grundschullehrkräfte individuelle Schullaufbahneempfehlungen für ihre Schüler/innen abgeben. Dabei sollen sie eine Empfehlung für eine Sekundarschulform aussprechen, die dem Potential der Schüler/innen am besten entspricht. Die bisherigen IGLU-Studien haben gezeigt, dass Lehrkräfte sich orientieren neben den gezeigten Leistungen der Kinder auch an den Bildungsaspirationen der Eltern, welche bei Eltern aus höheren Schichten besonders stark ausgeprägt sind (Arnold, Bos, Richert, & Stubbe, 2010). Aufgrund der Bedeutung dieses Übergangs für die Bildungskarrieren ist der Einfluss die relativ starke Bedeutung der sozialer Herkunft (laut den Ergebnissen von IGLU 2001 immerhin als dritt-stärkster Faktor (Bos et al., 2004)) mit Sorge zu betrachten.

Es wird argumentiert, dass die Bildungsempfehlung der Lehrkräfte eine bewusste Entscheidung ist, welche innerhalb je nach Bundesland unterschiedlichen institutionellen Vorgaben stattfindet, und somit als Funktion von erwartetem Nutzen, Realisierungswahrscheinlichkeit und Kosten konzeptualisiert werden kann. Es wird weiterhin argumentiert, dass die Erwartungen der Lehrkräfte durch die Reform (insbesondere vor allem durch deren Berichterstattung) negativ beeinflusst werden, was zu einer Reduzierung der Empfehlungswahrscheinlichkeit führen sollte. Dies sollte besonders der Fall sein bei Schüler/innen, die weniger bildungsrelevante Ressourcen aufbringen können oder deren Leistungen „auf der Kippe“ stehen. Hier sollten „leistungsfremde“ Faktoren an Einfluss gewinnen.

Für die Analyse werden die nationalen Erweiterungsstudien von IGLU 2001 und 2006 genutzt ( $n(\text{gepooled}) = 13.051$ ). Der kausale Effekt der Reform auf die Schullaufbahneinschätzung wird mittels einer difference-in-differences-Strategie identifiziert. Dazu werden logistische Regressionsmodelle unter korrekter Berücksichtigung der komplexen Stichprobenstruktur geschätzt (Cameron, Gelbach, & Miller, 2011). Aufgrund der großen Anzahl von fehlenden Werten wurden diese mit der chained-equations-Methode (Royston & White, 2011) multipel imputiert.

Erste Ergebnisse zeigen, dass die G8-Reform unter Kontrolle von zentralen Einflüssen (Notendurchschnitt in Deutsch und Mathematik, Lesefähigkeit, Geschlecht, Alter und sozialer Status) entgegen den Erwartungen nicht zu einer generellen Tendenz von weniger Gymnasialempfehlungen durch die Grundschullehrkräfte führt. Allerdings zeigen sich durchaus Effekte auf das Verhalten der Lehrer: Lehrkräfte sprechen nach der Reform häufiger eine Empfehlung für das Gymnasium für Kinder mit hohem sozialen Status aus, während der Einfluss der bisherigen Leistungen in Form von Schulnoten abnimmt. Somit spielen leistungsfremde Kriterien aufgrund der Reform eine noch größere Rolle, was als Zeichen für eine Verschärfung der sozialen Selektivität gewertet werden muss.

- Arnold, K.-H., Bos, W., Richert, P., & Stubbe, T. (2010). Der Übergang von der Grundschule in die Sekundarstufe: Schullaufbahnpräferenzen von Lehrkräften und Eltern im Ländervergleich. In: Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K., Tarelli, I., & Valtin, R. (Hrsg.): IGLU 2006 – Die Grundschule auf dem Prüfstand. Münster u.a.: Waxmann, 13-32.
- Bayerischer Lehrer- und Lehrerinnenverband e.V. (2013). Durchgefallen - Vom Flexi-jahr halten Lehrer nichts. Pressemitteilung Nr. 17. München: BLLV.
- Bos, W., Voss, A., Lankes, E.-M., Schwippert, K., Thiel, O., & Valtin, R. (2004). Schullaufbahneempfehlungen von Lehrkräften für Kinder am Ende der vierten Jahrgangsstufe. In Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., and Walther, G. (Eds.), IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Münster: Waxmann.
- Cameron, A. C., Gelbach, J., & Miller, D. (2011). Robust Inference with Multiway Clustering. *Journal of Business and Economic Statistics* 29(2), 238-249.
- Homuth, C. (2012). Der Einfluss der Einführung des achtjährigen Gymnasiums auf die Kompetenzentwicklung. BAGSS Working Paper. Bamberg.
- Kühn, S. M., Ackeren, I. van, Bellenberg, G., Reintjes, C., Brahm, G. im (2013). Wie viele Schuljahre bis zum Abitur? *Zeitschrift für Erziehungswissenschaft* 16(1), 115 – 136.
- Royston, P., & White, I.R. (2011). Multiple imputation by chained equations (MICE): Implementation in Stata. *Journal of Statistical Software* 45(4).

## **Soziale Ungleichheiten an den Schnittstellen des österreichischen Schulsystems - Eine Zerlegung in primäre und sekundäre Effekte**

Das österreichische Schulsystem ist nach wie vor durch ein hohes Ausmaß an sozialer Selektivität gekennzeichnet, d.h. die formale Bildungskarriere von Kindern und Jugendlichen wird erheblich von deren sozialer Herkunft beeinflusst. Unter dem Begriff der sozialen Herkunft werden vornehmlich die klassischen vertikalen Ungleichheitsdimensionen der formalen Bildung, des beruflichen Status und des Einkommens der Eltern subsumiert (Geißler, 1990). Diese spiegeln sich in der differentiellen Verfügbarkeit von kulturellem, ökonomischem und sozialem Kapital wider (Bourdieu, 1983) und sind weiterhin mit Benachteiligungen in vielen Lebensbereichen verbunden. Zur Erklärung von sozialen Disparitäten in Bildungssystemen hat sich die Unterscheidung in primäre und sekundäre Herkunftseffekte als bedeutsam erwiesen (Boudon, 1974). Während primäre Effekte die über die herkunftsbedingt variierenden schulischen Leistungen vermittelten Einflüsse auf die Schulwahl repräsentieren, beziehen sich die nachgelagerten sekundären Effekte auf die bei schulbezogenen Entscheidungsprozessen divergierende „Kosten-Nutzen-Kalkulation“ der Eltern in Abhängigkeit von deren sozialer Position, selbst bei gleichen Leistungen der Kinder. Es wird erwartet, dass zu Beginn der Schullaufbahn (am ersten Übergang) die sekundären Ungleichheitseffekte dominieren, da die Kompetenzen der Kinder – vor allem von bildungsfernen Herkunftsfamilien – noch nicht präzise beurteilt werden können (Erikson & Jonsson, 1996). Am zweiten Übergang ist mit einer relativen Zunahme der primären Effekte zu rechnen, die nicht zuletzt den sinkenden perzipierten Kosten für die erfolgreiche Realisierung eines höheren Bildungsgangs geschuldet ist und die nun hinlänglich bekannten schulischen Leistungen als Entscheidungskriterium an Bedeutung gewinnen lässt.

Daraus können die folgenden Fragestellungen abgeleitet werden:

- (i) Lassen sich primäre und sekundäre Ungleichheitseffekte im österreichischen Schulsystem nachweisen? Falls ja, wie stark sind diese ausgeprägt?
- (ii) In welchem Verhältnis stehen die primären und sekundären Effekte? Kann der erwartete relative Rückgang der sekundären Effekte zwischen dem ersten und zweiten Bildungsübergang beobachtet werden?
- (iii) Verändern sich die Befunde, wenn anstelle der gemessenen Kompetenzen die Schulnoten als Indikatoren für die schulische Leistung herangezogen werden?
- (iv) Bleiben die Ergebnisse der primären und sekundären Effekte stabil, wenn unterschiedliche Prozeduren zu deren analytischer Trennung angewendet werden?

Als Datenbasis werden die österreichischen Ausgangsmessungen für die Überprüfung der Bildungsstandards in der 4. und 8. Schulstufe (BIST-BL4, 2010 und BIST-BL8, 2009) herangezogen (Breit & Schreiner, 2010, 2011). Zur Validierung der Befunde für den zweiten Übergang wird auf PISA 2012 zurückgegriffen. Dies ist insofern von Relevanz, dass mit BIST-BL8 lediglich der geplante Schulbesuch in der Sekundarstufe II berücksichtigt werden kann, während die PISA Daten die bereits realisierte Schulwahlentscheidung abbilden. Zur Dekomposition der primären und sekundären Herkunftseffekte wird auf drei statistische Verfahren rekurriert: (i) OLS-basierte lineare Regressionsmodelle (u.a. Bacher et al., 2012; Bruneforth et al., 2012), (ii) kontrafaktische Dekompositionsmodelle (u.a. Jackson et al., 2007; Schindler & Reimer, 2010) sowie (iii) die von Karlson und Holm (2011) vorgeschlagene Dekompositionsmethode.

Da die Daten aus PISA 2012 erst ab Anfang Dezember 2013 verfügbar sein werden, liegen erste Ergebnisse bislang nur für die Bildungsstandarterhebungen vor. Basierend auf linearen Modellen konnten an beiden Bildungsübergängen erhebliche herkunftsbedingte Effekte festgestellt werden. Während theoriekonform den sekundären Effekten beim ersten Übergang eine höhere relative Bedeutsamkeit zukommt, gleicht sich das Verhältnis beim zweiten Übergang weitgehend aus. Zudem zeigt sich, dass beim ersten Übergang der elterlichen Bildung ein höherer Einfluss zukommt, während beim zweiten Übergang der berufliche Status der Eltern als ungleichheitsgenerierende Dimension dominiert.

Gemäß den formulierten Forschungsfragen gilt es noch zu klären, wie sensitiv die Effekte gegenüber unterschiedlichen Messungen der schulischen Leistungen (Kompetenzskalen oder Schulnoten) sind und ob verschiedene Dekompositionsverfahren stabile Ergebnisse hervorbringen.

Bacher, J., Leitgöb, H. & Weber, C. (2012). Bildungsungleichheiten in Österreich. Vertiefende Analysen der PISA2009-Daten. In F. Eder (Hrsg.), PISA 2009. Nationale Zusatzanalysen für Österreich (S. 432-457). Münster: Waxmann.

Boudon (1974). Education, opportunity, and social inequality – changing prospects in Western society. New York: Wiley & Sons.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), Soziale Ungleichheiten (Soziale Welt, Sonderband 2, S. 183-198). Göttingen: Schwartz.

Breit, S. & Schreiner, C. (Hrsg.) (2010). Bildungsstandards: Baseline 2009 (8. Schulstufe). Technischer Bericht. BIFIE Zugriff am 30.10.2013 <https://www.bifie.at/buch/1116>.

Breit, S. & Schreiner, C. (Hrsg.) (2011). Bildungsstandards: Baseline 2010 (4. Schulstufe). Technischer Bericht. BIFIE Zugriff am 30.10.2013 <https://www.bifie.at/buch/1332>.

Bruneforth, M., Weber, C. & Bacher, J. (2012). Chancengleichheit und garantiertes Bildungsminimum in Österreich. In B. Herzog-Punzenberger (Hrsg.), Nationaler Bildungsbericht Österreich 2012. Band 2. Fokussierte Analysen bildungspolitischer Schwerpunktthemen (S. 187-226). Graz: Leykam.

Erikson, R. & Jonsson, J. (1996). Explaining Class Inequality in Education. The Swedish Test Case. In R. Erikson & J. Jonsson (Hrsg.), Can education Be Equalized? The Swedish Case in Comparative Perspective (S. 1-63). Stockholm: Westview Press.

Geißler, R. (1990). Schichten in der postindustriellen Gesellschaft. In P. A. Berger & S. Hradil (Hrsg.), Lebenslagen – Lebensläufe – Lebensstile (Soziale Welt, Sonderband 7, S. 81-102). Göttingen: Schwartz.

Jackson, M., Erikson, R., Goldthorpe, J. H. & Yaish, M. (2007). Primary and Secondary Effects in Class Differentials in Educational Attainment. The Transition to A-Level Courses in England and Wales. Acta Sociologica, 50(3), 211-229.

Karlson, K. B., & Holm, A. (2011). Decomposing primary and secondary effects: a new decomposition method. Research in Social Stratification and Mobility, 29(2), 221-237.

Schindler, S. & Reimer, D. (2010). Primäre und sekundäre Effekte der sozialen Herkunft beim Übergang in die Hochschulbildung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 62(4), 623-653.



## ***Der Einfluss des Explorationsverhaltens auf die Stabilität zentraler Indikatoren zur Erklärung sozialer Ungleichheit am Hochschulübergang***

In der Literatur zu sozialen Disparitäten beim Hochschulzugang finden die Wert-Erwartungsmodelle von Erikson und Jonsson (1996) sowie von Esser (1999) häufig Anwendung, um herkunftsspezifische Determinanten in der Bildungsentscheidung zu untersuchen. Neben dem wahrgenommenen Bildungsertrag und der Erfolgswahrscheinlichkeit, gelten vor allem die wahrgenommenen Kosten wie auch das Motiv des Statusverlusts als relevante Einflussfaktoren, welche Unterschiede in der Studienwahlentscheidung zwischen Schülerinnen und Schülern mit akademischen und nicht-akademischen Hintergrund erklären. Aus entwicklungstheoretischer Sicht handelt es sich bei der Studienwahlentscheidung jedoch nicht - wie in bisherigen Querschnittstudien unterstellt - um ein statisches Ereignis, sondern vielmehr um einen lang andauernden biographischen Prozess (Holling, 2000). In Anlehnung an das Rubikon-Modell von Heckhausen (1999) sollten insbesondere während des letzten Schuljahres vor dem Übergang sukzessive neue Informationen eine Selektion von Entscheidungsoptionen herbeiführen sowie eine mit dem Entschluss konforme Abwertung der nicht-gewählten Optionen bedingen. Inwiefern auch die für die Analyse von Ungleichheitsmechanismen verwendeten Parameter gewissen Veränderungen unterliegen, wurde bislang kaum empirisch untersucht. Der vorliegende Beitrag geht daher der Frage nach, inwieweit die Einschätzungen des Bildungsertrags, der Erfolgswahrscheinlichkeit, der Kosten sowie das Motiv des Statusverlusts zeitlich stabil sind, welchen Einfluss das Explorationsverhalten der Studienberechtigten auf die Entwicklung dieser Parameter hat und inwiefern sich Unterschiede zwischen den Herkunftsgruppen identifizieren lassen. Dazu wurden N = 261 Studienberechtigte zu zwei Messzeitpunkten im mittleren Abstand von 5 Monaten zu ihren studienbezogenen Einstellungen sowie zu ihrem Explorationsverhalten im letzten Schuljahr vor dem Erwerb der Hochschulreife befragt. Für die Operationalisierung des Wert-Erwartungsmodells wurde auf die Variablen von Becker (2000) zurückgegriffen. Die Explorationsaktivitäten wurden über die Subskalen Exploration in die Breite und Exploration in die Tiefe von Kracke und Dietrich (2008) erfasst.

Die zeitliche Stabilität der Indikatoren des Wert-Erwartungsmodells wurde über die Korrelation der Variablen zwischen beiden Messzeitpunkten ermittelt. Um den Einfluss des Explorationsverhaltens zu bestimmen, wurden multiple lineare Regressionen auf die Einschätzungen zum zweiten Messzeitpunkt berechnet, in die zum einen die Ausgangswerte der ersten Messung und zum anderen die Explorationsaktivitäten in die Breite und in die Tiefe eingingen. Darüber hinaus wurde geprüft, ob signifikante Interaktionseffekte zwischen den Herkunftsgruppen (Akademiker / Nicht-Akademiker) und dem Explorationsverhalten (in die Breite / in die Tiefe) vorliegen.

Die Ergebnisse zeigen, dass sich bei den Studienberechtigten akademischer Herkunft insbesondere das Motiv des Statusverlusts und bei den Studienberechtigten nicht-akademischer Herkunft vor allem die wahrgenommenen Kosten als zeitlich stabil erweisen. Der wahrgenommene Bildungsertrag ist dagegen in beiden Herkunftsgruppen über die Zeit weitestgehend instabil. Die Regressionsanalysen zeigen, dass das Ausmaß der in die Tiefe gehenden Explorationsaktivität einen signifikant positiven Einfluss auf die Einschätzung des Bildungsertrags zum zweiten Messzeitpunkt hat. Signifikante Interaktionseffekte zwischen der Herkunft und der Explorationsaktivität verweisen darauf, dass der Unterschied in der Einschätzung des Bildungsertrags bei steigender Explorationsaktivität zu Gunsten der Studienberechtigten mit akademischem Hintergrund anwächst. Bezogen auf die Entwicklung der subjektiven Erfolgswahrscheinlichkeit lässt sich dagegen beobachten, dass sich der Unterschied zwischen den Herkunftsgruppen bei ansteigender Explorationsaktivität in die Breite zu Gunsten der Studienberechtigten mit nicht-akademischem Hintergrund verringert. Die Studie belegt einerseits eine hohe Stabilität wahrgenommener Kosten und Statusverlustmotive und verweist andererseits auf eine Veränderbarkeit der wahrgenommenen Erträge und Erfolgswahrscheinlichkeiten, die sich zum Teil durch das Explorationsverhalten der Studienberechtigten erklären lässt. Damit liefern die Befunde erste Hinweise darauf, an welchen Stellen Interventionen ansetzen müssten, um soziale Ungleichheit am Hochschulübergang zu verringern. Eine Diskussion der Ergebnisse vor dem Hintergrund der Verringerung kognitiver Dissonanzen bei der Studien- und Berufswahlentscheidung sowie ein Ausblick auf weiterführende Analysen werden gegeben.

Erikson, R. & Jonsson, J. O. (1996). Explaining class inequality in education: The Swedish test case. In R. Erikson & J. O. Jonsson (Eds.), *Can education be equalized? The Swedish case in comparative perspective* (pp. 1-63). Boulder: Westview Press.

Esser, H. (1999). *Soziologie. Spezielle Grundlagen: Bd 1. Situationslogik und Handeln*. Frankfurt a.M.: Campus.

Holling, H.; Lüken, K.H., Preckel, F. & Stotz, M. (2000). Berufliche Entscheidungsfindung. Beiträge zur Arbeitsmarkt- und Berufsforschung. 236. Nürnberg.

Kracke, B. & Dietrich, J. (2008). *Career Exploration Scales*. Unpublished Manuscript.



## **Implementation von Maßnahmen zur Leseförderung: Perspektiven der Lesedidaktik und der Pädagogischen Psychologie**

Chair(s): Maik Philipp (Fachhochschule Nordwestschweiz), Elmar Souvignier (Universität Münster)

Diskutant/in: Andreas Gold (Universität Frankfurt/Main)

In dem Symposium soll die Frage nach der Implementation evidenzbasierter Konzepte zur Förderung der Lesekompetenz mit zwei Beiträgen aus der Deutschdidaktik und zwei Beiträgen aus der Psychologie aus einer interdisziplinären Perspektive und unter Anwendung eines breiten Spektrums unterschiedlicher Forschungsmethoden betrachtet werden. Vor einer Darstellung der Einzelbeiträge soll zunächst die Ausgangslage skizziert werden: In einer Vielzahl kontrollierter Trainingsstudien werden wirksame Maßnahmen zur Förderung der Lesekompetenz beschrieben. Die Ergebnislage lässt sich – verkürzend – so zusammenfassen, dass strategieorientierte Fördermaßnahmen ergänzt durch Prozeduren selbstregulierten Lernens in kooperativen Settings besonders erfolgversprechend sind (Slavin et al., 2008, 2009). Gleichzeitig weisen Befunde aus Unterrichtsbeobachtungen darauf hin, dass solche Prinzipien wirksamer Leseförderung kaum in der schulischen Praxis zu finden sind (z.B. Ness, 2008). Wenngleich die Frage nach Hindernissen für eine breite Implementation des Wissens über wirksame Leseförderung in die Praxis bereits vor einiger Zeit formuliert wurde (Kline, Deshler, Schumaker, 1992), gibt es bis heute kaum befriedigende Antworten – wie Lehrkräfte mit den Erkenntnissen und mit entsprechenden Materialien zur Leseförderung umgehen, wurde in ausgesprochen wenigen Untersuchungen thematisiert. Vor diesem Hintergrund gehen die Studien dieses Symposiums der Frage nach, wie Lehrkräfte das Wissen über wirksame Leseförderung nutzen und wie sie die Implementation strukturierter Fördermaßnahmen reflektieren. Damit wird in den Beiträgen des Symposiums weniger die Perspektive eingenommen, nach der Wirksamkeit bestimmter Förderansätze zu fragen, sondern es wird vielmehr untersucht, was in der schulischen Praxis geschieht, wenn Lehrkräfte theoretisch begründete, innovative Förderansätze in den schulischen Alltag implementieren. In dem Beitrag von Scherf wird auf der Basis von Interviewdaten und Gruppendiskussionen eine Kategorisierung vorgenommen, wie Lehrkräfte mit lesedidaktischen Wissensangeboten umgehen. Auf diese Weise wird die Unterschiedlichkeit in der Rezeption von Prinzipien wirksamer Leseförderung analysiert. Philipp und Kollegen werten Rückmeldungen von Lehrkräften aus, die ein Leseförderprogramm implementierten, das auf dem Ansatz „Peer Assisted Learning Strategies“ basiert. Aus den mit den Lehrkräften durchgeführten Interviews werden Implementationshindernisse und der Umgang mit solchen Schwierigkeiten dargestellt. In der Studie von Schulte und Kollegen wird die Implementation einer individualisierten Leseförderung untersucht, bei der Methoden zur Förderung der Leseflüssigkeit und des Leseverständnisses auf der Basis diagnostischer Informationen eingesetzt wurden. Anhand von Unterrichtsbeobachtungen und Befragungen von Lehrkräften und Schülern wird hier der Implementationsprozess analysiert. Auch Koch und Kollegen implementieren mit dem Reciprocal Teaching Ansatz – angereichert durch Selbstregulationsprozeduren – ein wirksames Förderkonzept und vergleichen diesen mit Lehrerfortbildungen, bei denen Lehrkräfte unter Anleitung ein strategieorientiertes Lesekompetenztraining entwickelten. Neben Lehrerbefragungen zur Einstellung gegenüber den Unterrichtskonzepten wurden Unterrichtsbeobachtungen und Erhebungen der Schülerleistungen durchgeführt.

- Kline, F.M., Deshler, D.D. & Schumaker, J.B. (1992). Implementing learning strategy instruction in class settings: A research perspective. In M. Pressley, K.R. Harris, & J.T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 361-406). San Diego, CA: Academic Press.
- Ness, M. K. (2008). Supporting secondary readers: When teachers provide the “what,” not the “how”. *American Secondary Education*, 37, 80-95.
- Slavin, R.E., Cheung, A., Groff, C. & Lake, C. (2008). Effective reading programs for Middle and High Schools: A best-evidence synthesis. *Reading Research Quarterly*, 43, 290-322.
- Slavin, R.E., Lake, C., Chambers, B., Cheung, A. & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research*, 79, 1391-1466.





## **Was Wissenschaftler wünschen und was Lehrer tun: Fallskizzen zum Umgang mit lese-didaktischen Wissensangeboten**

### Theoretischer Hintergrund

Zu erfassen, auf welchen Wissensgrundlagen Lehrende Leseförderung betreiben, ist aufgrund der Bedeutung des Lesens als kultureller Basiskompetenz von hohem gesellschaftlichen sowie wissenschaftlichen Interesse. Da die Lesedidaktik – derzeit die innovationsreichste und meistbeachtete deutschdidaktische Teildisziplin – mit ihren Konzeptionen zudem auf die Veränderung diesbezüglichen Lehrerwissens ebenso wie auf die Umgestaltung der schulischen Förderpraxis zielt (vgl. hierzu z. B. Ossner, 1999, Prediger et al., 2012), ist außerdem zu klären, wie die Verwendung der von ihr produzierten Wissensangebote durch Praktiker verläuft.

Was mit ihren Wissensprodukten geschieht, nachdem sie ins ‚unterrichtspraktische Feld entlassen‘ wurden, erforschte die Wissenschaft bisher allerdings genauso wenig wie die Wissensgrundlage praktischen lesefördernden Handelns – obschon mittlerweile als gesichert gilt, dass auf Akteursseite häufig gerade nicht-wissenschaftliches Wissen handlungsbestimmend ist und mit wissenschaftlichen Wissensangeboten zudem höchst eigensinnig umgegangen wird (vgl. hierzu z. B. Terhart et al., 1994; Schmelz, 2009, Bräuer, 2010).

### Fragestellung

Das Anliegen der hier vorzustellenden qualitativ-empirischen Studie war es daher, Lehrerwissen zu Leseförderung zu explorieren sowie im Besonderen den Umgang von Lehrenden mit innovativen lese-didaktischen Wissensangeboten zu rekonstruieren (vgl. Scherf, 2013).

### Methode

Im begrifflichen Rückgriff auf Wieser (2008) wurden aus Interview- sowie Gruppendiskussionsdaten Vorstellungen und Orientierungen zu Leseförderung von fünf Teilkollegien hessischer Gesamtschulen (n=21) extrahiert. Methodologisch orientierte sich die Studie dabei an der Grounded Theory und der Dokumentarische Methodologie (vgl. hierzu z. B. Strauss/Corbin, 1996; Bohnsack, 2003). Auf Grundlage eines breiten Wissensbegriffs, der subjektiv-wertende Anteile einschließt, wurden individuelle sowie in Teilkollegien sozial geteilt vorliegende (tlw. implizite) Wissensbestände zu Leseförderung und deren Veränderungen exploriert, fallvergleichend interpretiert sowie eine Typenbildung vorgenommen (vgl. hierzu z. B. Nohl, 2006).

### Ergebnisse

Im Zentrum des Vortrags sollen Erkenntnisse der Studie zum Umgang mit lese-didaktischen Wissensangeboten stehen.

Die vorgenommenen Rekonstruktionen zeigen, dass – obschon das Wissen der Lehrenden zu Leseförderung zahlreichen, auch nicht-wissenschaftlichen Quellgebieten entstammt – lese-didaktische Wissensangebote durchaus wahrgenommen und für (handlungs-)relevant gehalten werden. Der Umgang mit wissenschaftlichen Wissensangeboten gestaltet sich dabei allerdings keineswegs einheitlich – typologisch lassen sich verschiedene Ausprägungen aber deutlich voneinander abgrenzen.

Im Einzelnen wurden folgende Umgangsweisen mit lese-didaktischen Wissensangeboten rekonstruiert:

- Leseförderverfahren werden ‚gewusst‘, sie werden aber im Verweis auf schlechte Unterrichtsbedingungen nicht handlungswirksam.
- Neu erworbene lese-didaktische Wissensbestände werden in bereits bestehende (dominante) lesebezogene Konventionsrahmen eingepasst und in diesen Rahmungen im Unterricht wirksam.
- Neu erworbene lese-didaktische Wissensbestände stellen neue (dominante) lesebezogene Konventionsrahmen, an denen der Unterricht ausgerichtet wird.
- Neu erworbene lese-didaktische Wissensbestände stellen neue Orientierungsrahmen, die neben anderen im Unterricht Verwendung finden.

Die Erkenntnis, dass lese-didaktische Wissensangebote auf die dargestellten Weisen für lese-fördernde Akteure (handlungs-)relevant werden, führt zu der Frage, ob die Wissenschaft von einer solchen Verwendung ihrer Wissensprodukte auch ausgeht: Sollen Akteure durch sie (in einer der/den dargestellten Weisen) zum Handeln angeleitet werden? Wie dominant muss bzw. soll sich zudem das Lehrerhandeln an Normen ausrichten, welche den Konzeptionen inhärent sind?

Auf Grundlage der vorgelegten Erkenntnisse sowie einer (dokumentarischen) Dokumentenanalyse lese-didaktischer Wissensangebote wird abschließend aufgezeigt, dass es eine drängende Aufgabe darstellt, lese-didaktische Modellierungen professionellen Wissens zu explizieren und zu elaborieren – schon, da sich die Ansprüche der Wissenschaft mit den rekonstruierten Umgehensweisen der Praktiker kaum in Einklang befinden.



Maik Philipp (Fachhochschule Nordwestschweiz), Martin Brändli (Fachhochschule Nordwestschweiz), Katharina C. Kirchhofer (Fachhochschule Nordwestschweiz)

## **Scharniere – Sollbruchstellen – Strategieexperten? Zur Rolle der Lehrpersonen bei der Implementierung von Lesefördermaßnahmen**

### Theoretischer Hintergrund

Die Unterrichtsforschung zeigt, dass es einen klaren Bruch zwischen empirisch wirksamen Lesefördermaßnahmen einerseits und beobachtetem Leseunterricht ohne Intervention andererseits gibt – gewissermaßen eine Sollbruchstelle bei den Lehrpersonen und ihrem Unterricht. So gelten aus Sicht der Interventionsforschung kooperative Fördermaßnahmen sowie solche mit Fokus auf umfassende Wissensbestände (darunter zu Lesestrategien) als besonders gut geeignet, das Leseverstehen zu erhöhen. Im Sprachunterricht hingegen dominieren eher oberflächliche bzw. unspezifische und nur bedingt geeignet wirkende Förderpraktiken (Philipp, 2013). Damit stellt sich die Frage nach den Gründen für diese Diskrepanz. Möglicherweise sind es die Merkmale der in Interventionsstudien eingesetzten konkreten Förderansätze selbst, die die ökologische Validität der Fördermaßnahmen gefährden. Erst vereinzelt ist jedoch untersucht worden, wie Lehrpersonen sich zum einen aktiv (kooperative) Leseförderansätze aneignen (Duffy, 1993; Seymour & Osana, 2003) und zum anderen welche Modifikationen sie vornehmen und sie wie begründen (Hacker & Tenent, 2002; Marks et al., 1993). Solche Untersuchungen sind aber dringend nötig, um mehr darüber zu erfahren, wie wissenschaftlich gut begründete Maßnahmen in den Schulalltag Einzug halten können (Klingner, Boardman, & McMaster, 2013). Möglicherweise sind nämlich sogar andere Merkmale als die reine Treatmentintegrität entscheidender für den Erfolg, z. B. ein motivierender Unterricht (Souvignier & Mokhlesgerami, 2005), Leistungsrückmeldungen zum Lernerfolg (Behrmann & Souvignier, 2013) oder eben auch begründete Modifikationen (Klingner et al., 2013). Die Thematik Modifikationen und Hindernisse ist noch weitgehend unbearbeitet und deshalb besonders für qualitative Zugänge geeignet (Duffy, 1993; Hacker & Tenent, 2002; Marks et al., 1993; Seymour & Osana, 2003).

### Fragestellung

Im Rahmen einer Interventionsstudie wurden mit den Lehrpersonen der Experimentalgruppe Interviews durchgeführt, bei denen folgendes Erkenntnisinteresse bestand:

- Welche Elemente des Förderansatzes warfen Probleme bei der Implementierung und Durchführung auf?
- Wie sind die Lehrpersonen mit diesen Hindernissen umgegangen?
- Sind die Reaktionen auf die Hindernisse funktional oder dysfunktional?

### Methode

In einer Interventionsstudie implementierten Deutschlehrkräfte im Schuljahr 2012/2013 ein Leseförderprogramm, das auf dem Ansatz „Peer-Assisted Learning Strategies“ basiert (Fuchs & Fuchs, 2007). Neben der Frage nach der Effektivität, die für verschiedene abhängige Variablen über ein Prä-Post-Follow-up-Design mit randomisierter Zuweisung zu Experimental- und Wartekontrollgruppe bei leseschwachen Sechst- und Siebtklässlern ermittelt wurde, stellte sich jene nach der Umsetzung im Unterricht. Im qualitativen Forschungsteil des Projekts wurden am Ende der Förderdauer mit den N = 21 Lehrpersonen leitfadengestützte Experteninterviews durchgeführt (Gläser & Laudel, 2010). In den Interviews wurden systematisch Hindernisse, Modifikationen und Begründungen für etwaige Veränderungen erfragt. Diese Interviews werden derzeit inhaltsanalytisch ausgewertet.

### Ergebnisse

Zum derzeitigen Stand sind erst Teile der Interviews faktisch ausgewertet und die Ergebnisse dezidiert vorläufig. Aber schon jetzt zeichnen sich eindeutige Problembereiche aus den Daten ab. So sind einerseits die sozialen und intellektuellen Fähigkeiten der Jugendlichen ein Problem gewesen, weil Lehrpersonen wiederholt berichteten, dass schwache Leser überfordert waren. Andererseits scheinen auch Teile der Förderprozedur selbst problematisch, zum Beispiel die Texte und ihre Auswahl. Für die Intervention sind eigens Texte in verschiedenen Schwierigkeitsstufen verfasst worden, aus denen die Jugendlichen selbst auswählen konnten. Zum Teil konnten sie die zu trainierenden Lesestrategien bei schwierigeren Texten nicht adäquat anwenden. Eine (sinnvolle) Modifikation bestand darin, dass Lehrpersonen zum Teil zusätzliche Feedback-Prozeduren eingebaut haben, um den Jugendlichen die Richtigkeit der Strategieanwendung zu demonstrieren. Wie bereits erwähnt, sind dies erste Ergebnisse aus einem Teil der bislang ausgewerteten Interviews. Bis zur Tagung werden reichhaltigere Befunde zu erwarten sein, die systematisiert dargestellt werden. Damit trägt der Beitrag dazu bei, den Transfer wirksamer Lesefördermaßnahmen in den Schulalltag zu unterstützen und damit dafür zu sorgen, dass effektive Programme endlich im Schulalltag ankommen.

## **Implementation eines Förderkonzeptes zur diagnosebasierten individuellen Leseförderung in der Grundschule**

### Theoretischer Hintergrund

Untersuchungen zur Lesekompetenz im Grundschulalter weisen darauf hin, dass aufgrund der Leistungsheterogenität in diesem Alter individuell spezifischer Förderbedarf im Hinblick auf die Leseflüssigkeit oder das Leseverständnis besteht. Während sich zur Förderung der Leseflüssigkeit insbesondere Ansätze zum lauten Lesen (z.B. Repeated Reading) bewährt haben, erweisen sich strategieorientierte Ansätze (z.B. Reciprocal Teaching) als wirksam zur Förderung des Leseverständnisses. Es wäre daher naheliegend, auf der Grundlage diagnostischer Informationen eine individualisierte Leseförderung vorzunehmen, zumal die Wirksamkeit einer solchen Vorgehensweise generell nachgewiesen ist (Connor et al., 2007). Dennoch zeigen Unterrichtsbeobachtungen, dass in der schulischen Praxis der Einsatz evidenzbasierter Methoden, die Nutzung diagnostischer Information und Ansätze zur Individualisierung noch als Ausnahmen gelten (zusammenfassend Philipp, 2013). Studien zur Implementation von Prinzipien wirksamer Fördermethoden legen dabei nahe, dass insbesondere materialbasierte Ansätze erfolgversprechend sind, um Lehrkräfte zu unterstützen, solche Methoden in den Unterrichtsalltag zu integrieren (Souvignier & Trenk-Hinterberger, 2010; van Keer & Verhaeghe, 2005).

### Fragestellung

Vor diesem Hintergrund stellt sich die Frage, ob sich ein diagnosebasiertes individualisiertes Förderkonzept, das die genannten Elemente wirksamen Leseunterrichts berücksichtigt, in den Regelunterricht implementieren lässt.

### Methode

Zur Beantwortung dieser Frage wurde 43 Lehrkräfte im Rahmen von zwei Informationsveranstaltungen ein Konzept diagnosebasierten individualisierten Fördermaterials vorgestellt. Dieses Fördermaterial berücksichtigt die genannten Elemente wirksamer Leseförderung und besteht aus einer Methode zur Förderung der Leseflüssigkeit (Repeated Reading, im Material Lese-Sprinter genannt) sowie einer Methode zur Förderung des Leseverständnisses (Reziprokes Lehren, im Material Lese-Tandem genannt). Nach einer lehrergeleiteten Einführung in diese Methoden arbeiteten die Schüler eigenständig in Dyaden. Übungsmaterial wurde für beide Methoden auf je drei Schwierigkeitsstufen bereitgestellt. Die Zuteilung zu einer konkreten Methode und Schwierigkeitsstufe erfolgte diagnosebasiert im Rahmen sogenannter Trainingsgespräche auf der Grundlage von Leseleistungen (Leseflüssigkeit und Leseverständnis), die mit standardisierten Lesetests ermittelt wurden. Die Implementation des Förderkonzeptes wurde durch Unterrichtsbeobachtungen sowie durch eine Befragung von Lehrkräften und Schülern zum Material überprüft. Unterrichtsbeobachtungen konnten in 39 Klassen realisiert werden. Hier beurteilten jeweils zwei Beobachterinnen anhand eines vorgegebenen Beobachtungsinstrumentes die Umsetzungsgüte zentraler Aspekte der beiden Fördermethoden.

### Ergebnisse

Unterrichtsbeobachtungen der 242 beobachteten Sprinter- und 128 Tandemintervalle zeigten zunächst eine hohe Beobachterübereinstimmung, die zwischen  $.75 < K < .83$  liegt. Die Methode des Lese-Sprinters wurde in fast allen Klassen korrekt umgesetzt. Bei der richtigen Umsetzung des Lese-Tandems kam es im Vergleich dazu etwas häufiger zu Schwierigkeiten. Insgesamt wurden aber in nur drei von 39 Klassen mehr als zwei spezifische Aspekte der Methode nicht vorgabengetreu umgesetzt. Somit kam es in mehr als 92% der Klassen zu einer adäquaten eigenständigen Anwendung der Methoden des Lese-Sprinters und des Lese-Tandems.

Die Ergebnisse der Lehrerbefragung unterstützen diese positiven Ergebnisse. So gaben 35 der 43 Lehrkräfte an, im Mittel 35 Minuten ( $SD = 12$ ) und damit wie geplant ca. eine Schulstunde pro Woche mit dem Material gearbeitet zu haben. Darüber hinaus wurden - wie erwartet - durchschnittlich zwei Trainingsgespräche pro Schüler durchgeführt, um die Methode und die Stufe des Materials erneut der Leseentwicklung des Schülers anzupassen. Insgesamt bewerteten die Lehrkräfte das Fördermaterial im Hinblick auf die Aspekte a) der Qualität der Instruktion für beide Methoden, b) der Verbesserung der Schüler, c) des investierten Zeitaufwandes und d) der Weiternutzung des Lese-Sportlers im folgenden Schuljahr für alle Bereiche mit Werten  $>3$  auf einer vierstufigen Skala (von 1-4).

Ähnlich positiv bewerteten auch die Schüler die Anwendung des Materials. Auf einer vierstufigen Skala von 1 = stimmt gar nicht bis 4 = stimmt genau gaben sie zum einen an zu wissen, wie mit dem Material gearbeitet wird ( $M=3,78; SD=0,51$ ). Zum anderen stimmten sie der Aussage zu, sich durch das Training verbessert zu haben ( $M=3,44; SD=0,69$ ).

Die Ergebnisse der Unterrichtsbeobachtungen sowie der Befragungen sprechen daher gleichermaßen dafür, dass die Implementation des diagnosebasierten individualisierten Förderkonzeptes in den Regelunterricht gelang.

## ***Sind positivere Fortbildungserfahrungen auch die effektiveren? Regellehrkräfte als Implementierer. Effekte verschiedener Programme zur Förderung der Lesekompetenz von Fünftklässlern***

### Theoretischer Hintergrund

In der vorliegenden Implementationsstudie setzten Lehrkräfte zwei verschiedene Trainingsprogramme zur Förderung der Lesekompetenz von Fünftklässlern im Regelunterricht um. Eines der Treatments war das evidenzbasierte Programm ‚Reciprocal Teaching‘ (Palincsar & Brown, 1984), welches durch Selbstregulationsprozeduren angereichert wurde (RT+SRL), das andere war ein von Lehrkräften konzipiertes strategieorientiertes Lesekompetenztraining (Best-Practice, BP). Die Wirksamkeit des Programms RT+SRL im Vergleich mit No-Treatment-Kontrollklassen wurde unter der Bedingung, dass es externe Lerntrainer durchführen, bereits mehrfach belegt (Schünemann, Spörer, & Brunstein, 2013; Spörer & Schünemann, 2013). Der Nachweis, dass das Treatment RT+SRL auch unter der Bedingung effektiv ist, dass es Regellehrkräfte implementieren, steht noch aus. Forschungsbedarf besteht des Weiteren darin zu prüfen, wie wirksam das Training RT+SRL im Vergleich zu einem alternativen Treatment ist.

### Fragestellungen

Diesen Desiderata folgend sollten mit der Untersuchung folgende Fragestellungen beantwortet werden: (1.) Werden die Treatments den Vorgaben entsprechend implementiert? (2.) Zeigen sich bei Schülern der RT+SRL-Bedingung höhere Leistungszuwächse in Indikatoren der Lesekompetenz als bei Schülern die ein alternatives Treatment erhalten? (3.) Sind die Lesekompetenztrainings auch in dem Sinne wirksam, dass sich die Einstellungen der Lehrkräfte ändern?

### Methode

Insgesamt nahmen an der Studie N = 12 Lehrpersonen und N = 207 Schüler aus sechs Brandenburger Grundschulen teil. Das Design war ein Prä-, Post- (Lehrerdaten) bzw. ein Prä-, Post-, Follow-Up-Testdesign (Schülerdaten). Die Zuweisung der Schulen zu den Bedingungen erfolgte randomisiert. An drei Schulen wurde das Treatment RT+SRL und an drei anderen Schulen wurde das Treatment BP implementiert.

Vor den Implementierungen wurden die Lehrkräfte der jeweiligen Bedingung zunächst im Rahmen zweier akkreditierter Veranstaltungen fortgebildet. Für die Bedingung RT+SRL wurden sechs Lehrkräfte in einer Fortbildung nach dem mutual-adaptierten Ansatz (Snyder et al. 1992) geschult. Den Lehrkräften dieser Bedingung stand ein umfangreiches Trainingsmanual (Schünemann, Spörer & Brunstein, 2013) zur Verfügung. In der Bedingung BP nahmen vier der sechs anderen Lehrpersonen an einem Workshop teil. In diesem Workshop, welcher dem symbiotischen Implementationsansatz (Gräsel & Parchmann, 2004) folgte, konzipierten die BP-Lehrpersonen in kooperativer Zusammenarbeit und unter Supervision einer wissenschaftlichen Mitarbeiterin ein auf praxeologisch bewährten Elementen basierendes strategieorientiertes Lesekompetenztraining. Diese Programmkonzeption wurde in Stundenverlaufsplänen verschriftlicht. Nach Absolvieren der Fortbildungen implementierten die Lehrkräfte sodann ihr jeweiliges siebenwöchiges Treatment in den Regelunterricht für Fünftklässler im Fach Deutsch.

Implementationchecks (1.) wurden über Unterrichtsbeobachtungen durch geschulte Beurteiler (Beobachtungsprotokolle) gesichert und es wurden Lehrerselebensschätzungen nach jeder Unterrichts- bzw. Trainingsstunde (Unterrichtsprotokolle) erhoben (Fuchs et al., 2002). Die Indikatoren der Lesekompetenz der Schüler (2.) (Leseverständnis, Lesegeschwindigkeit, Kenntnis über und Beherrschung von Lesestrategien, lesebezogene Selbstwirksamkeit) wurden sowohl mit standardisierten Schulleistungstests als auch mit selbstkonstruierten Instrumenten erfasst. Die Einstellungen der Lehrkräfte (3.) wurden über schriftliche Befragungen ermittelt. Die Einstellungsmaße bezogen sich unter anderem auf die Arbeit mit Sachtexten im Unterricht (Souvignier & Mokhesgerami, 2005), auf das selbstregulierte Lernen (Steinbach & Stöger, 2011) und auf die Erfahrungen mit Fortbildungen (Klein-Heßling & Röder, 2007).

### Ergebnisse

Es konnte belegt werden, dass sich das evidenzbasierte Treatment den Vorgaben entsprechend in den Regelunterricht implementieren ließ. Auch die Best-Practice-Lehrer wichen bei der Umsetzung ihres Treatments nur marginal von der Programmkonzeption ab. Die Entwicklung der Lesekompetenz der Schüler verlief in beiden Bedingungen positiv. Für die Bedingung RT+SRL zeigten sich in den Indikatoren Leseverständnis und lesebezogene Selbstwirksamkeit Vorteileffekte gegenüber der Bedingung BP. Die Lehrer selbst profitierten auch von der Implementierung der Treatments in dem Sinne, dass sich Effekte für die Einstellungsmaße ergaben. RT+SRL-Lehrer erachteten nach dem Training direktiv-instruktionale und kognitiv-konstruktivistische Elemente im Unterricht als nahezu gleich wichtig, während die BP-Lehrer nach Abschluss des Trainings eher die kognitiv-konstruktivistische Herangehensweise bei der Lesekompetenzförderung bevorzugten. Darüber hinaus zeigte sich, dass die Erfahrungen mit Fortbildungen tendenziell positiver von denjenigen Lehrpersonen eingeschätzt wurde, die an der Fortbildung mit symbiotischem Ansatz teilnahmen.

## ***Das Verständnis von Wissenschaft und von experimentellen Designs im Grundschulalter: Zusammenhang und Einflüsse der Theory of Mind und kognitiver Variablen***

Chair(s): Christopher Osterhaus (Pädagogische Hochschule Freiburg)

Ein Verständnis von Wissenschaft (Nature of Science) und die Kompetenz in der Anwendung der experimentellen Methode werden als zentrale Aspekte des wissenschaftlichen Denkens betrachtet. Bis dato ist jedoch wenig bekannt über die Zusammenhänge dieser Kompetenzen in der Grundschule sowie über mögliche Faktoren, die diese Einzelkompetenzen befördern. In diesem Beitrag wird zunächst ein Instrument vorgestellt, das die Methodenkompetenz von Grundschulern reliabel erfasst.

Methodenkompetenz wird hier, im Gegensatz zu anderen Studien (z.B. Kuhn et al., 1996; Chen & Klahr, 1999), nicht nur als Kompetenz in der Anwendung von Strategien der Variablenkontrolle verstanden, sondern auch als metakonzeptuelles Verständnis des Experiments – ein Aspekt, der im Grundschulalter bisher nur vereinzelt (z.B. Koerber, Mayer, Osterhaus, Schwippert, & Sodian, eingereicht; Mayer, Sodian, Koerber, & Schwippert, 2014) betrachtet wurde. Zudem werden Zusammenhänge zwischen dieser Kompetenz und dem Wissenschaftsverständnis sowie anderer kognitiver Variablen (Theory of Mind, Intelligenz, Inhibition, Arbeitsgedächtnis, Textverständnis) untersucht, um ein differenziertes Entwicklungsmodell für diesen Altersbereich aufzustellen.

Studien aus dem frühen und späten Jugendalter (z.B. Bullock, Sodian, & Koerber, 2009) zeigen einen starken Zusammenhang zwischen dem Wissenschaftsverständnis von Schülerinnen und Schülern und ihrer Fähigkeit, kontrollierte Experimente zu produzieren. Im Grundschulalter konnte zum einen gezeigt werden, dass ein Training des Wissenschaftsverständnisses die Methodenkompetenz (Strategieproduktion) von Viertklässlern erhöht (Sodian, Thoermer, Kircher, Grygier, & Günther, 2002). Zum anderen legen rezente Studien (z.B. Koerber et al., eingereicht; Mayer et al., 2014) dar, dass das wissenschaftliche Denken in dieser Altersgruppe eine starke konzeptuelle Kohärenz aufweist, das heißt, es sich als eindimensionales Konstrukt fassen lässt. Während diese Studien einen Hinweis darauf geben, dass auch für das Grundschulalter von einem Zusammenhang auszugehen ist, bleibt unklar, ob dieser auch für ein umfassendes metakonzeptuelles Verständnis experimenteller Designs besteht und welche weiteren (kognitiven) Faktoren interindividuelle Unterschiede im wissenschaftlichen Denken bedingen.

Variablen, die in Zusammenhang mit dem wissenschaftlichen Denken von Grundschulern gebracht wurden, sind unter anderen Intelligenz (Koerber et al., eingereicht; Mayer, 2014) und Inhibition (Kuhn & Pease, 2008). Bezüglich dieser letzteren wird davon ausgegangen, dass sie einen Einfluss auf die Methodenkompetenz und insbesondere die Variablenkontrollstrategie hat (Kuhn et al., 2008). Hinsichtlich des wissenschaftlichen Denkens in seiner Gesamtheit konnte dahingegen kein Zusammenhang in einer Studie in der vierten Klassenstufe gefunden werden (Mayer et al., 2014). In einer längsschnittlichen Untersuchung grundlegender Fähigkeiten im wissenschaftlichen Denken im Vor- und Grundschulalter, konnten Piekny und Mähler (2013) zudem zeigen, dass frühe Kompetenzen in der Theory of Mind spätere Fähigkeiten im wissenschaftlichen Denken vorhersagten.

In einer Studie mit N = 402 Zweit-, Dritt- und Viertklässlern wurde der Zusammenhang des Verständnisses experimenteller Designs und des Wissenschaftsverständnisses untersucht. Zudem wurde die Intelligenz, das Arbeitsgedächtnis, die Inhibitionsfähigkeit sowie das Textverständnis und Fähigkeiten der Schülerinnen und Schüler in Aufgaben zur „fortgeschrittenen“ Theory of Mind erfasst.

Die Ergebnisse zeigen, dass sich sowohl das Verständnis experimenteller Designs als auch das Wissenschaftsverständnis eine reliable Skala bilden. Strukturgleichungsmodelle zeigten darüberhinaus, dass ein starker Zusammenhang zwischen dem Verständnis experimenteller Designs und dem Wissenschaftsverständnis besteht. Zudem ergaben sich signifikante Zusammenhänge zwischen dem Verständnis experimenteller Designs und der Inhibition und dem Alter der Schülerinnen und Schüler. Hinsichtlich der interindividuellen Unterschiede im Wissenschaftsverständnis zeigte sich dahingegen kein unabhängiger Einfluss des Alters; jedoch waren sowohl Intelligenz und Textverständnis als auch Fähigkeiten in der Theory of Mind signifikante Prädiktoren für diese Kompetenz.

Zusammenfassend zeigen die Ergebnisse einen starken Zusammenhang von Methodenkompetenz und dem Wissenschaftsverständnis von Grundschulern bei jedoch gleichzeitig unterschiedlichen Einflüssen verschiedener kognitiver Variablen. Während dieser Befund die konzeptuelle Kohärenz im wissenschaftlichen Denken unterstreicht, zeigt er, dass Schülerleistungen in verschiedenen Komponenten des wissenschaftlichen Denkens von diversen kognitiven Fähigkeiten abhängen und daher eine differenzierte Betrachtung einen Beitrag leisten kann bei der Herleitung eines Entwicklungsmodells.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70, 1098–1120.

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., & Sodian, B. (eingereicht). The development of early scientific reasoning. *Child Development*.

Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the society for research in child development*, 60.

Kuhn, D., & Pease, M. (2008). What needs to develop in the development of inquiry skills? *Cognition and Instruction*, 26, 512 – 559.

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, 29, 43–55.

doi:10.1016/j.learninstruc.2013.07.005

Piekny, J., Grube, D., & Mähler, C. (2013). The relation between pre-school children's false-belief and understanding and domain-general experimentation skills. *Metacognition Learning*, 8, 103-119. doi:10.1007/s11409-013-9097-4

Sodian, B., Thoermer, C., Kircher, E., Grygier, P., & Günther, J. (2002). Vermittlung von Wissenschaftsverständnis in der Grundschule. *Zeitschrift für Pädagogik*, 45. Beiheft, 192-206.





## Die Erfassung des Dichteverständnisses von Grundschulern mit schriftlichen Testaufgaben

Ein Verständnis von Dichte ist eine wichtige Voraussetzung für das Lernen anspruchsvoller Inhalte. Zum Erlangen dieses Verständnisses ist die Umstrukturierung von zuvor existierenden intuitiven Theorien notwendig (Smith, 2007). Studien belegen, dass bereits fünf- bis siebenjährige Kinder ein qualitatives Dichteverständnis erreichen können (Smith, Carey & Wisner, 1985). Dennoch finden sich häufig auch in höheren Altersstufen Probleme bei der Bearbeitung von entsprechenden Aufgaben und ein proportionales Dichteverständnis kann erst später erreicht werden (Smith, 2007). Eine gängige Schwierigkeit ist dabei eine mangelnde Differenzierung von Gewicht und Dichte: Die Dichte eines Materials wird lediglich anhand des (gefühlten) Gewichts beurteilt und nicht in Bezug zum Volumen gesetzt (Smith, 2007).

Zur Erfassung des Dichteverständnisses werden sowohl mündliche wie auch schriftliche Testaufgaben mit größtenteils offenem Antwortformat eingesetzt. In einem möglichen Aufgabentyp werden die Probanden gebeten aus Paaren von Gegenständen (z.B. Klötze) den mit der höheren Dichte auszuwählen (vgl. Hardy & Stern, 2011; Pollmeier, Hardy, Koerber & Möller, 2011). Das spezifische Verhältnis der Charakteristika der Gegenstände zueinander kann dabei die Urteile der Probanden signifikant beeinflussen:

Inkongruente Aufgaben, d.h. Aufgaben in denen der Gegenstand mit der höheren Dichte das geringere absolute Gewicht aufweist, waren schwieriger zu lösen als kongruente Aufgaben (der Gegenstand mit der höheren Dichte weist das höhere Gewicht auf) (Fassoulopoulos, Kariotoglou & Koumaras, 2003; Smith et al., 1985). Im vorliegenden Beitrag soll untersucht werden, ob das Merkmal Kongruenz Schwierigkeitsunterschiede zwischen geschlossenen, schriftlichen Aufgaben für die Grundschule aufklären kann.

In einem ersten Schritt wurden dazu die Antworten von 1820 Schülerinnen und Schülern der zweiten, dritten und vierten Jahrgangsstufe der Grundschule auf neun Dichteitems in einem Test zur Erfassung naturwissenschaftlicher Kompetenz einer explorativen Analyse mit explanatory item response models (Janssen, Schepers & Peres, 2004) unterzogen. Die Daten wurden im Rahmen des Projektes Science-P erhoben (Hardy et al., 2010; Kleickmann, Hardy, Pollmeier & Möller, 2011). Obwohl die Ergebnisse dieser Studie daraufhin wiesen, dass die Kongruenz einen signifikanten Einfluss auf die Aufgabenschwierigkeit hatte, lagen aufgrund des Untersuchungsdesigns einige methodische Probleme vor.

Daher wurden in einer anschließenden Experimentalstudie 12 dichotom kodierte Dichteitems, von denen jeweils drei die gleiche Merkmalsausprägung (inkongruent – gleiches Material – gleiches Gewicht – kongruent) trugen, 73 Schülerinnen und Schülern der vierten Jahrgangsstufe der Grundschule zur Bearbeitung vorlegt. Auch in dieser Studie wiesen inkongruente Items ( $M=.47$ ,  $SD=.06$ ) niedrigere Lösungshäufigkeiten auf als kongruente ( $M=.86$ ,  $SD=.05$ ). Die inferenzstatistischen Analysen zeigten, dass die Kongruenz einen Großteil der Schwierigkeitsvarianz der Items aufklärte ( $R^2=.91$ ) und dass kongruente Items signifikant leichter zu lösen waren als Aufgaben mit inkongruentem Gegenstandspaar ( $b=3.06$ ,  $p<.001$ ). Deskriptive Analysen der Antwortmuster der Probanden zeigten, dass von den 54 Probanden, die in der Lage waren alle kongruenten Items richtig zu lösen, nur 23 ebenso alle inkongruenten Items erfolgreich bearbeiten konnten. Von den restlichen 31 konnten allerdings 11 alle Items korrekt lösen, in denen das Gegenstandspaar das gleiche Gewicht aufwies (das Gewicht konnte somit nicht als Indikator für Dichte genutzt werden). Dieser Befund kann als Hinweis für ein Zwischenstadium des Verständnisses interpretiert werden, bei dem das richtige Verständnis noch nicht auf alle möglichen Kontexte ausgeweitet werden kann.

Die Ergebnisse zeigen, dass das Aufgabenmerkmal Kongruenz einen wichtigen Faktor zur Erklärung der Itemschwierigkeiten darstellt. Potentiell kann dieses Merkmal künftig dazu genutzt werden, gezielt Aufgaben unterschiedlicher Schwierigkeit zu entwickeln, die Validität von Instrumenten zur Erfassung des Dichteverständnisses zu verbessern und wertvolle Hinweise für die Modellierung von Kompetenzentwicklung liefern (Klieme & Leutner, 2006). Die gezielte Analyse von Antwortmustern auf Aufgaben mit unterschiedlicher Merkmalsausprägung kann darüber hinaus auf das Vorliegen unterschiedlicher intuitiver Theorien hinweisen (Smith, 2007).

Fassoulopoulos, G., Kariotoglou, P., & Koumaras, P. (2003). Consistent and inconsistent pupils' reasoning about intensive quantities: The case of density and pressure. *Research in Science Education*, 33, 71-87.

Hardy, I., Kleickmann, T., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., Schwippert, K., & Sodian, B. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. *Zeitschrift für Pädagogik*, 56. Beiheft, S.115-125.

Hardy, I., & Stern, E. (2011). Visuelle Repräsentationen der Dichte: Auswirkungen auf die konzeptuelle Umstrukturierung bei Grundschulkindern. *Unterrichtswissenschaft*, 39(1), 35-48.

Janssen, R., Schepers, J. & Peres, D. (2004). Models with Item and Item Group Predictors. In P. De Boeck & M. Wilson (Hrsg.), *Explanatory Item Response Models* (S. 198-212). New York: Springer.

Kleickmann, T., Pollmeier, J., Hardy, I. & Möller, K. (2011). Die Struktur naturwissenschaftlichen Wissens von Grundschulkindern – eine person- und variablenzentrierte Analyse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43(4), 200-212.

Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52, 876-903.

Pollmeier, J., Hardy, I., Koerber, S. & Möller, K. (2011). Lassen sich naturwissenschaftliche Lernstände im Grundschulalter mit schriftlichen Aufgaben valide erfassen? *Zeitschrift für Pädagogik*, 57, 834-853.

Smith, C. (2007). Bootstrapping Processes in the Development of Students' Commonsense Matter Theories: Using Analogical Mappings, Thought Experiments, and Learning to Measure to Promote Conceptual Change. *Cognition and Instruction*, 25(4), 337-398.

Smith, C., Carey, S. & Wisner, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, 21, 177-237.



## **Biologische Informationen aus Texten und Bildern erschließen – eine Untersuchung zu Fähigkeiten von Schülerinnen und Schülern**

### Theoretischer Hintergrund

Ausgangspunkt der Studie sind Besonderheiten naturwissenschaftlicher Kommunikation, die am Beispiel der Biologie spezifiziert werden. Naturwissenschaftliche Kommunikation ist von unterschiedlichen Repräsentationen beispielsweise Texten, Tabellen, Diagrammen, Fotos und Zeichnungen geprägt (Anderson, 1999). Bei der Erklärung eines Sachverhaltes treten sie oft kombiniert als Multiple External Representations (MERs) auf (Ainsworth, 1999). In Anlehnung an Befunde und Theorien zum Text- und Bildverstehen (Schnotz, 2005) wird davon ausgegangen, dass Informationen aus Texten im verbalen und Informationen aus Bildern im piktoralen Kanal verarbeitet und zu einem kohärenten mentalen Modell integriert werden. Auf dieser Basis werden bezüglich der Fähigkeiten der Lernenden die folgenden Teilkonstrukte unterschieden: biologische Informationen erschließen aus Texten (T), Bildern (B) und Bild-Text-Kombinationen (MER).

In der vorliegenden Studie werden naturwissenschaftstypische Charakteristika in den Bild-Text-Kombinationen umgesetzt, aus denen sich die Lernenden biologische Informationen erschließen müssen. Hierzu gehört, dass sich diese aus nicht redundanten Repräsentationen zusammensetzen. Beispielsweise sind in einem Text über die Fotosynthese (beteiligte Substanzen, ablaufende Prozesse, Wege der unterschiedlichen Stoffe) und einer Zeichnung zum Aufbau des Blattes (Blattschichten, Zellorganellen, die an der Fotosynthese beteiligt sind) unterschiedliche biologische Informationen zu einem übergreifenden Konzept (Fotosynthese) enthalten. Zusätzlich geht die Funktion der Bilder über eine Veranschaulichung von Informationen aus dem Text hinaus und in den Texten sind nicht ausschließlich erläuternde Informationen zum Verständnis des Bildes, beispielsweise zur Bedeutung bestimmter Farben enthalten. In Abgrenzung zu anderen Arbeiten (Schnotz et al., 2010) kann somit jede der Repräsentationen für sich stehen und trotzdem einen biologischen Sachverhalt vermitteln. Nach Mayer (2005) ist bei Bild-Text-Kombinationen ein Multimedia Effekt zu erwarten. Diesen zeigen auch die Ergebnisse von Hartmann (2013) für naturwissenschaftliche Aufgaben in denen ein Teil der allgemeinen Information aus dem Text (z. B. ein Versuchsaufbau) in einem Bild veranschaulicht wird. In der vorliegenden Arbeit wird davon ausgegangen, dass aufgrund der Besonderheiten der MERs kein Multimedia Effekt auftritt. Vielmehr wird prognostiziert, dass die Verarbeitung der untersuchten Bild-Text-Kombinationen für Lernende eine schwierige Aufgabe darstellt.

### Fragestellungen

1. Inwiefern lassen sich die Fähigkeiten von Schülerinnen und Schülern im Bereich biologische Informationen erschließen mithilfe eines schriftlichen, biologiebezogenen Leistungstests empirisch abbilden?
2. Zeigen sich zwischen den Teilkonstrukten Unterschiede in der Itemschwierigkeit und in der Leistung?
3. Welche qualitativen Unterschiede und systematischen Zusammenhänge zwischen den Fähigkeiten und Faktoren wie Fachwissen, Schultyp und Geschlecht lassen sich innerhalb der Schülerpopulation beschreiben?

### Methodik

Zur Datenerhebung wurden 968 Probanden (Jahrgang 9/10) mithilfe eines quantitativen Messinstruments (Fähigkeitstest) in einer Querschnittstudie getestet. Die Untersuchungsteilnehmer bearbeiteten 57 Items (Multimatrixdesign). Die Testaufgaben bestehen aus einem Aufgabenstamm (Text/Bild) und mehreren Items, in denen die Probanden aufgefordert werden biologische Informationen aus einem Text oder Bild oder aus einer Bild-Text-Kombination zu erschließen. Somit können die Items jeweils einem der drei Teilkonstrukte zugeordnet werden. Jedes Items wurde im Mittel von 155.9 Schülerinnen und Schülern beantwortet ( $SD=23.8$ ,  $Min=86$ ,  $Max=187$ ). Ein Proband bearbeitete durchschnittlich 9.7 Items. Die Auswertung erfolgte mit Hilfe von klassischer (SPSS) und probabilistischer Testtheorie (Winsteps, Conquest).

### Ergebnisse

Der fachspezifische Test weist moderate bis gute Kennwerte auf ( $RelEAP/PV=.54$ ,  $0.9 < MNSQ < 1.1$ ). Weitere Befunde zur empirischen Trennbarkeit der Teilkonstrukte werden auf der Tagung vorgestellt und diskutiert. In Bezug auf die zweite Forschungsfrage zeigen die Ergebnisse der einfaktoriellen ANOVA signifikante Schwierigkeitsunterschiede zwischen Teilkonstrukten ( $F(2,54)=6,792$ ,  $p < .05$ ,  $r=0.45$ ). Paarweise Vergleiche zeigen signifikante Unterschiede zwischen der Schwierigkeit von Items zu den Teilkonstrukten T und MER ( $p < .05$ ) und den Teilkonstrukten B und MER ( $p < .05$ ). Generell deuten die Ergebnisse wie prognostiziert darauf hin, dass das Erschließen von biologischen Informationen aus Bild-Text-Kombinationen im Falle der vorliegenden Arbeit eine höhere Schwierigkeit darstellt als das Erschließen von biologischen Informationen aus einzelnen Repräsentationen. Somit scheint von den beschriebenen biologiespezifischen Bild-Text-Kombinationen kein verstehensförderlicher (Multimedia-) Effekt auszugehen. Vielmehr kann eine besondere Schwierigkeit im Umgang mit naturwissenschaftlichen Repräsentationen am Beispiel der Biologie beschrieben werden.

Anderson, C. W. (1999). Preface: Inscriptions and science learning. *Journal of Research in Science Teaching*, 36(9), 973-974.

Ainsworth, S. (1999). The Functions of Multiple Representations. *Computers & Education*, 33(2), 131-152.

Hartmann, S. (2013). Die Rolle von Leseverständnis und Lesegeschwindigkeit beim Zustandekommen der Leistungen in schriftlichen Tests zur Erfassung naturwissenschaftlicher Kompetenz. Universität Duisburg-Essen, Duisburg-Essen.

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Schnotz, W., McElvany, N., Horz, H., Schroeder, S., Ullrich, M., Baumert, J., Hachfeld, A., & Richter, T. (2010). Das BITE-Projekt: Integrative Verarbeitung von Bildern und Texten in der Sekundarstufe I. *Zeitschrift für Pädagogik*, 56. Beiheft.



## **Förderung systemischen Denkens in biologischen Kontexten: Können Wirkungen einer Lehrerfortbildung auf der Ebene der SchülerInnen nachgewiesen werden?**

Systemisches Denken gilt als Schlüssel zur Bearbeitung komplexer naturwissenschaftlicher, ökonomischer und sozio-kultureller Fragestellungen im Umfeld einer Bildung für eine nachhaltige Entwicklung (BNE). In bisherigen Studien ging es vor allem um die direkte Förderung von SchülerInnen in systemischem Denken. Ziel der vorliegenden Studie war es, im Rahmen des Projekts SysThema (System Thinking in ecological and multidimensional areas) eine Lehrerfortbildung zur Förderung systemischen Denkens im Biologieunterricht durchzuführen und deren Wirkungen auf der Ebene der SchülerInnen zu untersuchen. Die Ergebnisse zeigen, dass die Förderung systemischen Denkens bei SchülerInnen durch eine LehrerInnenfortbildung möglich ist.

### Theoretischer Hintergrund und Fragestellung

Eine systemtheoretische Betrachtungsweise gilt als Schlüssel zur nachhaltigen Lösung komplexer Problembereiche (z.B. Klimawandel). In der Folge wurden verschiedene Ansätze zur Förderung systemischen Denkens bei SchülerInnen durchgeführt, die zeigen konnten, dass sich systemisches Denken in verschiedenen Klassenstufen tatsächlich fördern lässt (vgl. Ossimitz, 2000; Assaraf et al., 2010; Rieß & Mischo, 2008; Bräutigam et al., 2011). Gleichzeitig spielt die Förderung systemischen Denkens in der Lehrerausbildung und -fortbildung bisher eine eher untergeordnete Rolle. Da erfolgreiches Lernen von SchülerInnen in starkem Maße von der Lehrerexpertise abhängt (vgl. Hattie, 2009), war es Ziel dieser Studie, ein erfolgreiches Lehrerfortbildungsprogramm zu entwickeln und zu evaluieren. Lipowsky und Rzejak (2012) haben wesentliche Merkmale einer erfolgreichen Fortbildung benannt: sie darf nicht zu kurz sein, sollte zu einer Vertiefung des fachdidaktischen Wissens (vgl. Shulman, 1986) und zu einer Verbesserung der diagnostischen Fähigkeiten führen. Insbesondere wird die „Verschränkung von Input-, Erprobungs- und Reflexionsphasen“ von Lipowsky & Rzejak (2012, S.7) als Schlüssel erachtet, um über die Fortbildung dauerhafte Ergebnisse auf SchülerInnenebene erzielen zu können. Im Sinne von Gräsel et al. (2004) sollen die Lehrkräfte nicht durch die Bereitstellung fertiger Unterrichtskonzepte sondern durch Ideen und Bausteine zur Eigenaktivität angeregt werden.

### Methode

Von Februar bis Juni 2013 wurde eine quasiexperimentelle Interventionsstudie im Prä-Posttest-Design mit Kontrollgruppe durchgeführt. Dazu nahmen Realschullehrkräfte, die im Schuljahr 2012/13 eine neunte Klasse in Biologie unterrichteten, an einer Fortbildung zum systemischen Denken im Umfang von 4 Tagen teil. Im Anschluss hielten sie jeweils eine achtstündige Unterrichtseinheit zu systemischem Denken. Die Wirksamkeit der Intervention wurde quantitativ mittels Fragebögen direkt bei den SchülerInnen ermittelt. Es wurden zwei Experimentalgruppen und eine Kontrollgruppe realisiert. Die Experimentalgruppen bestanden aus 4 bzw. 3 Lehrkräften, die Kontrollgruppe aus 3 Lehrkräften. Die Experimentalgruppen variierten in den Anteilen an fachwissenschaftlichem Wissen (CK) zu den Aspekten biologischer Systeme und an fachdidaktischem Wissen (PCK) über die Vermittlung systemischen Denkens. Die Kontrollgruppe erhielt keinen Input. Die Gesamtanzahl der getesteten SchülerInnen belief sich auf N = 289. Erhebung und Intervention waren an einem heuristischen Kompetenzmodell orientiert, das sich aus vier Kompetenzdimensionen zusammensetzt: (1) Deklaratives/Konzeptionelles systemisches Wissen, (2) Systemmodellierungsfähigkeit, (3) systemische Modellierfähigkeit und (4) Fähigkeit zur Bewertung von Systemmodellen. Das Erhebungsinstrument stellt dabei eine Weiterentwicklung des in Klassenstufe 6 erfolgreichen Projekts SYSDENA von Bräutigam et al. (2011) dar. Aus Gründen der Umsetzbarkeit für den durch die Lehrkräfte zu planenden Unterricht, wurden nur die ersten drei Kompetenzdimensionen in die Studie miteinbezogen.

### Ergebnisse

Die Pilotierung des Fragebogens zeigte, dass das Messinstrument geeignet ist, um systemisches Denken bei Schülerinnen und Schülern zu erheben. Die Resultate der zwei untersuchten Experimentalgruppen weisen zwischen Prä- und Posttest einen Leistungszuwachs von hoher Signifikanz auf, wobei die gemessenen Effekte zwischen den beiden Experimentalgruppen (Exp.1:  $F(1,175) = 8.233$ ,  $p = .000$ ,  $\eta^2 = .045$ ; Exp.2:  $F(1,216) = 9.739$ ,  $p = .002$ ,  $\eta^2 = .043$ ) und der Kontrollgruppe schwach sind. Beim Vergleich der Zuwächse zwischen den beiden Experimentalgruppen nach der Intervention zeigt sich kein signifikanter Unterschied. Des Weiteren weist die Kontrollgruppe einen hoch signifikanten Zuwachs vom Prä- zu Posttest auf. Bei der Aufspaltung in die ersten drei Dimensionen des heuristischen Kompetenzmodells ergeben sich keine neuen Erkenntnisse, da kaum Unterschiede zum Gesamtergebnis festzustellen sind.

Assaraf, O., Dodick, J. & Tripto, J. (2010). High School Students' Understanding of the Human Body System. *Research in Science Education*, 43(1), 33-56.

Bräutigam, J., Nerb, J. & Rieß, W. (2011). Entwicklung und Validierung eines Fragebogens zur Evaluierung einer Unterrichtseinheit zur Förderung systemischen Denkens. 75. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), 2011, Otto-Friedrich-Universität Bamberg, 67-68.

Hattie, H. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London.

Gräsel, C., Parchmann, I., Puhl, T., Baer, A., Fey, A. & Demuth, R. (2004). Lehrerfortbildung und ihre Wirkungen auf die Zusammenarbeit von Lehrkräften und die Unterrichtsqualität. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule*. Münster: Waxmann, 136-137.

Lipowsky, F. & Rzejak, D. (2012). Lehrerinnen und Lehrer als Lerner – Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. *Schulpädagogik heute*, 5, 3-9.

Ossimitz, G. (2000). *Entwicklung systemischen Denkens: Theoretische Konzepte und empirische Untersuchungen*. München: Profil-Verlag, 121-178.

Rieß, W. & Mischo, C. (2008). Förderung systemischen Denkens im Bereich von Ökologie und Nachhaltigkeit. *Unterrichtswissenschaft*, 36, 346 – 364.

Shulman, Lee S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher* 15 (2), 4-14.



## **Beziehungen zwischen argumentativen Fähigkeiten in den naturwissenschaftlichen Fächern und im Fach Deutsch**

### Theoretischer Hintergrund

Ein zentrales Ziel bei der Erlangung naturwissenschaftlicher Grundbildung ist es, Lernende zum kritischen Denken anzuleiten. In den Bildungsstandards der KMK für die naturwissenschaftlichen Fächer wird dieses Ziel im Kompetenzbereich „Bewertung“ konzeptualisiert. Die Fähigkeit, Argumente schlüssig zu formulieren und nachzuvollziehen, wird dabei als zentrales Mittel identifiziert. Der Fokus liegt hier auf fachlich korrekten sowie schlüssigen Argumenten. Im Fach Deutsch steht beim Argumentieren insbesondere die Vermittlung sprachlicher und struktureller Elemente im Vordergrund.

Beispielsweise konnten Osborne, Erduran und Simon (2004) beobachten, dass sich ein explizites Training von Argumentationsfähigkeit - d.h. zu Strukturen und Prinzipien eines Arguments - positiv auf die Qualität naturwissenschaftlicher Argumentationen auswirkte. Hingegen konnte dieser Effekt in anderen Studien nicht festgestellt werden (z.B. von Aufschnaiter, Erduran, Osborne, & Simon, 2008).

In der Literatur wird weiterhin die Fähigkeit zum schlussfolgernden Denken diskutiert, die die Struktur und Qualität eines (naturwissenschaftlichen) Entscheidungsprozesses beeinflussen könnte: Einige Studien zeigten einen engen Zusammenhang zum Argumentieren auf (z.B. Means & Voss, 1996); andere konnten diese Beziehung für spezifischere Faktoren wie dem verbalen schlussfolgernden Denken nicht replizieren (z.B. Eggert & Bögeholz, 2010).

Zusammengefasst sind die Befunde zum Einfluss einer allgemeinen Argumentationsfähigkeit wie im Fach Deutsch und schlussfolgerndem Denken auf naturwissenschaftliches Argumentieren inkonsistent und auf eher geringer Datengrundlage beruhend. Weitere empirische Untersuchungen mit hinreichend großen Datensätzen sind notwendig, um zu belastbaren Aussagen zur Beziehung zwischen den Konstrukten zu gelangen. Dies ist insbesondere zur Entwicklung von unterrichtspraktischen Implikationen für die Förderung naturwissenschaftlicher Argumentationsfähigkeit relevant. Lehrerinnen und Lehrer sollten informiert sein, inwiefern diese Fähigkeit ein Denken über die Fächergrenzen hinaus erfordert.

### Fragestellung

Diese Studie untersucht die Beziehungen zwischen den beiden Formen des Argumentierens sowie schlussfolgerndem Denken. Dafür wird die Fähigkeit zum allgemeinen, fächerübergreifenden Argumentieren mit Hilfe von Aufgaben für das Fach Deutsch erhoben, da diese kein fachspezifisches Wissen erfordern. Die zentrale Forschungsfrage ist:  
- Wie hängen die Konstrukte Argumentieren in den naturwissenschaftlichen Fächern, Argumentationsfähigkeit im Fach Deutsch und schlussfolgerndes Denken zusammen?

### Methode

Die Daten wurden im Rahmen einer umfangreichen empirischen Schulleistungsstudie im Jahr 2011 erhoben, an der 3274 Schülerinnen und Schüler der 10. Jahrgangsstufe (Alter  $M = 16.1$ ,  $SD = 0.7$  Jahre, 48.5 % weiblich) teilnahmen. In die Analysen gingen 209 Items mit offenem Antwortformat zur Erhebung der Argumentationsfähigkeit in den Naturwissenschaften und zwei komplexe Schreibaufgaben zur Messung der Argumentationskompetenz für das Fach Deutsch mit ein. Verbales, numerisches und figurales schlussfolgerndes Denken wurden mit jeweils 16 Multiple-Choice Items erhoben.

Mit Hilfe der Software Mplus 7.1 wurden zwei Strukturgleichungsmodelle geschätzt. Modell 1 war ein mehrdimensionales Strukturmodell, das je eine latente Dimension für die beiden Formen des Argumentierens und die drei Inhaltsbereiche des schlussfolgernden Denkens spezifizierte, um die korrelativen Zusammenhänge auf Konstruktebene zu bestimmen. In Modell 2 wurde eine multiple Regression mit naturwissenschaftlicher Argumentationsfähigkeit als Kriterium durchgeführt.

### Ergebnisse

Die Analysen zeigen eine substantielle Korrelation der Argumentationsfähigkeit in den Naturwissenschaften mit der im Fach Deutsch ( $\rho = .51$ ). Ähnlich hohe Zusammenhänge finden sich zwischen naturwissenschaftlichem Argumentieren und numerischem ( $\rho = .50$ ), figuralem ( $\rho = .56$ ) und verbalem schlussfolgerndem Denken ( $\rho = .63$ ), wobei die letztgenannte Korrelation erwartungsgemäß am höchsten ausfiel. Insgesamt konnte 50% der Varianz im naturwissenschaftlichen Argumentieren mit diesem Modell aufgeklärt werden, was auf eine moderate Beziehung sowie Eigenständigkeit der Konstrukte hinweist.

Im Gegensatz dazu sind die Korrelationen der Argumentationsfähigkeit im Fach Deutsch mit numerischem ( $\rho = .11$ ), figuralem ( $\rho = .29$ ) und verbalem ( $\rho = .41$ ) schlussfolgerndem Denken geringer ausgeprägt. Dies deutet darauf hin, dass bei beiden Formen des Argumentierens unterschiedliche kognitive Prozesse erforderlich sind. Weitere Unterschiede zeigen sich im Schreibprodukt der Schülerinnen und Schüler: Lernende tendieren in den Naturwissenschaften dazu, ihre Argumentationen lediglich durch die Aufzählung von Belegen zu gestalten anstatt diese in eine elaborierte Begründungshierarchie einzubinden.

Eggert, S., & Bögeholz, S. (2010). Students' use of decision-making strategies with regard to socioscientific issues: An application of the Rasch partial credit model. *Science Education*, 94(2), 230–253. doi:10.1002/sce.20358

Means, M. L., & Voss, J. F. (1996). Who Reasons Well? Two Studies of Informal Reasoning Among Children of Different Grade, Ability, and Knowledge Levels. *Cognition and Instruction*, 14(2), 139–178. doi:10.1207/s1532690xci1402\_1

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994–1020. doi:10.1002/tea.20035

von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45(1), 101–131. doi:10.1002/tea.20213



Chair(s): Monika Oberle (Georg-August-Universität Göttingen)

Diskutant/in: Rainer Lehmann (Humboldt-Universität zu Berlin)

Was „guten Unterricht“ ausmacht, lässt sich nicht allgemeingültig bestimmen. Viele Faktoren wirken auf den Erfolg oder den Misserfolg von Lehr-Lern-Prozessen ein, die je spezifische Anforderungen an Unterrichtsgestaltung stellen (vgl. Angebots-Nutzungs-Modell, Helmke, 2012: 71). Es wurden allerdings einige fachübergreifende Qualitätsbereiche identifiziert, die für erfolgreichen Unterricht zentral scheinen (Helmke, 2012: 168 ff.; Kunter & Trautwein, 2013). Um sie für eine Analyse der Kompetenzvermittlung im Fachunterricht nutzbar zu machen, müssen diese Unterrichtsmerkmale je fachspezifisch gefüllt werden. Zur Realität des Politikunterrichts und seinen spezifischen Qualitätskriterien liegen bislang kaum systematisch gewonnene Erkenntnisse vor. Das Symposium möchte sich diesem Forschungsbedarf stellen.

Das Symposium nähert sich der Qualität von Politikunterricht aus unterschiedlichen Perspektiven. Erstens werden Schülerwahrnehmungen zu Qualitätsmerkmalen ihres Politikunterrichts in den Blick genommen und Zusammenhänge mit Ausprägungen einzelner Aspekte ihrer Politikkompetenz (Detjen et al., 2012) analysiert (Weißeno & Landwehr; Oberle & Forstmann); zweitens werden Vorstellungen von Unterrichtsqualität und Aussagen zu ihrer eigenen Unterrichtspraxis bei Lehrer/innen untersucht (Biedermann & Oser); drittens werden mittels Videostudien Prozesse des Politikunterrichts direkt in den Blick genommen und auf fachspezifische Qualitätsmerkmale hin analysiert (Manzel). Im Fokus stehen in allen vier empirischen Studien Unterrichtsmerkmale, die für den Kompetenzerwerb von Schülerinnen und Schülern als bedeutsam erachtet werden, so die kognitive Aktivierung (Weißeno & Landwehr, Manzel), Konstruktionsorientierung (Biedermann & Oser, Manzel, Weißeno & Landwehr), Unterstützungsverhalten (Weißeno & Landwehr), Unterrichtsklima (Biedermann & Oser; Oberle & Forstmann; Weißeno & Landwehr), Angebotsvariation (Manzel; Oberle & Forstmann) und Lebensweltorientierung (Oberle & Forstmann).

Die Beiträge untersuchen Aspekte der Qualität von Politikunterricht unter Verknüpfung einer fachdidaktischen und pädagogisch-psychologischen Perspektive. Die Referent/innen haben einen erziehungswissenschaftlichen (Biedermann, Oser) bzw. politikdidaktischen (Manzel, Oberle, Weißeno) Hintergrund. Im Rahmen des Symposiums sollen die Ergebnisse der empirischen Studien diskutiert werden und der Frage nachgegangen werden, ob und wie allgemeine Merkmale von Unterrichtsqualität auf den politischen Fachunterricht übertragen werden können. Dabei sollen auch die unterschiedlichen empirischen Zugangsweisen kritisch und konstruktiv in den Blick genommen werden. Diskutant ist Prof. Dr. Rainer Lehmann, der u.a. an der Durchführung der Civic Education Studie (CivED) der International Association for Educational Achievement (IEA) 2001 für Deutschland maßgeblich beteiligt war.

Detjen, J., Massing, P., Richter, D., & Weißeno, G. (2012). Politikkompetenz – Ein Modell. Wiesbaden: Springer VS.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. 4. Aufl., Seelze-Velber: Kallmeyer.

Kunter, M. & Trautwein, U. (2013). Psychologie des Unterrichts. Paderborn: Schöningh.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
PEG-Gebäude  
Raum 1.G150



## ***Effekte qualitätvollen Politikunterrichts, des Selbstkonzepts und der sozialen Interaktion auf das Fachwissen***

Der Politikunterricht muss nach dem Angebots-Nutzungs-Modell (Helmke, 2012) verschiedenen Qualitätsdimensionen gerecht werden. Qualitätvoller Unterricht unterstützt den Konzeptaufbau durch das Auslösen mentaler Aktivitäten. Die kognitive Aktivierung ist die bedeutsamste leistungsrelevante Qualitätsdimension (Kunter & Trautwein, 2013). Ein weiteres Qualitätsmerkmal der Tiefenstruktur des Unterrichts ist die konstruktive Unterstützung (Ebenda, S. 97 ff). Eine hohe konstruktive Unterstützung geht mit hohem fachlichen Interesse der Schüler/-innen einher (Klieme & Rakoczy, 2003), da die Lehrkraft sich auf die Bedürfnisse der Schüler/-innen einstellt und bedarfsbezogene Hilfestellungen leistet.

Ein Indikator für die Motivierung ist das Selbstkonzept der eigenen schulischen Leistungsfähigkeit. Ein besseres oder ungünstiges Selbstkonzept basiert auf sozialen Vergleichen (Dijkstra et al., 2008). Es gibt einen Interaktionseffekt zwischen den Interessen der Lernenden und aktivierendem Unterricht hinsichtlich der Leistungsentwicklung (Lipowsky et al., 2009). Erste Hinweise für die Bedeutung des Fachinteresses und des fachspezifischen Selbstkonzepts für die Lernleistung im Politikunterricht liegen vor (Weißenö & Eck, 2013).

Die Beziehungskultur zwischen den Lehrer/-innen und den Schüler/-innen schafft ein lernförderliches Klima. Im Kern geht es dabei um die soziale Interaktion und emotionale Sicherheit. Positive Beziehungen haben positive Effekte auf die Motivation und die Leistung. Eine Arbeit von Buhl (2003) fand Zusammenhänge zwischen dem Schüler-Lehrer-Verhältnis und explizit politischen Merkmalen auf Schülerseite.

Dem Modell des Fachwissens (Detjen et al., 2012) liegen 30 Fachkonzepte und dazu gehörende konstituierende Begriffe zugrunde. Nach dem Informationsparadigma und den soziokonstruktivistischen Lerntheorien sind Fachkonzepte keine isolierten Wissenseinheiten im Gedächtnis, sondern sie liegen in der Form verschiedener grundlegender Schemata für Verstehen und Schlussfolgern (Urteilen) vor. Je stärker Lernende sich mental aktiv mit dem Lerngegenstand auseinandersetzen, umso besser werden die Konzepte verstanden und umso nachhaltiger wird das Lernen (Kunter & Trautwein, 2013, S. 86). Darin besteht der Nutzen des Unterrichts für die Schüler/-innen nach dem Angebots-Nutzungs-Modell (Helmke, 2012).

Politikstunden scheinen nur sehr wenig zum Aufbau der Vorstellungen zu Fachkonzepten beizutragen (Manzel & Gronostay, 2013). Das Komplexitätsniveau ist sehr gering und das Faktenniveau liegt bei 67 % pro Unterrichtsstunde. Eine auf der deutschen Stichprobe der Civic Education-Studie basierende Untersuchung von Watermann (2003) belegt auf Individualebene signifikant günstige, wenn auch nur schwach ausgeprägte Effekte eines diskursiven Unterrichtsstils auf das politische Wissen. Die hohen Anforderungen führen offenbar dazu, dass nur Gymnasiasten von der Aktivierung profitieren (Gruehn, 2000).

Die Studie geht folgenden Fragestellungen nach: Wirkt sich effektiver Unterricht auf das Wissen aus? Hat die Unterrichtswahrnehmung der Schüler/-innen einen Effekt auf das Lernergebnis? Beeinflussen Hintergrundvariablen wie Migrationshintergrund, Geschlecht oder Note in Politik die Schülerwahrnehmung und die Leistung? Erhoben werden die Einschätzungen der Schüler/-innen und das Schülervwissen. Auch ohne Videos ergibt die Schülerperzeption von Unterricht ein verlässliches Bild der Unterrichtsqualität (Gruehn, 2000). Das Sample besteht aus 1260 Schüler/-innen der 9. & 10. Klassen in Realschulen & Gymnasien in Baden-Württemberg. Zur Erfassung werden ausschließlich geschlossene Items aus ICCS (2009) und IGLU (2006) eingesetzt. Die Erhebung erfolgt auf Basis einer vierstufigen Likertskala. Die 35 Wissensfragen wurden aus POWIS, TEESAEC und WEUS entnommen. DIF-Tests wurden für Geschlecht, Migrationshintergrund und Schulart gerechnet. Die Fits aller Messmodelle sind gut.

Das Pfadmodell gibt die Zusammenhänge zwischen konstruktiver Unterstützung, kognitiver Aktivierung, Leistungsfähigkeit, sozialer Interaktion, dem Fachwissen und den Hintergrundvariablen wieder (akzeptable Fitwerte). Berichtete Unterstützung und Aktivierung hängen stark positiv zusammen. Die Aktivierung weist einen Effekt auf das Wissen auf. Erwartungskonform berichten Realschüler/-innen weniger Aktivierung und Gymnasiast/-innen wissen mehr. Die Lehrer-Schüler-Interaktion hat einen Effekt auf das leistungsbezogene Selbstkonzept, das mit dem Wissen positiv zusammen hängt. Die Note wirkt sich leicht auf alle Batterien aus, Geschlecht nur auf die Interaktion. Der Migrationshintergrund zeigt die erwarteten Effekte.

Buhl, M. (2003): Schulqualität und politische Sozialisation. Zeitschrift für Soziologie der Erziehung und Sozialisation, 4, 385-398.

Detjen, J., Massing, P., Richter, D. & Weißenö, G. (2012). Politikkompetenz - ein Modell. Wiesbaden: Springer.

Dijkstra, P. et al. (2008). Social comparison in the classroom: A review. Review of Educational Research, 78, 828-879.

Gruehn, S. (2000). Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster: Waxmann.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (4. Aufl.). Seelze: Klett-Kallmeyer.

Klieme, E. & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive. In Baumert et al. (Hrsg.), PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland (S. 333-359). Opladen: Leske u. Budrich.

Kunter, M. & Trautwein, U. (2013). Psychologie des Unterrichts. Paderborn: Schöningh.

Lipowsky, F. (2009). Unterricht. In E. Wild & J. Möller (Hrsg.), Pädagogische Psychologie (S. 73-102). Berlin: Springer.

Manzel, S. & Gronostay, D. (2013). Videografie im Politikunterricht – Erste Ergebnisse einer Pilotstudie zu domänenspezifischen Basisdimensionen. In U. Riegel & K. Macha (Hrsg.), Videobasierte Kompetenzforschung in den Fachdidaktiken (S. 198-215). Münster: Waxmann.

Watermann, R. (2003): Diskursive Unterrichtsgestaltung und multiple Zielerreichung im politisch bildenden Unterricht. Zeitschrift für Soziologie der Erziehung und Sozialisation, 4, 356-370.

Weißenö, G. & Eck, V. (2013). Wissen, Selbstkonzept und Fachinteresse. Ergebnisse einer Interventionsstudie zur Politikkompetenz. Münster/New York: Waxmann.



## **Qualität von Politikunterricht aus Schülersicht: Alltagsorientierung, Angebotsvariation und diskursive Unterrichtsgestaltung im Zusammenhang mit motivationalen Orientierungen und Geschlecht**

Für die Qualität von Unterricht wurden allgemeine Merkmale identifiziert, darunter Motivation, Angebotsvariation und lernförderliches Klima (Helmke, 2012: 121 ff.). Welche spezifischen Ausprägungen hierbei für einen qualitätsvollen Politikunterricht wünschenswert sind, wurde bislang nur ansatzweise systematisch erforscht. So liegen zum demokratischen Unterrichtsklima zahlreiche Studien vor, die dessen positive Effekte u.a. auf politisches Wissen und politikbezogene Motivationen belegen (z.B. Oser & Biedermann, 2003; Watermann 2003). Allerdings wurden hier selten weitere Aspekte des Fachunterrichts berücksichtigt, sodass unklar bleibt, ob das Unterrichtsklima selbst oder mit diesem einhergehende Merkmale der Unterrichtsgestaltung wirksam werden. Als weitere wirksame Aspekte von Unterrichtsqualität kommen u.a. die Lebensweltorientierung und die Angebotsvariation (Methoden- und Medieneinsatz) in Frage. Im theoretischen Diskurs der Politikdidaktik hat Lebensweltorientierung einen hohen Stellenwert und wird meist dem Unterrichtsprinzip Schülerorientierung (Schelle, 2005) zugeordnet (bei Helmke: Motivierung). Es wird vermutet, dass besonders weibliche Lernende Bedarf nach lebensweltorientierten Unterrichtsangeboten haben (Richter, 2005; Oberle, 2013). Empirisch gesicherte Ergebnisse fehlen allerdings sowohl bezüglich der geschlechtsspezifischen Bedürfnisse von Lernenden im Fach Politik als auch bezüglich der Praxis des alltäglichen Politikunterrichts.

Die vorliegende Studie untersucht Erwartungen an und Realitäten des Politikunterrichts aus Schülersicht und fokussiert dabei die Aspekte Lebensweltorientierung, Methoden- und Medieneinsatz sowie diskursive Unterrichtsgestaltung. Ziel ist es, zum einen den Alltag des Politikunterrichts in Niedersachsen zu beleuchten, zum anderen dessen Qualität mit Erwartungen der Lernenden abzugleichen. Darüber hinaus sollen die Zusammenhänge der erhobenen fachspezifischen Qualitätsmerkmale mit dem Geschlecht der Lernenden und deren motivationalen Politikkompetenzen (Sachinteresse, Fachinteresse, Selbstwirksamkeitsüberzeugungen; vgl. Modell der Politikkompetenz von Detjen et al. 2012) untersucht werden. Die Fragestellungen lauten, wie lebensweltorientiert, angebotsvielfältig und diskursiv gestaltet „alltäglicher“ Politikunterricht heute ist, in welchem Zusammenhang diese Merkmale der Unterrichtsqualität mit den motivationalen Facette von Politikkompetenz der Schüler/innen sowie deren Erwartungen an ihren politischen Fachunterricht stehen. Dabei interessieren besonders geschlechtsspezifische Unterschiede der Ausprägungen und Zusammenhänge.

Die Untersuchung setzt einen standardisierten Fragebogen ein, dem eine Pilotstudie mit qualitativen Interviews (N = 15) sowie eine quantitative Vorerhebung mittels Fragebogen (N = 200) vorgeschaltet waren. Das Sample der Hauptstudie besteht aus ca. N = 800 (Datenbereinigung läuft) Schüler/innen der 10. Klasse an Niedersächsischen Gymnasien und Gesamtschulen. Erhoben wurden die Erwartungen an und Wahrnehmungen von Politikunterricht der Lernenden hinsichtlich Lebensweltorientierung, Methoden- und Medieneinsatz sowie diskursiver Unterrichtsgestaltung. Die Merkmale der Unterrichtsqualität werden mit motivationalen Facetten der Politikkompetenz wie Sachinteresse, Fachinteresse und Selbstwirksamkeitsüberzeugungen der Schüler/innen in Beziehung gesetzt. Aspekte der Unterrichtsqualität und Schülermotivationen werden latent modelliert und in latenten Strukturgleichungen (MPlus) miteinander in Beziehung gesetzt. Dabei kommen u.a. adaptierte Erhebungsinstrumente von Köller, Schnabel und Baumert (2000; Sach- und Fachinteresse), Vetter (1997, Selbstwirksamkeitsüberzeugung bzw. efficacy) sowie der Civic Education Studie (CIVEd) der International Association for the Evaluation of Educational Achievement (IEA) (diskursive Unterrichtsgestaltung) zum Einsatz. Die Batterie zur Alltagsorientierung wurde unter Einbeziehung von Studien anderer Fächer in den qualitativen und quantitativen Vorstudien entwickelt und wird hier validiert. Die Datenauswertung läuft aktuell, Ergebnissen liegen bis Ende 2013 vor. Die Auswertung erfolgt mit SPSS und MPlus, die Ergebnisse werden u.a. mittels latenter Strukturgleichungen und Pfadmodelle präsentiert werden.

Detjen, J., Massing, P., Richter, D. & Weißeno, G. (2012). Politikkompetenz - ein Modell. Wiesbaden: Springer.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (4. Aufl.). Seelze: Klett-Kallmeyer.

Köller, Olaf/Schnabel, Kai/Baumert, Jürgen (2000): Der Einfluss der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 32(2), S. 70-80.

Oberle, M. (2013). Geschlechtsspezifische Differenzen in politischen Kompetenzen. In S. Frech & D. Richter (Hg.), Politische Kompetenzen fördern (S. 164-184). Schwalbach/Ts.: Wochenschau Verlag.

Oser, F., & Biedermann, H. (2003). Jugend ohne Politik. Zürich: Rüegger.

Schelle, C. (2005): Adressatenorientierung. In: Sander, W. (Hrsg.): Handbuch der politischen Bildung (S. 79-92). Bonn: Bundeszentrale für politische Bildung.

Richter, D. (2005): Geschlechtsspezifische Aspekte politischen Lernens. In: Sander, W. (Hrsg.): Handbuch der politischen Bildung (S. 407-418). Bonn: Bundeszentrale für politische Bildung.

Vetter, A. (1997). Political Efficacy. Reliabilität und Validität, alte und neue Messmodelle im Vergleich. Wiesbaden: Deutscher Universitätsverlags-Verlag.

Watermann, R. (2003): Diskursive Unterrichtsgestaltung und multiple Zielerreichung im politisch bildenden Unterricht. Zeitschrift für Soziologie der Erziehung und Sozialisation, 4, 356-370.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
PEG-Gebäude  
Raum 1.G150



## ***Auf die Lehrpersonen kommt es an? – Politische Bildung in der Schweiz im Vergleich mit 37 anderen Ländern (ICCS 2009)***

### Theoretischer Hintergrund und Fragestellungen

„Die Legitimität politischer Entscheidungen geht in der modernen Kultur auf die Idee eines politischen Herrschafts-Vertrags zurück, dem alle Staatsbürger zustimmen können, weil ihre grundlegenden Rechte und Interessen durch ihn gesichert und gewahrt bleiben“ (Meyer, 2006, S. 164). Die IEA (International Association for the Evaluation of Educational Achievement) ging mit ihrer Vergleichsstudie ICCS 2009 (International Civic and Citizenship Education Study) der Frage nach, wie ausgeprägt sich die politische Bildung in 38 Ländern zeigt und damit verbunden, welcher Stellenwert der Erreichung des von Meyer erwähnten Ideals eingeräumt wird – so auch in der Schweiz. Während dazu die Ergebnisse der Schüler/innen bereits breit dargestellt und vertiefend diskutiert worden sind (vgl. Biedermann et al., 2010, 2011; Schulz et al., 2010), richtete sich der Blick noch kaum auf die Daten der Lehrpersonen und damit verbunden auf die Vorbereitung und Realisierung von Unterricht im Bereich der politischen Bildung.

Auch wenn für den Begriff „Qualität“ im vorliegenden Fach keine allgemeingültige Definition besteht, so herrscht doch weitgehende Übereinstimmung, dass damit der Grad der Deckung von Erwartungen (Soll-Zustand) und Leistungen (Ist-Zustand) zum Ausdruck gebracht wird. Qualität von politischem Unterricht kann damit verbunden z.B. dem Anspruch gerecht werden, dass Schüler/innen über hohes politisches Wissen verfügen, was überprüft worden ist. Oder „guter“ politischer Unterricht wird dann festgestellt, wenn sich Schüler/innen gesellschaftlich engagieren, was wiederum empirisch untersucht worden ist.

Wenn nun hier der Blick auf die Lehrpersonen gerichtet wird, so verbindet sich damit jedoch eine andere Perspektive auf die Qualität von politischem Unterricht: im Mittelpunkt steht die Frage nach dem inhaltlich-professionellen Verständnis und Handeln der Lehrperson. Dabei interessieren Dimensionen wie theoretisch fundierte Ausbildung, abgegrenzte Kompetenzdomänen, Autonomie der Berufsausübung und das berufliche Ethos (vgl. z.B. Oser, 1998).

Vor dem Hintergrund dieser Zielsetzungen werden in diesem Beitrag folgende Fragestellungen verfolgt:

Über was für Ausbildungen, Vorstellungen, Konzepte und Kompetenzen bezüglich des Unterrichts politischer Bildung verfügen Lehrpersonen der Schweiz im Vergleich zu anderen ICCS-Ländern?

Wie bereiten Lehrpersonen in der Schweiz im Vergleich zu anderen ICCS-Ländern den Unterricht in politischer Bildung vor und wie setzen sie ihn um?

### Methode

In dieser Studie wurden insgesamt 62110 Lehrpersonen aus 4954 Schulen befragt. In der Schweiz liegen Antworten von 1571 Lehrpersonen aus 144 Schulen vor – wobei es zu beachten gilt, dass davon nur 326 Lehrpersonen während des befragten Zeitraums auch tatsächlich politischen Unterricht erteilten, wodurch die Analysen daher jeweils für die beiden Gruppen gesondert durchgeführt werden.

### Ergebnisse

Die Ergebnisse lassen erkennen, dass politische Bildung in der Schweiz im Vergleich mit den ICCS-Ländern auffallend selten als eigenständiges Fach unterrichtet (19% vs. 53% im ICCS-Durchschnitt –kurz ICCS) und auch nicht als Aufgabe aller Fächer (19% vs. 55% ICCS), sondern als Aufgabe der geistes- und sozial-wissenschaftlichen Fächer betrachtet wird (92% vs. 77% ICCS). Auch zeigen die Schweizer Lehrpersonen wenig Vertrautheit mit politischen Themen (M=45 vs. M=49 ICCS). Des Weiteren greifen die Schweizer Lehrpersonen in der Vorbereitung des Unterrichts viel stärker auf eigene Materialien (95% vs. 88%) und weniger auf amtliche Anforderungen (43% vs. 82% ICCS) zurück. Der Unterricht wird bedeutend weniger konstruktionsorientiert gestaltet (M=46 vs. M=49 ICCS) und die Bewertung der aktiven Beteiligung der Schüler/innen liegt mit am Ende des Ländervergleichs (M=43). Dennoch bringen die Lehrpersonen mit ihren Zielsetzungen politischen Unterrichts stärker ein kommunitarisches als funktionalistisches oder demokratisches Verständnis zum Ausdruck. Verschiedene korrelative Zusammenhänge sind erwartungskonform; so etwa wird gezeigt, dass politischer Unterricht und politische Selbstwirksamkeit negativ korrelieren. Eine regressionsanalytische Modellierung macht deutlich, dass Interesse, Selbstwirksamkeit und Wissensausprägungen höhere aufgeklärte Werte haben als die Qualität des Lehrerhandelns.

Die Ergebnisse erhellen, dass die Qualität politischen Unterrichts in der Schweiz unter der Perspektive des professionellen Verständnisses und Handelns deutlich defizitär ausfällt. Im Rahmen des Beitrags wird dieses Ergebnis kritisch diskutiert.

Biedermann, H., Oser, F., Konstantinidou, L. & Widorski, D. (2010). Staatsbürgerinnen und Staatsbürger von morgen: Zur Wirksamkeit politischer Bildung in der Schweiz – ein Vergleich mit 37 anderen Ländern. Universität Fribourg: Departement Erziehungswissenschaften. Online unter: [www.Universitätfr.ch/pedg/iccs/bericht1.pdf](http://www.Universitätfr.ch/pedg/iccs/bericht1.pdf).

Meyer, T. (2006). Was ist Politik. Stuttgart: UTB.

Oser, F. (1998). Ethos – die Vermenschlichung des Erfolgs. Zur Psychologie der Berufsmoral von Lehrpersonen. Opladen: Leske+Budrich.

Schulz, W., Ainley, J., Fraillon, J., Kerr, D. & Losito, B. (2010). ICCS 2009 International Report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).



## **Politikunterricht im Fokus der Kamera – Eine Videostudie zur Identifikation von Faktoren für wirksamen Fachunterricht**

In Videostudien, z.B. der Didaktik der Naturwissenschaften (z.B. Janik & Seidel, 2009), hat sich bestätigt, dass vor allem Tiefenstrukturen von Unterricht maßgeblichen Einfluss nehmen auf die Kompetenzentwicklung von Schülern, sowohl bzgl. des kognitiven Wissenserwerbs als auch der motivational-affektiven Entwicklung. Dabei gelten u.a. effizientes Class Room Management und ein motivierendes Lernklima als lernförderlich (vgl. Helmke, 2010). Zudem wird „der Beschaffenheit der Aufgaben und deren kognitiven Aktivierungspotential [...] eine wichtige Rolle beigemessen“ (Kunter et al., 2004, S. 176). Vor dem Hintergrund der Kompetenzorientierung sowie der Einführung von Bildungsstandards zum kumulativen Lernen erfordert wirksamer Unterricht neue Aufgabenformate: Zur Förderung der Politikkompetenz müssen Lernaufgaben „Fachbegriffe und Konzepte begreifbar machen, Stoffgebiete vernetzen [...] und KommUniversitätskation anstoßen“ (Weißeno, 2008, S. 415).

In der Politikdidaktik mangelt es an empirisch abgesicherten Kenntnissen über „alltäglichen“ Politikunterricht. Hier setzt die Studie mit der standardisierte Videografie an. Ziel ist es, mittels videogestützter Unterrichtsaufzeichnungen „core patterns of social science teaching“ zu erfassen. Dabei werden Unterstützungsleistungen unterrichtlicher Lehr-Lernprozesse beim Aufbau von Politikkompetenz systematisch identifiziert und Politikstunden von Praktikant/-innen und Lehrkräften verglichen, um erste Hypothesen zur Wirksamkeit von Politikunterricht zu generieren.

Folgende Forschungsfragen stehen u.a. im Zentrum der Erhebung: Inwiefern lassen sich allgemeine Merkmale von Unterrichtsqualität auf Politikunterricht übertragen? Werden im Politikunterricht gemäß der situierten Kognition aktuelle politische Ereignisse thematisiert und authentische Materialien eingesetzt? Lassen sich politische Fachkonzepte nach dem Konzepte-Modell (Weißeno et al., 2010) im alltäglichen Politikunterricht der Sek. I identifizieren? Welchen Komplexitätsgrad weisen diese auf? Unterscheiden sich Praktikant/-innen von Lehrkräften unter der Annahme des Experten-Novizenparadigmas?

Die theoretische Grundlage bildet das Angebots-Nutzungs-Modell sowie die Modelle der Expertiseforschung zu Praktikant/-innen und Referendaren als Novizen sowie länger praktizierenden Lehrkräften als Experten (vgl. König & Seifert, 2012). Für den Politikunterricht bietet das Modell der Basis- und Fachkonzepte nach Weißeno et al. (2010) zusammen mit dem ESNaS-Modell (Neumann et al., 2010, S. 290) ein Analyseinstrument, um den Aufbau konzeptuellen Fachwissens, den Vernetzungsgrad sowie die Anregung kognitiver Prozesse zu bestimmen. Das Kategoriensystem zum quantitativen Video-Rating wird sowohl deduktiv aus dem Kompetenzmodell von Detjen et al. (2012) als auch induktiv am Video-Material nach der qualitativen Inhaltsanalyse entwickelt (Mixed-Methods-Ansatz nach Mayring, 2010). Die Videos der Stichprobe von 8 Praktikant/-innen und 11 Lehrkräften in NRW werden standardisiert mit Videograph anhand von Kodiermanualen mit Ankerbeispielen von drei unabhängigen und geschulten Rater/-innen ausgewertet und kodiert. Die Rater/-innen-Übereinstimmung liegt bei den nominalskalierten Ratings bei einem Kappawert  $> 0.7$ . Die ordinalskalierten hoch-inferenten Ratings zeigen ebenfalls gute Übereinstimmungswerte von mindestens Spearman Rho  $> 0.7$  und Kendall-W  $> 0.8$ .

Erste Ergebnisse (vgl. Manzel & Gronostay, 2012) zeigen, dass im alltäglichen Politikunterricht überwiegend mit traditionellem Unterrichtsmaterial wie dem Schulbuch gearbeitet wird. Für auditive oder visuelle Lerntypen wird kein differenziertes Lernangebot gemacht, der analytische Zugang über gedruckte Texte steht im Vordergrund des Politikunterrichts. Die fehlende Einbettung des Stundenthemas in aktuelle politische Ereignisse und weitgehend abstrakte Lerninhalte verhindern die kognitive Auseinandersetzung der Schüler/-innen mit realer Politik. Fachkonzepte als Gradmesser für „gehaltvollen Fachunterricht“ werden mäßig thematisiert. Aufgabenstellungen sind überwiegend faktenorientiert und regen nur äußerst selten ein komplexes Vernetzungsniveau beim Aufbau eines Fachkonzepts bei den Lernenden an. Dies deckt sich mit Befunden aus den Naturwissenschaften (Wadouh 2007; Knobloch, Sumfleth, & Walpuski, 2011). Praktikant/-innen setzen für fachliche Aufgabenstellungen signifikant mehr Gruppenarbeit als Lehrer/-innen ein, während Lehrkräfte signifikant mehr Einzel- und Partnerarbeit für fachliche Aufgabenstellungen anwenden. Zudem unterscheiden sich beide Gruppen signifikant bei der Anregung von kognitiven Prozessen auf den Niveaus des Reproduzierens, Organisierens und Integrierens.

Detjen, J. et al. (2012). Politikkompetenz – Ein Modell. Wiesbaden: Springer VS.

Helmke, A. (2010). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. 3. Aufl., Seelze-Velber: Kallmeyer.

Janik, T. & Seidel, T. (Hrsg., 2009). The power of video studies in investigating teaching and learning in the classroom. Münster: Waxmann.

Manzel, S. & Gronostay, D. (2013). Videografie im Politikunterricht – Erste Ergebnisse einer Pilotstudie zu domänenspezifischen Basisdimensionen. In U. Riegel & K. Macha (Hrsg.), Videobasierte Kompetenzforschung in den Fachdidaktiken (Fachdidaktische Forschungen: Bd.4). (S. 198-215). Münster: Waxmann.

Mayring, P. (2010). Qualitative Inhaltsanalyse. 11., akt. u. überarb. Aufl., Weinheim u. Basel: Beltz.

Knobloch, R., Sumfleth, E. & Walpuski, M. (2011). Analyse der Schüler-Schüler-Kommunikation im Chemieunterricht. Chemie konkret: CHEMKON, Forum für Unterricht und Didaktik, 18 (2), 1-6.

König, J. & Seifert, A. (2012). Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung. Münster: Waxmann

Kunter, M. et al. (2004). Mathematikunterricht in den PISA-Klassen 2004: Rahmenbedingungen, Formen und Lehr-Lernprozesse. In PISA-Konsortium (Hrsg.), PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres. Münster: Waxmann.

Wadouh, J. (2007). Vernetzung und kumulatives Lernen im Biologieunterricht der Gymnasialklasse 9 [Internetdokument]. Verfügbar unter: [http://duepublico.Universität-duisburg-essen.de/servlets/DerivateServlet/Derivate-20492/DISSERTATION\\_komplett\\_5\\_BIB\\_end.pdf](http://duepublico.Universität-duisburg-essen.de/servlets/DerivateServlet/Derivate-20492/DISSERTATION_komplett_5_BIB_end.pdf) [10.10.2012]

Weißeno, G. et al. (2010). Konzepte der Politik. Ein Kompetenzmodell, Schwalbach/Ts.: Wochenschau.

Weißeno, G. (2008). Testaufgaben für die politische Bildung – Ergebnisse einer Pilotstudie. In: GPJE (Hrsg.): Testaufgaben und Evaluationen in der politischen Bildung. Schwalbach/Ts.: Wochenschau.



Chair(s): Michel Knigge (Martin-Luther-Universität Halle-Wittenberg)

Diskutant/in: Britta Pohlmann (Institut für Bildungsqualität Hamburg)

Der Lehrkraft kommt eine Schlüsselrolle im schulischen Geschehen zu, denn ihre Gestaltung des Unterrichts nimmt maßgeblichen Einfluss auf die Kompetenzentwicklung von Schülerinnen und Schülern. Dabei hängt es von den professionellen Kompetenzen einer Lehrpersonen ab, ob sie ihren Schülerinnen und Schülern im Sinne einer Kompetenzentwicklung einen qualitativ hochwertigen Unterricht anbietet (Kunter et al., 2013). Zu den professionellen Kompetenzen zählen neben dem Fach- und pädagogischen Wissen auch motivationale Faktoren, Einstellungen und Überzeugungen, die die Lehrkräfte im Laufe ihrer Biografie und natürlich auch in ihrem Studium erwerben und ausbilden. Ob Lehrkräfte Unterricht erfolgreich gestalten (können), so dass möglichst alle Schülerinnen und Schüler in hohem Maße gefördert werden, hängt darüber hinaus auch maßgeblich von Faktoren des Kontexts ab (Helmke, 2012).

Die Beiträge im eingereichten Symposium richten ihren Fokus aus unterschiedlichen wissenschaftlichen Disziplinen auf Lehrermerkmale, die die Bereitstellung von einem qualitativ hochwertigen Unterricht erleichtern oder erschweren können. Aus einer soziologischen Sichtweise werden vor allem Bildungsbiografien von Lehramtsstudierenden in Deutschland, Dänemark und Finnland als Kontextfaktoren betrachtet. Aus einer psychologisch-soziologischen Perspektive heraus werden dabei ergänzend motivationale Merkmale und Einstellungen der angehenden Lehrkräfte untersucht. Noch stärker psychologisch ist dann die Berücksichtigung von wahrgenommenen Einstellungen gegenüber Lehrerinnen und Lehrern in der Bevölkerung in Deutschland und ihre Effekten auf die Lehrkräfte. Zum Schluss werden aus einer erziehungswissenschaftlich-psychologischen Sicht Einstellungen und Überzeugungen sowie Selbstwirksamkeitserwartungen bei Lehramtsstudierenden in Bezug auf ihre Lehrplanungen untersucht. Das Symposium beginnt mit einem Beitrag von David Reimer und Hans Dorf von der Universität Aarhus. Sie vergleichen in ihren soziologischen Analysen Lehramtsstudierende zu Beginn ihres Studiums in Bezug auf Bildungswege, ihren soziodemografischen Hintergrund, Selbstkonzept und berufliche Motivation. Im zweiten Beitrag beschäftigt sich Johanna Ray ebenfalls mit Auswahlprozessen von Lehrkräften in Deutschland und Finnland, wobei auch hier ein besonderes Augenmerk auf Einstellungen und Überzeugungen gelegt wird. Der dritte Beitrag von Toni A. Ihme und Jens Möller untersucht aus einer psychologischen Perspektive, ob Lehramtsstudierende aufgrund der Wahrnehmung eines negativen Stereotyps über die Leistungsfähigkeit von Lehrkräften Leistungseinbußen zeigen, wenn das Stereotyp kognitiv aktiviert wird. Im vierten Beitrag werden Lehramtsstudierende in Halle und Hamburg aus einer psychologisch-erziehungswissenschaftlichen Zusammenarbeit von Michel Knigge und Carolin Rotter zu ihren Einstellungen zur Inklusion verglichen und es wird geprüft, inwieweit diese Einstellungen in einer Situation mit hoher Unsicherheit relevant für Unterrichtsplanungen sind. Abschließend wird Britta Pohlmann die Beiträge im Gesamtüberblick diskutieren.

Helmke, A. (2012). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (4. Auflage). Seelze: Klett-Kallmeyer.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805-820.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
IG-Farben-Haus  
Raum 0.251



## ***Lehrerselektion im Norden: Unterschiede zwischen dänischen und finnischen Lehramtsstudierenden***

Die nordischen Bildungssysteme sind in den letzten Jahren, vor allem wegen des guten Abschneidens finnischer Schülerinnen und Schüler in den PISA Untersuchungen vermehrt in den Fokus der internationalen und deutschen Bildungsforschung gerückt (vgl. Kobarg und Prenzel, 2009). Im vorliegenden Beitrag untersuchen wir Unterschiede in akademischen Selbstkonzepten, Motivationslagen und soziodemographischen Merkmalen von Lehramtsstudierenden in Dänemark und Finnland. Der Vergleich dieser beiden nordischen Länder ist vor dem Hintergrund der sehr unterschiedlichen Auswahlverfahren, Ausbildungsinhalten und Status der Lehrprofession in diesen beiden Ländern interessant. Finnland hat im Gegensatz zu Dänemark ein hochselektives Auswahlverfahren, in dem nur ca. 10-15% der Bewerber zugelassen werden (vergleiche auch den vorgeschlagenen Beitrag von Ray). In Dänemark hingegen wurden in den letzten Jahren nahezu alle Bewerber zugelassen (vgl. Dorf und Rasmussen, 2010). In einer eigenen Survey-Studie (N: 1313 (DK), 1383 (FI)) wurden sowohl Gymnasiasten im Abschlussjahrgang sowie Lehramtsstudierende im ersten Jahr zu ihren Wahlmotiven und anderen Merkmalen befragt. Um länderspezifische Faktoren von Charakteristika der Lehrerstudierenden analytisch trennen zu können führen wir eine einfache difference-in-difference Schätzung durch und vergleichen Unterschiede zwischen Gymnasiasten und Lehramtsstudierenden über beide Länder hinweg. Unsere Analysen zeigen, dass die dänischen Lehramtsstudierenden bei der Selbsteinschätzung ihrer Mathematikfertigkeiten deutlich hinter den finnischen Studierenden liegen. Darüber hinaus zeigen sich wie erwartet auch deutliche Unterschiede in den Motivationslagen zwischen den beiden Vergleichsgruppen. Diese Unterschiede verschwinden auch nicht, wenn in Kontrollanalysen nur die nach der Abiturnote besten 10% der dänischen Lehramtsstudierenden herangezogen werden. Implikationen unserer Befunde – und ein möglicher Bezug zur deutschsprachigen Forschung zum Thema der Lehramtsrekrutierung werden diskutiert.

## ***Komparative Betrachtungen zur Studierendenauswahl in der Lehrerausbildung anhand einer Analyse der aktuellen Vorgehensweise in Finnland***

Finnland hat eine lange Tradition, Lehramtsinteressenten vor Aufnahme des Studiums auf ihre Eignung zu überprüfen und ggf. bei nicht erfülltem Anspruch die Zulassung zu verwehren (Räihä, 2010). Im Gegensatz zum gegenwärtigen Diskurs in Deutschland (Rothland & Terhart, 2011) ist man im „PISA-Spitzenland“ von der Richtigkeit des Auswahlverfahrens bei der Studienplatzvergabe in der Lehrerausbildung überzeugt. Die selektierende Vorgehensweise geht allerdings nicht nur auf die hohen Bewerberzahlen zurück, die bei den beliebtesten Lehramtsstudiengängen eine Absagequote von bis zu 90% mit sich bringen (Sahlberg, 2012). Auch bei weniger beliebten Studiengängen wie etwa dem Studium von zukünftigen Mathematik- und Physiklehrkräften werden Aufnahmeprüfungen mit einem im Vorfeld festgelegten Cut-off-Wert flächendeckend eingesetzt (z.B. Kansanen, 2003). Dieses Vorgehen basiert auf der Überzeugung, dass der Lehrerberuf nicht von jedem erfolgreich ausgeübt werden kann bzw. dass es möglich ist, die „richtigen“ Personen im Vorfeld ausfindig zu machen (vgl. Valli & Johnson, 2007). Welche berufsrelevanten Kompetenzmerkmale stehen in Finnland im Vordergrund und wie werden diese erfasst?

Obwohl das finnische Bildungssystem aufgrund von PISA ein sehr großes internationales Interesse auf sich gezogen hat (siehe z.B. Kobarg & Prenzel, 2009), wurde dem Thema der strikten Selektionsverfahren in der Lehrerausbildung bisher erstaunlich wenig Beachtung geschenkt. Dies gilt sowohl in nationaler als auch in internationaler Hinsicht. Hinzu kommt, dass die vorhandenen Veröffentlichungen zum Thema mit diversen Einschränkungen behaftet sind. Mit dem Ziel, die Gestaltung der Bewerberselektion in der Lehrerausbildung in Finnland in ihren einzelnen Aspekten systematisch aufzuschlüsseln wurden die im Jahr 2013 eingesetzten Auswahlverfahren in einer deskriptiv angelegten Studie analysiert. Die Studie schloss die seit Jahren sehr begehrte, breit angelegte Lehrerausbildung für die Jahrgangsstufen 1-6 sowie die – wie in Deutschland – unter Bewerbermangel leidenden Studiengänge für Mathematik- und Physiklehrkräfte für die Jahrgangsstufen 7-9, ein. Die zentrale Fragestellung lautete: (i) Wie, (ii) auf welcher Grundlage und (iii) mit welchem Erfolg werden Bewerbende, die eine Tätigkeit in der finnischen „Pflichtschule“ anstreben, ausgewählt. Das empirische Material bestand aus Universitätsdokumenten und offiziellen Texten zu Konzeption, Zielsetzung und Durchführung der Aufnahmeprüfungen sowie zu den Aufnahmekriterien. Die Studie umfasste sieben finnischsprachige Universitäten sowie die schwedischsprachige Universität Finnlands.

Anders als in früheren Arbeiten nahegelegt wird, geht aus der vorliegenden Analyse deutlich hervor, dass nicht pauschal von einem einheitlichen Modell im Sinne einer landesweit praktizierten Vorgehensweise gesprochen werden kann. Nichtsdesto trotz gilt, dass an allen untersuchten Standorten großer Wert auf die persönliche Sichtung potentieller Studierender gelegt wird, wobei Einzelgespräche das am häufigsten eingesetzte Verfahren der hochschulspezifischen Auswahlverfahren ist. In Deutschland sind Interviews jedoch nur vereinzelt, und erst seit relativ kurzer Zeit (Faust et al., 2003), als Teil des Bewerbungsverfahren für Lehramtsinteressenten vorgesehen (Nieskens & Demarle-Meusel, 2012). Selbsterkundungsverfahren wie die in Deutschland gängigen Instrumentarien FIT oder FIBEL (siehe z.B. Rothland & Tirre, 2011) sind in Finnland nicht Bestandteil des Bewerbungsverfahrens. Eine Direktzulassung zum Studium ausschließlich anhand formaler Kriterien wie etwa der Abiturnote oder aber ausgewählter Einzelnoten – wie vielerorts in Deutschland üblich – kommt ebenfalls nicht vor. Die Studienplatzvergabe erfolgt in einigen Fällen sogar ohne jegliche Berücksichtigung schulischer Leistungen.

Anhand der Ergebnisse werden Möglichkeiten und Herausforderungen der Eignungsabklärung vor Beginn eines Lehramtsstudiums aus komparativer Sicht im Rahmen des Symposiums ausführlich diskutiert. Dabei wird der defizitären wissenschaftlichen Absicherung der in Finnland eingesetzten Auswahlinstrumente besondere Aufmerksamkeit gewidmet.

***“He who can, does; he who cannot, teaches”*: Stereotype Threat for teacher students**

Stereotype Threat (Steele & Anderson, 1995) bezeichnet eine situationale Bedrohung, die von einem negativen Stereotyp über die eigene soziale Gruppe ausgeht und zu Leistungseinbußen und Rückzug von den betroffenen Leistungsbereichen führt.

Lehramtsstudierende werden in der Literatur als hinsichtlich ihrer Leistungsfähigkeit negativ stereotypisiert dargestellt (Spinath, van Ophuysen & Heise, 2005) und sind somit potenziell durch Stereotype bedroht. Doch sowohl das Vorhandensein der postulierten Stereotype wie auch die Wirkung von Stereotype Threat auf Lehramtsstudierende ist bisher nicht untersucht worden.

Die hier vorgestellten Studien sollen ermitteln, ob die in der Literatur postulierten negativen leistungsbezogenen Stereotype gegenüber Lehramtsstudierenden existieren (Studien 1 & 2) und ob Lehramtsstudierende durch Stereotype Threat in ihrer Leistung beeinträchtigt werden (Studie 3).

Zur Beantwortung der ersten Frage zwei Befragungen (Fragebogen & Online) auf Basis des Stereotype Content Model (Fiske, Xu, Cuddy & Glick, 1999) durchgeführt. In Studie 1 (N = 82) wurden Lehramtsstudierende zu ihrer Wahrnehmung der Stereotype anderer zu ihrer Gruppe befragt. Es zeigte sich in offenen Antworten, dass Lehramtsstudierende hauptsächlich dem paternalistischen Stereotyp (niedrige Kompetenz, hohe Soziabilität) zugehörige Fremdwahrnehmungen schildern. In Studie 2 (N = 120) wurden Personen verschiedenster Gruppen (Arbeitnehmer, Studierende verschiedener Disziplinen etc.) befragt. Es bestätigte sich der Befund der ersten Studie mit einem geschlossenen Antwortformat im Vergleich zu Studierenden anderer Fächer. Studierenden jeder der Vergleichsgruppen (Psychologie, Jura & Informatik) wurde eine signifikant höhere Kompetenz zugeschrieben. Damit kann davon ausgegangen werden, dass ein Stereotyp niedriger Leistungsfähigkeit Lehramtsstudierender besteht.

In Studie 3 (N = 262) wurden Lehramtsstudierende im Rahmen eines Experiments durch dieses Stereotyp bedroht. Die TeilnehmerInnen führten einen kognitiven Leistungstest durch; in der Stereotype Threat-Bedingung wurde ihnen mitgeteilt, der Zweck dieser Testung sei, Leistungsunterschiede zwischen verschiedenen Studiengängen zu untersuchen. In Übereinstimmung mit dem Stereotype Threat-Ansatz zeigten Lehramtsstudierende, die durch das Stereotyp bedroht wurden, geringere Leistungen in einem kognitiven Fähigkeitstest als nicht bedrohte Kommilitonen, während sich für Studierende der Psychologie (vom Stereotyp nicht betroffene Kontrollgruppe) keine Unterschiede in der Leistung in Abhängigkeit von der Bedrohung ergaben. Mit der vorliegenden Studie sind somit erstmals Effekte von Stereotype Threat auf Lehramtsstudierende nachgewiesen worden.

Michel Knigge (Martin-Luther-Universität Halle-Wittenberg), Carolin Rotter (Universität Hamburg)

## ***Einstellungen zur Inklusion, Selbstwirksamkeit und Unterrichtsplanungen bei Lehramtsstudierenden – Ein Vergleich zwischen Studierenden in Halle und Hamburg***

Seit der Unterzeichnung der UN-Konvention für die Rechte mit Behinderungen (2009) ist Inklusion in aller Munde. Es kann als sicher gelten, dass Inklusion den beruflichen Alltag von angehenden Lehrkräften maßgeblich bestimmen wird. Auch im Rahmen der Universitätsversitären Lehrerbildung werden angesichts dieser Entwicklungen Veränderungen in der Gestaltung der Ausbildungsinhalte an vielen Orten vorgenommen.

Eine positive Einstellung gegenüber Inklusion kann als eine wichtige Grundvoraussetzung für eine produktive Auseinandersetzung mit dem Thema gesehen werden (Heyl, Janz, Trumpp, & Seifried, 2013). So konnte beispielsweise bereits gezeigt werden, dass eine positivere Einstellung gegenüber Inklusion mit einer höheren Selbstwirksamkeit zum Umgang mit Inklusion korrespondiert (Kopp, 2009). In diesem Sinne könnte die Entwicklung positiver Einstellungen gegenüber Inklusion selbst als ein Teil von professioneller Kompetenz verstanden werden.

In dem hier dargestellten Forschungsprojekt geht es darum, über bloße Zusammenhänge zwischen Selbstwirksamkeit und Einstellungen hinaus zu untersuchen, ob unterschiedliche Einstellungen und Selbstwirksamkeitserwartungen bei angehenden Lehrkräften in Situationen mit diagnostischer Unsicherheit zu unterschiedlichen Handlungsintentionen bezüglich der Unterrichtsgestaltung führen. Dabei ist die Annahme leitend, dass implizite Etikettierungsprozesse von Seiten der angehenden Lehrkräfte umso wahrscheinlicher werden, je negativer die Einstellung bzw. je niedriger die Selbstwirksamkeit ist. Dies folgt der Rationale, dass bei positiver Einstellung und hoher Selbstwirksamkeit im Umgang mit Inklusion und damit Heterogenität die subjektive Notwendigkeit der Anwendung von Etikettierungen weniger bedeutsam sein könnte, was sich auch in einem stärker inklusiven Unterricht ausdrücken würde.

Es wurden an zwei Standorten (Halle und Hamburg) jeweils mehrere hundert Lehramtsstudierende an Regel- sowie an Förder- bzw. Sonderschulen am Anfang der zweiten Hälfte ihres Studiums zu ihren emotionalen und kognitiven Einstellungen zur Inklusion und ihren korrespondierenden Selbstwirksamkeitserwartungen am Anfang und am Ende des Semesters befragt. Anschließend wurde ein Szenario vorgestellt, dass die Studierenden 4 Stunden in einer dritten Klasse unterrichten müssten. Um diagnostische Unsicherheit zu erzeugen, wurde ihnen ein kurzes Video aus dieser Klasse mit hoher Ambiguität in Bezug auf mögliche Förderbedarfe gezeigt. Im Anschluss wurden offene und geschlossene Fragen zu ihrer Vorbereitung auf ihren Unterricht in dieser Klasse gestellt.

Zum Zeitpunkt der Einreichung des Beitrages lagen die Daten noch nicht vollständig vor. Daher kann hier nur beschrieben werden, welche Analysen vorgenommen werden. Anhand dieser Daten werden nun die Hypothesen untersucht. Es werden Ergebnisse präsentiert inwieweit a) Unterschiede in Einstellungen und Selbstwirksamkeit in Bezug auf Inklusion sowie die Unterrichtsplanungen zwischen den Standorten und den Studiengängen bestehen, b) wie Zusammenhänge zwischen Einstellungen, Selbstwirksamkeit und Unterrichtsplanungen aussehen und ob sie sich zwischen den Standorten und Studiengängen unterscheiden. Darüber hinaus können c) Veränderungen in den Einstellungen, Selbstwirksamkeitserwartungen und Unterrichtsplanungen über ein Semester hinweg untersucht und mit Seminarbesuchen zum Thema Inklusion sowie schulischen und außerschulischen Erfahrungen mit Kindern mit Behinderung oder Förderbedarf in Beziehung gesetzt werden.

**GEBF** TAGUNG  
2014

D11 d

Symposiums  
beitrag

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
IG-Farben-Haus  
Raum 0.251

Lehrer werden in Deutschland und Skandinavien

2

## **Frühe mathematische Kompetenz - Entwicklung, Erfassung, Förderung**

Chair(s): Claudia Hildenbrand (Universität Hamburg), Anne-Katrin Jordan (IPN Kiel), Gabi Ricken (Universität Hamburg)

Diskutant/in: Jens Holger Lorenz (Universität Frankfurt)

In den letzten Jahren konnte immer besser herausgearbeitet werden, welche Konzepte und Kompetenzen sich bereits im Vorschulalter entwickeln und welche Bedeutung die vorschulische Förderung mathematischer Kompetenz hat. Bereits zu Beginn der Grundschule sind deutliche interindividuelle Unterschiede bezüglich mathematischer Kenntnisse zu beobachten. Ebenso kann belegt werden, dass sich dieses Wissen aufbauend aufeinander entwickelt. Dafür sprechen theoretische Überlegungen und zugleich empirische Daten. In allen Beiträgen des Symposiums steht die Frage des Erkennens und Nutzens mathematischen Potentials und damit verbunden die gezielte Förderung im Mittelpunkt. Gemeinsam ist allen, dass der Weg über die Qualifizierung der Erzieherinnen und Erzieher gewählt wird. Damit werden Fragen der Konzipierung der Qualifikation und der Erfassung der Veränderungen zentral. Grenzen und Möglichkeiten der alltagsintegrierten Förderung werden mit Blick auf die Wirkungen für Kinder und Erzieherinnen und Erzieher diskutiert.

Im ersten Beitrag wird eine theoretische Abfolge von zentralen Meilensteinen im Vorschulalter vorgestellt und empirisch anhand eines Niveaustufenmodells überprüft. Im zweiten Beitrag geht es um die aktionsbezogene Beschreibung und Erhebung von Kompetenzen von Vorschulpädagoginnen und -pädagogen mithilfe eines videobasierten Testinstruments. Im dritten und vierten Beitrag steht die Frage im Mittelpunkt, wie Förderung im Alltag der Kinder integriert werden kann. Dazu werden jeweils Fortbildungsprogramme entwickelt und in ihrer Wirkung beobachtet. Im dritten Beitrag werden ergänzend dazu Coachings durchgeführt in denen bedeutsame Situationen individuell analysiert und reflektiert werden. Im vierten Beitrag wird die Frage auf Varianten dieser Vorbereitung und Unterstützung von Pädagogen im Kindergarten erweitert.



## **Modellierung der Entwicklung mathematischer Konzepte im Vorschul- und Grundschulalter**

Mathematisches Wissen entsteht im Vorschulalter nicht beliebig. Eine Vielzahl von Befunden spricht dafür, dass das Erkennen von Zusammenhängen und Beziehungen jeweils eine spezifische Qualität aufweist, die wiederum jeweils eine Voraussetzung für die weitere Entwicklung darstellt. Zentrale Meilensteine im Vorschulalter sind das Zählzahlkonzept (Zahlen werden Objekten zugeordnet), ein ordinaler Zahlenstrahl (Zahlen werden als Folge in Beziehung gesetzt), die Kardinalität (Mengen bestehen aus einzelnen Elementen), ein Teil-Teil-Ganzes-Konzept (Mengen können variabel in Teilmengen aufgeteilt werden) und ein Relationalzahlaspekt (Beziehungen zwischen Mengen werden quantitativ bestimmt).

Diese aus der Zusammenstellung von einzelnen empirischen Befunden erhaltene theoretische Reihe ist empirisch durch eine Rasch-Modellierung zu prüfen, dabei ist von Bedeutung, ob sich eine eindimensionale Skala belegen lässt, die sich außerdem in der angenommenen Weise durch die Folge der Konzepte inhaltlich beschreiben lässt. Diese hätte dann Konsequenzen für die Diagnostik insofern, als das sich mit der Reihe ein individuelles Bezugssystem hinsichtlich der erworbenen, der aktuell entstehenden und zukünftigen Konzepte aufspannen ließe.

**Fragestellung und Methode**

Die Gültigkeit einer so zentralen Annahme von Entwicklungsfolgen ist in Stichproben mit unterschiedlich alten Kindern und mit unterschiedlicher Leistungsfähigkeit zu prüfen.

Dazu werden die MARKO-D Aufgaben verwendet und weitere Merkmale entsprechend der jeweiligen Fragestellung erhoben.

**Ergebnisse**

In einer ersten Hauptstudie wurden 1095 Kindern im Alter zwischen 48 und 87 Monaten mit 55 Items getestet, die die angenommenen Konzepte prüfen. Die Modellprüfung ergab für alle 55 Items einen MNSQ Infit im Bereich von  $1 \pm 0.3$ ; für 53 Items gilt sogar das strengste Kriterium  $1 \pm 0.2$  sowie eine ausreichend gute Trennung der Konzepte als Niveaustufen.

In Folgeuntersuchungen wurden Kinder mit Förderbedarf im Lernen (3. Klasse,  $n = 82$ ) untersucht und mit Grundschulkindern ohne Förderbedarf (1. Klasse,  $n = 422$ ) verglichen. Für beide Kindergruppen zeigt sich eine analoge Reihenfolge der Items auf einer Skala mit angemessenen Modellgütwerten. Daraus ist die Konsequenz abzuleiten, dass Entwicklungsverzögerungen als Verlangsamung bei gleicher Konzeptfolge darzustellen sind.

In einer dritten Studie wurde die Frage der Kulturabhängigkeit einbezogen. Für 236 südafrikanische Kinder mit den Muttersprachen Englisch, Sesotho und isiZulu konnte gezeigt werden, dass die Skala grundsätzlich gilt und nur in 5 Fällen Unterschiede in der Itemplatzierung zu beobachten waren. In einer aktuellen Studie werden in Gebärdensprache übersetzte Items des Modells auf ihre Tauglichkeit zur Skalierung bei gehörlosen Kindern geprüft.

Ricken, G., Fritz, A. & Balzer, L. (2013). Marko-D. Mathematik und Rechenkonzepte im Vorschulalter – Diagnose. Göttingen: Hogrefe.

Fritz, A., Ehlert, A., Ricken, G. & Balzer, L. (in Vorb.). Marko-D1. Mathematik und Rechnen. Göttingen: Hogrefe.

Dampier, G., Ehlert, A. & Fritz, A. (2013). Interkultureller Vergleich der Entwicklung arithmetischer Konzepte im Vorschulalter. South African Journal of Childhood Education.



Brigitte Hepberger (Hochschule für Heilpädagogik Zürich), Elisabeth Moser Opitz (Universität Zürich), Anke Lindmeier (IPN Kiel), Aiso Heinze (IPN Kiel), Imke Knievel (IPN Kiel), Franziska Vogt (Pädagogische Hochschule St. Gallen)

## **Aktions-bezogene mathematikdidaktische Kompetenzen von Vorschulpädagoginnen und -pädagogen erfassen**

Verschiedene Studien weisen insbesondere für Risikokinder auf die Bedeutung der vorschulischen mathematischen Förderung hin und es liegen empirische Daten zu erfolgreichen Förderprogrammen vor (z. B. Jordan, Kaplan, Ramineni, & Locuniak, 2009; Krajewski, Renner, Nieding, & Schneider, 2008). In der Mathematikdidaktik werden Ansätze präferiert, welche die Nutzung des mathematischen Potentials im Spiel und in alltäglichen Situationen betonen (Gasteiger, 2012; Rechsteiner, & Vogt, 2012). Die Umsetzung eines solchen Ansatzes setzt hohe fachdidaktische Kompetenzen der Pädagoginnen und Pädagogen voraus.

Während erste Untersuchungen zu deren Beliefs zum Mathematiklernen im Kindergarten vorliegen (Benz, 2012), sind stärker wissenschaftsbasierte Aspekte, wie etwa fachdidaktische Kompetenzen, bisher noch wenig untersucht worden. Insbesondere stellt sich die Frage, wie aktions-bezogene Kompetenzen beschrieben und erhoben werden sollen, die benötigt werden, um z. B. mathematikhaltige Situationen zu identifizieren und diese durch entwicklungsgerechtes Aufgreifen als Lerngelegenheiten zu nutzen.

Im vorgestellten Projekt wird zur Modellierung solcher Kompetenzen das Strukturmodell für Lehrpersonen von Lindmeier (2011) verwendet. Dieses schließt die oben skizzierten aktions-bezogenen Kompetenzen ein und unterscheidet zwei weitere Ebenen fachspezifischer Kognitionen: das Basiswissen (Fachwissen und fachdidaktisches Wissen, z.B. Prinzipien und Konzepte des Fachs) und die reflexiven Kompetenzen, die für die Vor- und Nachbereitung von Lerngelegenheiten benötigt werden (z.B. die Auswahl geeigneter Zugänge zu fachlichen Inhalten sowie die Analyse von Arbeiten von Lernenden). Offen bleibt die Frage, wie diese Kompetenzen gemessen werden können und ob sie als wissenschaftsbasierte Fähigkeiten mit starkem Praxisbezug vom Ausbildungsgang und der Erfahrung der Pädagoginnen und Pädagogen abhängig sind.

Fragestellung und Methode

Ziel der hier vorgestellten Studie ist die Entwicklung einer reliablen und validen Operationalisierung, die es erlaubt, die Kompetenzen von Vorschulpädagoginnen und -pädagogen in den drei Bereichen Basiswissen, reflexive und aktions-bezogene Kompetenzen zu erfassen. Dazu wurde ein Instrument entwickelt, für das ein klassisches Paper-Pencil-Verfahren durch video-basierte Items ergänzt wurde. Die befragten Personen müssen hierbei kurze Videovignetten betrachten und eine dazu präsentierte Frage direkt in ein Aufnahmegerät beantworten. Mit der dadurch implementierten Spontanität und Unmittelbarkeit sollen Situationen simuliert werden, wie sie im Kindergartenalltag auftreten, so dass aktions-bezogene Kompetenzen anforderungsnah erhoben werden können. Das in dieser Studie pilotierte Erhebungsinstrument enthält Items zu den drei Kompetenzbereichen mit folgenden Themen:

- Basiswissen (BK, 10 Items): Wissen zur Entwicklung des Zahlbegriffs und zu Fähigkeiten im Umgang mit Zahlen (z. B. Vorläuferfertigkeiten, Zählen, mathematischen Kompetenzen von Kindergartenkindern)
- Reflexive Kompetenzen (RC, 7 Items): Einschätzung von typischen Fähigkeiten, Schwierigkeiten sowie dem mathematischen Potential von Aktivitäten (z. B. Fehler beim verbalen Zählen eines Kindes erkennen, verschiedene Aktivitäten zur Förderung des Zahlbegriffs beurteilen).
  - Aktions-bezogene Kompetenzen (AC, 7 Items): Handeln in Situationen, präsentiert über Video-Vignetten (z. B. Aufgreifen authentischer mathematischer Kinderfragen zu großen Zahlen; Hilfestellung bei Abzählchwierigkeiten).

Zur Analyse der Antworten wurde ein – auf fachdidaktischer Literatur basierendes – Kodiermanual erstellt.

Erste Ergebnisse

Die Pilotierung stützt sich auf N = 66 Vorschulpädagoginnen und -pädagogen aus der Schweiz (N = 55) und aus Deutschland (N = 11). 41,8% verfügen über eine nicht-akademische Ausbildung, 58,2% über einen akademischen Abschluss.

In den ersten Analysen interessierte insbesondere die Reliabilität der Skalen (Cronbachs Alpha). Für das gesamte Instrument mit 24 Items ist diese mit  $\alpha = 0,86$  (rit .11 - .63) befriedigend. Die drei Subskalen weisen ähnlich reliable Werte auf (BK  $\alpha = 0,65$ , rit .27 - .43; RC  $\alpha = 0,79$ , rit .21 - .66; AC  $\alpha = 0,68$ , rit .23 - .59). Durch die Schätzung auf die Länge des Gesamttests mittels der Spearman-Brown-Formel wurde die Reliabilität der Subskalen auf folgende Werte korrigiert: BK = 0,86; RC = 0,90; AC = 0,88. Dies lässt die Folgerung zu, dass sich die drei theoretisch begründeten Kompetenzbereiche auch empirisch bestätigen lassen, wenn sie auch nicht klar voneinander zu trennen sind.

Erste Analysen weisen zudem auf Effekte von unterschiedlicher Ausbildung und Berufserfahrung hin.

Benz, C. (2012). Attitudes of Kindergarten Educators about Math. *Journal für Mathematikdidaktik*, 33, 203-232. DOI 10.1007/s13138-012-0037-7.

Jordan, N. C., Kaplan, D., Ramineni, C. & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Development Psychology*, 45 (3), 850–867.

Gasteiger, H. (2012). Fostering early mathematical competencies in natural learning situations - foundation and challenges of a competence-oriented concept of mathematics education in kindergarten. *Journal für Mathematik-Didaktik*, 33 (2), 181–201.

Krajewski, K., Renner, A., Nieding, A. & Schneider, W. (2008). Frühe Förderung von mathematischen Kompetenzen im Vorschulalter. In H.-G. Rossbach & H.-P. Blossfeld (Hrsg.) *Frühpädagogische Förderung in Institutionen. Zeitschrift für Erziehungswissenschaft*. 10 (11), 91–103 [Themenheft]. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lindmeier, A. (2011). Modeling and Measuring Knowledge and Competencies of Teachers. A Threefold Domain-Specific Structure Model for Mathematics. Münster: Waxmann.

Rechsteiner, K., Hauser, B., Vogt, F. (2012). Förderung der mathematischen Vorläuferfertigkeiten im Kindergarten: Spiel oder Training? In M. Ludwig & M. Kleine (Hrsg.), *Beiträge zum Mathematikunterricht 2012*. 2. Band (S.677-680). Münster: WTM Verlag für wissenschaftliche Texte und Medien.

Andrea Schulz (Universität Rostock), Timo Tresp (Universität Rostock), Tanja Jungmann (Universität Rostock), Katja Koch (Universität Rostock), Anne-Katrin Jordan (IPN Kiel), Christoph Duchhardt (IPN Kiel)

## Alltagsintegrierte Förderung mathematischer Kompetenzen im Kindergarten

### Projektbeschreibung

Eine Vielzahl von Alltagssituationen im Kindergarten ist für die mathematische Bildung nutzbar (u.a. Gasteiger 2010; van Oers, 2004). Im Rahmen des Projektes KOMPASS der Universität Rostock werden pädagogische Fachkräfte darin fortgebildet, mathematische Bildungspotenziale in Alltagssituationen zu erkennen, aufzugreifen und an die Interessen und Bedürfnisse der Kinder anzupassen. Inhalte der Fortbildung sind die Vermittlung theoriegeleiteten Wissens über mathematische Standards sowie didaktischer Prinzipien und Methoden, Beobachtungs- und Dokumentationsmöglichkeiten sowie die Förderung in Alltagssituationen. Die Inhalte werden in 8 Fortbildungsmodulen mit einem Gesamtumfang von 44 Stunden vermittelt.

Um Transferlücken zwischen dem theoretisch vermittelten Wissen und der Praxis zu schließen (Jungmann, Koch, & Etzien, 2013), wird die Fortbildung um ein pädagogisches Coaching ergänzt, in dem über 6 Monate in 7 Sitzungen didaktisch bedeutsame Situationen im Alltag des Kindergartens mit 5 Fachkräften individuell analysiert und reflektiert werden.

Weitere 11 Fachkräfte erhalten Interventionen in den Entwicklungsbereichen Sprache/Literacy (n=5) bzw. emotional-soziale Entwicklung (n=6). Insgesamt gehen Ergebnisse von 200 Kindern in die Auswertungen ein.

### Fragestellung

Zur Evaluation der Effektivität der Fortbildungsmodule auf die mathematische Entwicklung der Kinder werden deren mathematische Kompetenzen zu zwei Messzeitpunkten (im Abstand von 6 Monaten) mit dem Kieler Kindergartentest (KiKi) (Grüßing et al., 2013) erfasst. Mit der eingesetzten Form des KiKi können drei Bereiche („Mengen, Zahlen und Operationen“, „Veränderung und Beziehung“, „Raum und Form“) differenziert erfasst werden (Jordan et al., 2013). Es stellt sich nun die Frage, ob der Test auch zur Beschreibung der Kompetenzentwicklung geeignet ist und ob die Fortbildung diese zusätzlich positiv beeinflusst.

### Methode

Zunächst werden die Messinvarianz und die Stabilität der Aufgaben über die Zeit mit Methoden der Item-Response-Theorie geprüft. Im nächsten Schritt wird die mehrdimensionale Skalierung mit zwei Messzeitpunkten mit der Software ConQuest (Wu et al., 1997) durchgeführt. Die Effekte bereichsspezifischer Fortbildungen auf die mathematische Kompetenzentwicklung werden mittels Latent-Change-Modellen verglichen.

### Ergebnisse

Die Messinvarianzüberprüfung identifiziert sechs auffällige Items, die aus den weiteren Berechnungen ausgeschlossen werden. In der längsschnittlichen Skalierung weisen beide Dimensionen (Messzeitpunkt (MZP) 1 und 2) sehr gute Reliabilitäten auf (EAP/PVMZP1 = 0,896; EAP/PVMZP2 = 0,911). Die durchschnittliche Personenfähigkeit beträgt zum ersten MZP -1,422 Logits und zum zweiten MZP -0,274 Logits, d.h. dass bereits nach 6 Monaten ein beachtlicher Kompetenzzuwachs zu beobachten ist. Ergebnisse zu Fortbildungseffekten liegen derzeit noch nicht vor, werden aber im Beitrag vorgestellt.

### Ausblick

Es bleibt zu überprüfen, wie sich der Kompetenzzuwachs den einzelnen mathematischen Bereichen zuordnen lässt und wie sich additive pädagogische Coachings auf die kindliche Kompetenzentwicklung auswirken. Zur weiteren längsschnittlichen Erfassung der mathematischen Kompetenzen werden adaptierte, mit Linking-Items verknüpfte Testversionen des KiKi eingesetzt.

Gasteiger, H. (2010). Elementare mathematische Bildung im Alltag der Kindertagesstätte. Grundlegung und Evaluation eines kompetenzorientierten Förderansatzes. Münster: Waxmann.

Grüßing, M., Heinze, A., Duchhardt, C., Ehmke, T., Knopp, E. & Neumann, I. (2013). KiKi – Kieler Kindergartentest Mathematik zur Erfassung mathematischer Kompetenz von vier- bis sechsjährigen Kindern im Vorschulalter. in M. Hasselhorn, A. Heinze, W. Schneider, & U. Trautwein (Hrsg.), Diagnostik mathematischer Kompetenzen. (1. Ausg.) (S. 67-80). (Jahrbuch der pädagogisch-psychologischen Diagnostik. Tests und Trends). Göttingen: Hogrefe Verlag.

Jordan, A.-K., Duchhardt, C., Grüßing, M., Tresp, T., Knopp, E. & Heinze, A. (2013). Erfassung mathematischer Kompetenz im Kindergarten? Vortrag auf der AEPF-Tagung vom 25.-27.09. in Dortmund.

Jungmann, T., Koch, K. & Etzien, M. (2013). Effektivität alltagsintegrierter Sprachförderung bei ein- und zwei- bzw. mehrsprachig aufwachsenden Vorschulkindern. Frühe Bildung, 2 (3), 110-121.

Oers, B. van (2004). Mathematisches Denken bei Vorschulkindern. In Fthenakis, W. E., Oberhuemer, P. (Hrsg.), Frühpädagogik international. S. 313-330. Wiesbaden: VS Verlag.

Wu, M., Adams, R. J. & Wilson, M. R. (1997). ACER Conquest: Generalised item response modelling software. Melbourne: ACER Press.

## Förderung mathematischer Konzepte – Trainingsprogramm und alltagsintegrierte Bildung im Vergleich

Kinder verfügen schon zu Beginn der Grundschule über beachtliche mathematische Kenntnisse, die sich jedoch interindividuell deutlich unterscheiden können. Unterschiede im Wissen bleiben ohne spezielle Förderung über die Grundschuljahre hinweg stabil erhalten (Krajewski & Schneider, 2006) oder werden noch verstärkt (Aunola et al., 2006). Diese Befunde weisen auf die Bedeutsamkeit einer geeigneten frühen Förderung hin, durch die einerseits Entwicklungsrückstände ausgeglichen werden und andererseits die Kompetenzentwicklung aller Kinder unterstützt wird. Die Frage wie die Förderung mathematischer Kompetenzen in die Praxis umsetzen werden kann und wie pädagogische Fachkräfte gezielt auf diese Aufgabe vorbereitet werden können, ist jedoch noch weitgehend unbeantwortet und steht in der vorliegenden Studie im Fokus.

Es wird untersucht, ob sich die Effekte unterscheiden, wenn einerseits die Fortbildung von ErzieherInnen mit der Vorgabe und Anleitung eines strukturierten mathematischen Trainings verbunden wird und andererseits ErzieherInnen eine alltagsintegrierte und spielbasierte mathematische Anregung umsetzen.

Die empirische Untersuchung erfolgte nach einem quasi-experimentellen Design mit zwei Treatmentgruppen und einer Kontrollgruppe. Die ErzieherInnen der Treatmentgruppe 1 wurden in der Umsetzung des Fördertrainings Mina und der Maulwurf (Gerlach & Fritz, 2011) geschult. Dieses sollte von den ErzieherInnen in ihren Einrichtungen zu wöchentlich festen Terminen in Kleingruppen durchgeführt werden. Die TeilnehmerInnen der Treatmentgruppe 2 erarbeiteten sich theoriegeleitet vielfältige Förder-, Spiel- und Gestaltungsmöglichkeiten, die in einer alltagsintegrierten Förderung umgesetzt werden sollten. Die ErzieherInnen der Kontrollgruppe erhielten im Rahmen der Untersuchung keine Fortbildung.

Die mathematischen Konzepte der Kinder wurden mit einer gekürzten Version des Tests MARKO-D (Ricken, Fritz, & Balzer, 2013) in einem Prä-/Post-/Follow-up-Design gemessen. Sowohl die beiden Interventionsmaßnahmen als auch der eingesetzte Test basieren auf dem gleichen theoretischen Entwicklungsmodell (Fritz & Ricken, 2008), welches fünf aufeinander aufbauende Niveaus mathematischer Konzepte ausweist.

An der Studie waren insgesamt 290 Kinder (durchschnittliches Alter: 57 Monate) aus 30 Hamburger Kindertagesstätten beteiligt. Für die Analysen wurden die aus einer Rasch-Skalierung resultierenden WLE-Schätzer als Parameter für die individuell zugrundeliegenden mathematischen Konzepte genutzt. Fehlende Werte wurden mittels multiplen Imputation in SPSS generiert.

Eine einfaktorielle multivariate Varianzanalyse zeigte, dass sich die Untersuchungsgruppen vor der Intervention sowohl im Mittelwert der mathematischen Kompetenz, als auch hinsichtlich verschiedener Hintergrundvariablen signifikant unterscheiden. Aufgrund dieser Unterschiede wurden die drei Untersuchungsgruppen durch die Ziehung von vergleichbaren Teilstichproben parallelisiert (N = 90).

Die mathematischen Kompetenzen aller Kinder verbesserten sich im Untersuchungszeitraum, wobei jedoch in der parallelisierten Stichprobe keine statistisch signifikanten oder pädagogisch relevanten Unterschiede zwischen den Testergebnissen der beiden Programmgruppen und der der Kontrollgruppe festgestellt werden konnten. Dies betrifft sowohl den Posttest direkt nach der Intervention, als auch die Follow-up Erhebung ein Jahr später. Betrachtet man indessen die Entwicklung der Kinder, zeigt sich zwischen dem Prä- und Posttest ebenfalls kein signifikanter Unterschied. Zwischen dem Posttest und der Follow-up Erhebung gewinnen jedoch die Kinder der Kontrollgruppe signifikant mehr Punkte hinzu als die Kinder des Trainingsprogramms.

Für eine Extremgruppenanalyse wurden vergleichbare Kinder der nicht parallelisierten Ausgangsstichprobe ausgewählt, deren mathematischen Ausgangsleistungen auf dem niedrigsten Niveau lagen (N = 67). Im Vergleich mit der Kontrollgruppe zeigte sich, dass Kinder der Trainingsprogrammgruppe nach der Intervention signifikant höhere Werte erzielten und damit einen größeren Entwicklungsfortschritt erreichen konnten. Ein Jahr später ist dieser Unterschied knapp nicht mehr signifikant. Die schwächsten Kinder der Gruppe die eine alltagsintegrierte Förderung erhielten unterscheiden sich weder hinsichtlich der einzelnen Testleistungen noch in der Entwicklung von den Kindern der Kontrollgruppe.

Differentielle Ergebnisse hinsichtlich der Frage für welche Kinder welche Förderung wirksam ist, welche Faktoren als hemmend für die Effektivität und Implementation angesehen werden können und welche familiären und institutionellen Merkmale Varianz aufklären werden vorgestellt.

Aunola, K., Leskinen, E., Lerkkanen, M.-K. & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology*, 96(4), 699-713.

Fritz, A. & Ricken, G. (2008). *Rechenschwäche*. München: Reinhardt.

Gerlach, M. & Fritz, A. (2011). *Mina und der Maulwurf. Frühförderbox Mathematik*. Berlin: Cornelsen.

Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht*, 53, 246-262.

Ricken, G., Fritz, A. & Balzer, L. (2013). *MARKO-D. Mathematik- und Rechenkonzepte im Vorschulalter - Diagnose*. Göttingen: Hogrefe.

Dienstag,

04.03.2014,

14:15 - 16:00,

IG-Farben-Haus

Raum 0.254



Reinhold Nickolaus (Universität Stuttgart)

## ***Modellierung, Messung, erreichte Niveaus und Bezüge zwischen allgemeiner und berufsspezifischer Problemlösefähigkeit***

Chair(s): Reinhold Nickolaus (Universität Stuttgart), Susan Seeber (Universität Göttingen)

Diskutant/in: Detlev Leutner (Universität Duisburg-Essen)

Im Zuge der Kompetenzorientierung innerhalb des Bildungssystems hat die Fähigkeit problemhaltige Anforderungen bzw. Probleme zu bewältigen, einen wesentlich höheren Stellenwert als in der Vergangenheit erhalten. Ursächlich sind dafür letztlich stark ausgeprägte Bedeutungszuschreibungen, die der Umgang mit Problemen in alltags- und berufsbezogenen Situationen in den letzten Jahren erhalten hat und die zugleich verstärkte Forschungsaktivitäten zur Modellierung und Erklärung der Problemlösefähigkeit auslösten. Das gilt sowohl für das allgemeinbildende als auch das berufliche Bildungssystem und findet beispielsweise Ausdruck im Einbezug der komplexen Problemlösefähigkeit in den jüngeren PISA Studien und einer Reihe domänenspezifischer Arbeiten in der beruflichen Bildung, in welchen partiell auch die Bezüge zwischen komplexer Problemlösefähigkeit sensu Greiff, Wüstenberg und Funke (2012) und fachspezifischen Problemlöseleistungen analysiert wurden. Innerhalb der beruflichen Bildung kann domänenübergreifend dokumentiert werden, dass zumindest zwei Subdimensionen berufsfachlicher Kompetenz, das Fachwissen und die fachspezifische Problemlösefähigkeit, unterschieden werden können. Je nach beruflicher Domäne sind auch weitere Ausdifferenzierungen der fachspezifischen Problemlösefähigkeit naheliegend, jedoch bisher kaum empirisch geprüft. Gestützt wird die Annahme einer Mehrdimensionalität berufsfachlicher Problemlösefähigkeit bzw. der Wissensanwendung durch Winther (2010) im kaufmännischen Berufssegment. Weitgehend offen ist auch, inwieweit Problemlösekompetenzen in didaktischen Settings effektiv gefördert werden können. In diesem Symposium sollen Beiträge aus unterschiedlichen allgemeinbildenden und beruflichen Domänen auf verschiedenen Qualifikationsstufen (Medizinische Fachangestelltenberufe, zwei gewerblich-technische Berufe, Gymnasium im Hinblick auf Medizinalstudien) und der Psychologie zur „allgemeinen“ Problemlösefähigkeit zusammengeführt und ein interdisziplinärer Diskurs initiiert werden. Vorgestellt werden in diesem Kontext theoretische Modellierungen, Operationalisierungen, empirische Modellierungsergebnisse, Strukturanalysen zum Konstrukt der Problemlösefähigkeit in unterschiedlichen Kontexten und erste Erklärungsmodelle zur (berufsfachlichen) Problemlösefähigkeit.

Greiff, S., Wüstenberg, S. & Funke, J. (2012). Dynamic Problem Solving: A new assessment perspective. *Applied Psychological Measurement*, 36(3), 189-213.  
Winther, E. (2010). Kompetenzmessung in der beruflichen Bildung. Bielefeld: W. Bertelsmann.

**GEBF** TAGUNG  
2014

D13

Symposium

Dienstag,

04.03.2014,

14:15 - 16:00,

IG-Farben-Haus

Raum 1.411

Allgemeine und berufsspezifische Problemlösefähigkeit

2



## Die Erfassung komplexer Problemlösefähigkeiten auf Basis Finiter Automaten: Der MicroFIN Ansatz

### Theoretischer Hintergrund

Verbindungsstellen zwischen allgemeinen komplexen Problemlösefähigkeiten (KPL) und domänenspezifischen (beruflichen) Problemlösefähigkeiten bauen in Hinblick auf die Erfassung komplexer Problemlösefähigkeiten momentan wesentlich auf die Arbeiten von Greiff und Kollegen auf (Greiff, 2012; Greiff, Wüstenberg, & Funke, 2012; vgl. bspw. Abele et al., 2012).

Ein Vergleich und die empirische Annäherung beider Konzeptbereiche wird unter anderem durch spezifische Anforderungen an die Modellierung komplexer Probleme erschwert: Zur Erfassung individueller KPL Leistungen bauen Instrumente wie MicroDYN und Genetics Lab auf lineare Strukturgleichungen (LSE; Funke, 2001) zur Formalisierung von Problemen und erlauben damit eine reliable Schätzung individueller Problemlösefähigkeiten (vgl. Greiff et al., 2012).

Gleichzeitig weisen Probleme im beruflichen Kontext bzw. in schulbezogenen Fachbereichen wie Physik eine Reihe von Eigenschaften auf die sich schwerlich im Kontext linearer Strukturgleichungen in Instrumente übersetzen lassen (vgl. Scherer & Tiemann, 2012). Beispiele für derartige Problemeigenschaften sind etwa Gleichgewichtszustände oder ein Wechsel der Systemeigenschaften beim Überschreiten bestimmter Grenzwerte. Etablierte Instrumente zur Schätzung von KPL Leistungen lassen aufgrund der genutzten Formalisierung daraus resultierende Anforderungen an den Problemlöser jedoch außer acht.

In diesem Beitrag soll daher untersucht werden, in wie weit eine Nutzung des formalen Frameworks finiter Automaten (FSA; Funke, 2001) im Hinblick auf die Einschränkung in der Problembreite Abhilfe schaffen und damit eine empirische Lücke in der Verbindung zu fachspezifischen Problemlöseleistungen schließen kann. Finite Automaten wurden bis dato zwar in Laboruntersuchungen komplexer Problemlösefähigkeiten eingesetzt (Buchner & Funke, 1993; Buchner, 1999), nicht jedoch zur multi-item Erfassung von KPL Leistungen auf individueller Ebene.

### Fragestellung

Kann durch die Nutzung des formalen Frameworks finiter Automaten eine Erfassung komplexer Problemlöseleistung erreicht werden, die ein breites Spektrum von Problemen abdeckt? Wie gestalten sich empirische Bezüge zu etablierten Messinstrumenten des gleichen Konzepts und divergenten Maßen?

### Methode

Durch die Verwendung des FSA Frameworks lässt sich eine große Breite von Zusammenhängen modellieren, unter anderem auch die zuvor genannten Beispiele, wie Grenzwerte, bei deren Überschreiten sich das Systemverhalten grundlegend ändert.

Zur empirischen Prüfung der Nutzbarkeit des Frameworks wurden eine Reihe von komplexen Problemlöseaufgaben erstellt, die zwar in ihrem generellen Aufbau etablierten Instrumenten gleichen, jedoch gleichzeitig deren Beschränkungen im Hinblick auf Problemeigenschaften vermeiden. Einem Vorschlag von Greiff und Funke (2009) folgend, wurden die Aufgaben unter dem Namen MicroFIN zusammengefasst.

Zur Ermittlung ihres empirischen Nutzens wurden die Aufgaben von 576 Schülern bearbeitet (314 Mädchen, Alter zwischen 13 und 18,  $M = 14.95$ ,  $SD = 1.30$ ) und die Ergebnisse mit denen eines etablierten Instruments zur KPL Erfassung (MicroDYN, vgl. Greiff et al., 2012) und einem Test schlussfolgerndem Denkens verglichen (KFT; Heller & Perleth, 2000).

### Ergebnisse

Die fünf auf FSA aufbauenden Aufgaben zeigten die erwartete empirische Unterscheidbarkeit der beiden Dimensionen Wissenserwerb und Wissensanwendung ( $\chi^2(89) = 143.928$ ,  $p < .001$ ,  $RMSEA = 0.033$ ,  $CFI = 0.978$ ,  $TLI = 0.974$ , sign. Modellverschlechterung für eindimensionales Modell) und eine akzeptable Reliabilität bei der Erfassung dieser: McDonalds omega (Zinbarg, Revelle, Yovel, & Li, 2005) von  $\omega = .79$  (Wissenserwerb) und  $\omega = .78$  (Wissensanwendung).

Die Ergebnisse der verschiedenen Instrumente wurden mit Hilfe von multitrait-multimethod Analysen ausgewertet (CT-C(M-1) Modell; Eid, Lischetzke, Nussbeck, & Trierweiler, 2003). Ein Modell zum Vergleich mit MicroDYN Aufgaben bestätigte die konzeptionelle Nähe beider Instrumente ( $\chi^2(285) = 445.247$ ,  $p < .001$ ,  $RMSEA = 0.031$ ,  $CFI = 0.980$ ,  $TLI = 0.978$ ), zeigte jedoch gleichfalls die Notwendigkeit einer separaten Modellierung von Methodeneffekten auf. Der Bezug zu schlussfolgerndem Denken bestätigte die moderate Beziehung beider Konzepte ( $\chi^2(361) = 540.359$ ,  $p < .001$ ,  $RMSEA = 0.029$ ,  $CFI = 0.978$ ,  $TLI = 0.976$ , latent Korrelationen von .63 und .60 für die beiden KPL Dimensionen).

### Diskussion

Die Instrumente werden momentan im beruflichen Kontext eingesetzt, eine empirische Prüfung des Zusammenhangs der zusätzlichen Anforderungen mit solchen aus dem spezifischen Kontext steht daher bevor.

Abele, D.-G. S., Greiff, D. S., Gschwendtner, T., Wüstenberg, S., Nickolaus, R., Nitzschke, A., & Funke, J. (2012). Dynamische Problemlösekompetenz. Zeitschrift für Erziehungswissenschaft, 15(2), 363–391. doi:10.1007/s11618-012-0277-9

Buchner, A. (1999). Komplexes Problemlösen vor dem Hintergrund der Theorie finiter Automaten. Psychologische Rundschau, 50(4), 206–212.

Buchner, A., & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. The Quarterly Journal of Experimental Psychology Section A, 46(1), 83–118.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. Psychological Methods, 8(1), 38–60. doi:10.1037/1082-989X.8.1.38

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. Thinking & Reasoning, 7(1), 69–89. doi:10.1080/13546780042000046

Greiff, S. (2012). Individualdiagnostik der komplexen Problemlösefähigkeit [Diagnostics of problem solving ability on an individual level]. Münster: Waxmann.

Greiff, S., & Funke, J. (2009). Measuring complex problem solving: the MicroDYN approach. In The Transition to Computer-Based Assessment - Lessons learned from large-scale surveys and implications for testing. (pp. 157–163). Luxembourg: Office for Official Publications of the European Commission.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving: A new assessment perspective. Applied Psychological Measurement, 36(3).

Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. [Cognitive Abilities Test (CogAT); Thorndike, L. & Hagen, E., 1954-1986] - German adapted version. Göttingen: Beltz-Test.

Scherer, R., & Tiemann, R. (2012). Factors of problem-solving competency in a virtual chemistry environment: The role of metacognitive knowledge about strategies. Computers & Education, 59(4), 1199–1214. doi:10.1016/j.compedu.2012.05.020

Dienstag,

04.03.2014,

14:15 - 16:00,

IG-Farben-Haus

Raum 1.411



## Berufliche Problemlösekompetenzen bei Medizinischen Fachangestellten

Im Rahmen des Forschungsprojekts „Competence Measurement CoSMed wird ein Assessment zur Messung beruflicher Kompetenzen für Medizinische Fachangestellte (MFA) entwickelt, auf dessen Grundlage berufliche Fach- und Problemlösekompetenzen erfasst werden. In dem Beitrag werden Forschungsbefunde zu beruflichen Problemlösekompetenzen bei MFA diskutiert. In Anlehnung an ein Evidence-Centered Assessment-Design (Mislevy & Haertel 2006) wurde zunächst ein Domänenmodell entwickelt, in dessen Rahmen typische Arbeitsprozesse und -ergebnisse sowie erforderliches Wissen und notwendige Fähigkeiten näher bestimmt wurden. Mit Hilfe der Critiquel Incident Technique wurden kritische Situationen identifiziert, die typisch für Problemsituationen sind: Es sind keine Handlungsroutinen verfügbar, es tritt eine unbefriedigende Situation ein (Anderson 2001), intransparente Informationen sind zu entschlüsseln, Ziele des Handelns müssen identifiziert und Lösungswege gegeneinander abgewogen werden (vgl. Leutner et al. 2012). Auf der Grundlage der Beobachtungen und Interviews in Arztpraxen wurde ein mehrdimensionales Kompetenzmodell spezifiziert, das (1) Kompetenzen in der personenorientierten, situationsgerechten Beratung und Betreuung von Patienten, (2) medizinische Fachkompetenzen im Bereich von Hygiene- und Präventionsmanagement sowie Labordiagnostik und (3) kaufmännisch-verwaltende Fachkompetenzen im Bereich des Praxismanagement als Subdimensionen unterscheidet. Für die Erfassung konzeptuellen Wissens wurde ein adaptiver entwickelt, der Aufgaben aus allen drei Kompetenzdimensionen enthält. Um die beruflichen Handlungs- und Problemlösekompetenzen zu messen, wurden authentische Handlungssituationen für die drei Kompetenzdimensionen modelliert und auf der Grundlage von Videoclips dargestellt, ferner wurden für die Situierung und Lösung von Aufgaben fiktive Datenbestände und Arbeitsdokumente in eine web-basierte Testumgebung eingebunden. Im beruflichen Bereich liegen bisher nur begrenzt Forschungsbefunde zu den Problemlösekompetenzen von Auszubildenden in ausgewählten Berufen vor (vgl. Abele et al. 2012). Die MFA sind insofern ein interessanter Beruf, als sie zu jenen Berufen gehören, die zwar einen fachlichen Kern haben, der aus medizinisch-gesundheitsbezogene Anforderungen besteht, die jedoch gleichzeitig kaufmännische Qualifikationsanforderungen im Profil haben (Aufgaben des Praxismarketing, der Betriebsorganisation, der Beschaffungsprozesse, der Abrechnung ärztlicher Leistungen etc.). Bislang liegen keine Erkenntnisse vor, wie sich Problemlösekompetenzen in diesen sehr unterschiedlichen Handlungsbereichen, bei denen zur Problemlösung jeweils Sach- und Handlungswissen aus verschiedene Bezugswissenschaften, Medizin und Naturwissenschaft auf der einen Seite, Ökonomie und Mathematik auf der anderen Seite, heranzuziehen sind, in ihrer Struktur und Ausprägung unterscheiden.

### Forschungsfragen

Lassen sich die Problemlösekompetenzen von MFA besser in einem ein- oder zweidimensionalen Modell, bei dem den unterschiedlichen Bezugsdisziplinen Rechnung getragen wird, abbilden? In welchem Zusammenhang stehen die Problemlösekompetenzen mit dem Handlungs- und Sachwissen der jeweils affinen Bezugsdisziplin? Inwiefern zeigen sich differenzielle Effekte in der Ausprägung der Problemlösekompetenz nach verschiedenen Gruppen?

### Methodisches Vorgehen

Für die Analysen stehen Daten aus den Piloterhebungen im Sommer 2013, die in verschiedenen Bundesländern erhoben wurden, zur Verfügung. Für die Beantwortung der Forschungsfragen werden nur jene Fälle zugrunde gelegt, die sowohl den adaptiven Test zur Erfassung konzeptuellen Wissens im medizinischen und kaufmännischen Bereich bearbeitet haben als auch das video- und simulationsgestützt Testmodul, in das die Problemlöseaufgaben eingebettet sind (N=248). Die Tests wurden auf der Grundlage ein- und mehrdimensionaler Rasch-Modelle skaliert.

### Ergebnisse

Die Auswertungen sind noch nicht vollständig abgeschlossen, aber es zeichnet sich nach der ein- und zweidimensionalen Rasch-Skalierung und dem Vergleich verschiedener Güteindizes keine eindeutige Präferenz für das ein- oder zweidimensionale Modell der Problemlösekompetenz ab. In der Ausprägung der beruflichen Problemlösekompetenzen zeigen sich bedeutsame Unterschiede in Abhängigkeit vom schulischen Abschluss.

Abele, S., Greiff, S., Gschwendtner, T., Wüstenberg, S., Nickolaus, R., Nitzschke, A. & Funke, J. (2012). Dynamische Problemlösekompetenz. Ein bedeutsamer Prädiktor von Problemlöseleistungen in technischen Anforderungskontexten? Zeitschrift für Erziehungswissenschaft

Anderson, J. R. (2001). Kognitive Psychologie (3. Aufl.). Heidelberg: Spektrum Akademischer Verlag.

Leutner, D., Fleischer, J., Wirth, J., Greiff, S. & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien. Psychologische Rundschau, 63 (1), 34–42

Mislevy, Robert J. & Haertel, Geneva D. (2006). Implications of Evidence-Centered Design for Educational Testing. SRI International and University of Maryland, Ravenswood.



## **Problemlösekompetenzen von Maturandinnen und Maturanden im Hinblick auf medizinale Studiengänge und Berufe**

### Theoretischer Hintergrund

Als eines der Hauptziele des (Schweizer) Gymnasiums gilt die Erlangung einer allgemeinen Studierfähigkeit (MAR 95, 1995). Dieses Konstrukt lässt sich zunächst beschreiben als Gesamtheit aller Kompetenzen (Kenntnisse, Fähigkeiten, Fertigkeiten und Bereitschaften), die unabdingbare Voraussetzung für die erfolgreiche Bewältigung eines Universitätsstudiums sind, also befähigen, ein solches erfolgreich zu beginnen, durchzuführen und abzuschliessen (vgl. Huber, 2009, S. 108). Allgemeine Studierfähigkeit meint dann, dass sich diese Kompetenzen in Übereinstimmung mit der allgemeinen Hochschulzugangsberechtigung auf jedes Hochschulstudium in jedem Fach erstrecken. Über die Elemente der allgemeinen Studierfähigkeit existieren verschiedene Theoriefragmente (siehe auszugsweise in Eberle et al., 2008, S. 26 ff.). Sie umfasst in einer groben Aufteilung die folgenden Komponenten (Eberle, 2013, S. 122 f.): 1. Überfachliche kognitive und nicht kognitive Kompetenzen; 2. Fachliches Spezialwissen und Fachkönnen, das nur in einzelnen Studienfächern vorausgesetzt wird; 3. Fachwissen und Fachkönnen, das nicht nur von einzelnen, sondern von vielen Studienfächern vorausgesetzt wird (basale fachliche Kompetenzen für die allgemeine Studierfähigkeit (siehe dazu auch Huber, 2009, S. 118 ff.), dazugehörend insbesondere Wissen und Können aus der Erstsprache, der Wissenschaftssprache Englisch und aus der Mathematik sowie Informatikanwenderkompetenzen). Allgemeine Problemlösekompetenzen können der ersten Komponente, fachspezifische den Komponenten zwei und drei zugeordnet werden. Inwieweit Maturandinnen und Maturanden mit bestandener Matura tatsächlich über die so beschriebene allgemeine Studierfähigkeit und insbesondere über die notwendigen Problemlösefähigkeiten verfügen, ist angesichts der grossen Heterogenität und formalen Unbestimmtheit der im Schweizer System des Gymnasiums möglichen individuellen Bildungswege eine eher offene Frage. Bereits der für das Gymnasium verbindliche Rahmenlehrplan (EDK, 1994) ist wenig theoriegestützt und sehr offen. Es besteht eine Vielfalt von im Rahmen der gesetzlichen Vorgaben möglichen Strukturvarianten von Fächerzusammensetzungen (Schwerpunkt- und Ergänzungsfächer), zeitlichen Dotierungen, und die teilweise bis auf die Ebene von Einzelklassen gewährte Prüfungsautonomie (bei Erfahrungsnoten, aber auch bei den Maturaprüfungen) führt zu nur beschränkt vergleichbaren Noten, den Indikatoren für die allgemeine Studierfähigkeit.

### Fragestellung

Im Beitrag wird untersucht, wie weit Maturandinnen und Maturanden über die Problemlösekompetenzen verfügen, die im Studium und in der anschliessenden Berufstätigkeit vorausgesetzt werden. Die Fragestellung wird auf das potenzielle Studium der Medizin und der anschliessenden Berufstätigkeit eingegrenzt. Dabei werden sowohl die Komponenten der studien- und berufsbezogenen Problemlösefähigkeit untersucht als auch die Zusammenhänge zu den individuellen und systemischen Heterogenitätsmerkmalen aufgezeigt.

### Methode

Die Daten wurden im Rahmen der nationalen Studie EVAMAR II (Eberle et al., 2008) erhoben, in der ausgewählte Subkomponenten der drei beschriebenen Komponenten allgemeiner Studierfähigkeit mittels Tests und Selbsteinschätzung erfasst wurden. Dabei wurde in der Deutschschweiz (N = 2'598) auch ein Test zur Erfassung Problemlösekompetenzen eingesetzt, in dem – in Anlehnung an den Eignungstest für das Medizinstudium (Hänsgen & Spicher, 2012) – zu lösende Problemstellungen mit den notwendigen Informationen vorgegeben sind.

### Ergebnisse

Im Beitrag werden die Dimensionalität der Problemlösefähigkeit (Quantitative und formale Probleme, Textverständnis, Naturwissenschaftliches Grundverständnis, Interpretation von Diagrammen und Tabellen), der erreichte Kompetenzstand in Abhängigkeit von Heterogenitätsmerkmalen wie dem im Gymnasium gewählten Schwerpunktfach, der Dauer des Gymnasiums und des Geschlechts sowie die Zusammenhänge zu den Fachtests in Erstsprache, Biologie und Mathematik und weiteren Elementen der allgemeinen Studierfähigkeit aufgezeigt.

Eberle, F u.a. (2008). Evaluation der Maturitätsreform 1995. Bern.

Eberle, F. (2013). Kompetenzentwicklung in Gymnasien im Spannungsfeld von Studierfähigkeit und vertiefter Gesellschaftsreife. In S. Seufert & C. Metzger (Hrsg.). Kompetenzentwicklung in unterschiedlichen Lernkulturen. S. 116–129, Paderborn

EDK (1994). Schweizerische Konferenz der kantonalen Erziehungsdirektoren. Rahmenlehrplan für die Maturitätsschulen (Dossier 30A). Bern.

Huber, L. (2009). Von „basalen Fähigkeiten“ bis „vertiefte Allgemeinbildung“: Was sollen Abiturienten für das Studium mitbringen? In D. Bosse (Hrsg.), Gymnasiale Bildung zwischen Kompetenzorientierung und Kulturarbeit (S. 107-124). Wiesbaden.

MAR 95 (1995). Verordnung des Bundesrates über die Anerkennung von gymnasialen Maturitätsausweisen.

Hänsgen, K.-D., & Spicher, B. (2012). EMS Eignungstest für das Medizinstudium in der Schweiz 2012: Freiburg.



## **Berufsfachliche Problemlösekompetenzen von Kfz Mechatroniker/-innen – Modellierung und erreichte Niveaus**

### Theoretischer Hintergrund

Berufsfachliche Problemlösekompetenzen erhalten aufgrund sich dynamisch entwickelnder Qualifikationsanforderungen seit Jahrzehnten einen herausragenden Stellenwert in den Kompetenzdebatten zur beruflichen Bildung (Winther/Achtenhagen, 2008; Nickolaus/Seeber 2013). Im gewerblich-technischen Bereich werden die Problemlösekompetenzen in der Regel in konstruktive und analytische Varianten ausdifferenziert, wobei bei Kfz Mechatronikern unterstellt werden kann, dass primär analytische Problemlösekompetenzen im Kontext von Fehlerdiagnosen eingefordert werden (Abele et al., 2012; Nickolaus et al., 2012). Diagnostiziert werden die analytischen Problemlösekompetenzen gestützt auf Simulationen technischer Systeme, die hoch authentisch modelliert werden zuzüglich der diagnostischen Instrumente, die Fehlerdiagnosen analog zu jenen in realen Systemen ermöglichen. Erste Studien zu den erreichten Niveaus (Nickolaus et al., 2012) dokumentieren deutliche Diskrepanzen zwischen curricularen Ansprüchen und den erreichten Niveaus. Zudem verdeutlichen sie die Schwierigkeit, diese Kompetenzfacette mit einem authentischen (computerbasierten) Test in akzeptablen Testzeiten sowohl valide als auch reliabel zu erheben. Es erwies sich als empirisch nicht tragfähig, das Reliabilitätsproblem dadurch zu minimieren, zusätzlich zu den zeitaufwändigen computerbasierten Items Papier- und Bleistift-basierte Items vorzugeben.

### Fragestellung

Können mit neuen Aufgabenformaten, die in kürzerer Bearbeitungszeiten bewältigt werden können, jedoch wie die authentischen Fehlerfälle systematische Interaktionen mit den simulierten Systemen erfordern, die Reliabilitätsprobleme gelöst und gezieltere Informationen für die Niveaumodellierung gewonnen werden? Werden mit den modifizierten Aufgaben die gleichen Personfähigkeiten erfasst wie mit den authentischen Simulationen?

### Methode

Zur Klärung dieser Fragen wurde ein neues Aufgabenformat erprobt. Vorgestellt werden in diesem Beitrag a) das neue Aufgabenformat und einschlägige empirische Befunde und b) Niveaumodelle zur berufsfachlichen Problemlösekompetenz bei Kfz-Mechatroniker/-innen. Basis der Analysen ist eine im zweiten (n=100) und dritten Ausbildungsjahr (n=270) im SPP 1239 bzw. einem ASCOT Projekt durchgeführte Studie, in der neben der berufsfachlichen Problemlösekompetenz auch verschiedene Dimensionen des berufsfachlichen Wissens erhoben wurden.

### Ergebnisse

Die bisherigen Analysen legen nahe, dass das neue Format geeignet ist, das angesprochene Reliabilitätsproblem zu lösen. Außerdem ermöglichen die neuen Aufgaben im Vergleich zum bisherigen Vorgehen eine präzisere Identifikation relevanter Barrieren und damit auch eine präzisere inhaltliche Beschreibung von Kompetenz-niveaus. Mit den neuen Testformaten kann in wesentlich kürzerer Testzeit als bisher mit guter Reliabilität (0.81) und hoch valide eine Kompetenzabschätzung vorgenommen werden.

### Diskussion

Der neue Aufgabentypus könnte in gewerblich-technischen Kontexten, in welchen simulationsgestützte Testverfahren zur Erfassung der fachspezifischen Problemlösekompetenz in Frage kommen, generell geeignet sein, das Reliabilitätsproblem bei vertretbaren Testzeiten zu lösen. Vorteilhaft scheint der neue Aufgabentypus auch für die Diagnose relevanter Barrieren, da damit gezielt Teilleistungen angesprochen werden können, die in vielschrittigen und komplexen Fehleranalyseprozessen zu erbringen sind.

Abele, S. (2013). Modellierung, Entwicklung und Determinanten berufsfachlicher Kompetenz in gewerblich-technischen Ausbildungsberufen. Analysen auf Basis von testbasierten, berufsschulischen und betrieblichen Leistungsdaten sowie Prüfungsergebnissen. Aachen: Shaker, Universität Stuttgart, Dissertation.

Abele, S./Greiff, S./Gschwendner, T./Wüstenberg, S./Nickolaus, R./Nitzschke, A./Funke, J. (2012): Die Bedeutung über-greifender kognitiver Determinanten für die Bewältigung beruflicher Anforderungen – Untersuchung am Beispiel Dynamischen und Technischen Problemlösens. In: Zeitschrift für Erziehungswissenschaft, Bd. 15, H. 2, S. 363-391

Nickolaus, R./ Abele, S./ Gschwendner, T./ Nitzschke, A./ Greiff, S. (2012). Fachspezifische Problemlösefähigkeit in gewerblich-technischen Ausbildungsberufen – Modellierung, erreichte Niveaus und relevante Einflussfaktoren. In: Zeitschrift für Berufs- und Wirtschaftspädagogik (ZBW), Jg. 108, H. 2, S. 243-272.

Nickolaus, R./ Seeber, S. (2013). Berufliche Kompetenzen: Modellierungen und diagnostische Verfahren In: Frey, A./ Lissmann, U./ Schwarz, B. (Hrsg.). Handbuch berufspädagogischer Diagnostik (im Druck).

Winther, E./ Achtenhagen, F. (2008): Kompetenzstrukturmodell für die kaufmännische Ausbildung. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 104, (4) S. 511-538



Chair(s): Anke Wischgoll (Universität Freiburg)

Diskutant/in: Cordula Artelt (Universität Bamberg)

Von Studierenden der Bildungswissenschaften wird erwartet, dass sie in ihrer Ausbildungszeit Kompetenzen im Lesen und Schreiben von wissenschaftlichen Texten entwickeln. Sich mit wissenschaftlicher Literatur auseinanderzusetzen, daraus Schlussfolgerungen zu ziehen und anzuwenden sowie eigene fachliche Texte zu produzieren ist ein wesentlicher Bestandteil vieler bildungswissenschaftlicher Studiengänge. Jedoch ist in den jeweiligen Studiengängen für den Erwerb dieser Kompetenzen zumeist keine Ausbildungszeit und Ausbildungsform vorgesehen. Zudem sind die Kompetenzen der Studierenden im Umgang mit wissenschaftlicher Literatur empirisch kaum erforscht. Es fehlen sowohl Daten, die Aussagen über den Stand der einzelnen Lese- und Schreibkompetenzen zulassen, als auch darauf aufbauend geeignete Trainingsprogramme zur Verbesserung der einzelnen Teilkompetenzen. Wir hoffen mit unserem Symposium einen Beitrag zur Schließung dieser Lücke leisten zu können.

Das Symposium umfasst vier Präsentationen, die sich mit dem Umgang mit wissenschaftlicher Originalliteratur in bildungswissenschaftlichen Studiengängen (Erziehungswissenschaft und Psychologie) befassen. In den beiden ersten Beiträgen werden Kompetenzmodelle vorgestellt, auf denen die Erfassung der Lese- und Schreibkompetenzen mittels neu entwickelter Testverfahren basiert, in den beiden letzten Beiträgen werden Ergebnisse, die über Trainings zur Verbesserung von Teilkompetenzen erreicht wurden, vorgestellt:

1) Nach Verarbeitungszielen und –strategien werden die Kompetenzen im Beitrag "Diagnostik von Kompetenzen Studierender im Umgang mit wissenschaftlicher Originalliteratur" aus dem Projekt Kompetenzen Studierender im Umgang mit wissenschaftlicher Originalliteratur (KOSWO) entschlüsselt.

2) Der Beitrag "Wie vertraut sind Studierende der Bildungswissenschaften mit dem wissenschaftlichen Schreiben? Ein Test zu Genrewissen soll Aufschluss erbringen" aus dem Projekt Learning the Science of Education (LeScEd) stellt Ergebnisse eines Testverfahrens vor, das auf einem Modell von Genrewissen basiert.

3) Die Antwort auf die Frage, wie die geeignete Unterstützung beim Erwerb der Kompetenz des wissenschaftlichen Schreibens aussehen kann, sucht der Beitrag "Die Bedeutung des Selbst-Monitorings während des Schreibprozesses für die Entwicklung der wissenschaftlichen Schreibkompetenz" aus dem Projekt Learning the Science of Education (LeScEd).

4) Der Beitrag "Training der Kompetenz angehender Lehrkräfte zur Bewertung und argumentativen Nutzung von Evidenz" aus dem Projekt Evidence for teachers (E4teach) setzt sich mit der Frage auseinander, wie Lehrkräfte mit Forschungsbefunde aus den Bildungswissenschaften umgehen, um sie für ihre Unterrichtspraxis nutzbar zu machen.

Die Beiträge decken die Spannweite von der Erfassung der Teilkompetenzen des Lesens und Schreibens bis zum Training des gezielten Einsatzes von Teilkompetenzen in der Lehrpersonenausbildung ab. Die unterschiedliche Herangehensweise an das Thema Umgang mit wissenschaftlicher Literatur im Hochschulkontext und die verschiedenen Zielsetzungen um die Komplexität des Lesens und Schreibens von wissenschaftlichen Texten erlauben eine spannende Diskussion.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
IG-Farben-Haus  
Raum 251



## **Diagnostik von Kompetenzen Studierender im Umgang mit wissenschaftlicher Originalliteratur**

### Theoretischer Hintergrund

In den meisten Studiengängen wird von den Studierenden erwartet, dass sie sich mit wissenschaftlicher Originalliteratur auseinandersetzen. Dies dürfte für zahlreiche Studierende eine beträchtliche Herausforderung darstellen. Wissenschaftliche Originalliteratur unterscheidet sich von wissenschaftlicher Sekundärliteratur durch zahlreiche strukturelle, lexikalische, semantische und syntaktische Merkmale, verfolgt bestimmte Ziele und beruht auf besonderen epistemologischen Vorannahmen. Ihre kompetente Verarbeitung erfordert daher ein breites Repertoire von Teilkompetenzen, von denen viele weder in der Schule noch in der Universität gezielt vermittelt werden. Das Projekt „Kompetenzen Studierender im Umgang mit wissenschaftlicher Originalliteratur“ (KOSWO) zielt darauf ab, diagnostische Verfahren zu entwickeln, die der Bandbreite potenziell relevanter Kompetenzen gerecht werden. Zur theoretischen Vorstrukturierung wird zwischen zwei Arten von Verarbeitungszielen (rezeptive versus epistemische Ziele) und zwei Arten von Verarbeitungsstrategien (systematische versus heuristische Strategien) unterschieden. Die Unterscheidung rezeptiv–epistemisch bezieht sich darauf, ob es dem Leser/der Leserin darum geht, den Textinhalt zu erfassen und zu behalten (Lernen im Sinne von Wissensakkumulation) oder sich ein möglichst wahrheitsgetreues Bild der dargestellten Sachverhalte zu machen (Lernen im Sinne von Erkenntnisgewinn). Diesen Zielen dienen systematische und heuristische Strategien durch die Tiefenverarbeitung des Textinhalts bzw. die regelgeleitete Verarbeitung von Oberflächenmerkmalen. Die Kombination der beiden Zieltypen mit den beiden Strategietypen ergibt ein Schema, in dem vier Kompetenzbereiche unterschieden werden: (a) Rezeptiv-systematische Kompetenzen dienen der Anreicherung oder Strukturierung von Informationen mit dem Ziel, sie in späteren Situationen leicht abrufen zu können (z.B. Elaborieren und Organisieren). (b) Rezeptiv-heuristische Kompetenzen werden eingesetzt, um einen schnellen ersten Eindruck vom Inhalt eines Texts zu gewinnen (z.B. „scanning“, das schnelle Überfliegen eines Texts zur Lokalisation einer bestimmten Information). Derartige Prozesse sind insbesondere zur ersten Sichtung von Texten bei der Literaturrecherche unentbehrlich und werden von Wissenschaftlern/innen in solchen Situationen entsprechend häufig eingesetzt (Bazerman, 1985). (c) Epistemisch-systematische Kompetenzen werden verwendet, um die Argumentation des Texts auf ihre Stichhaltigkeit zu prüfen. Sie beinhalten die Identifikation der funktionalen Bestandteile von Argumenten wie Behauptung (claim) und Begründung (data/ground, Toulmin, 1958) sowie die Bewertung der Relevanz der angeführten Begründung für die Behauptung, für die argumentiert wird. Damit bilden epistemisch-systematische Kompetenzen eine Voraussetzung für eine rational begründete Beurteilung des Textinhalts. (d) Epistemisch-heuristische Kompetenzen zielen schließlich darauf ab, ein schnelles (vorläufiges) Urteil über die Glaubwürdigkeit des Textinhalts zu fällen (z.B. Nutzung von Quelleninformationen). Diese Prozesse sind besonders dann wichtig, wenn systematische Prozesse nicht angewendet werden können, weil relevantes domänenspezifisches Vorwissen fehlt oder motivationale und kognitive Ressourcen für eine weitergehende epistemisch-systematische Verarbeitung nicht zur Verfügung stehen (z.B. Richter, Schroeder, & Wöhrmann, 2009).

### Methode

Zur Erfassung der verschiedenen Teilkompetenzen wurde eine elektronische Testbatterie entwickelt und in einer ersten Studie mit 101 Studierenden der Psychologie geprüft (21 Frauen; 80 Männer; mittleres Alter: 24 Jahre). Da die Testbatterie 20 Subtests enthält, wurde ein Rotationsverfahren eingesetzt, bei dem Gruppen von Studierenden jeweils unterschiedliche Test-Zusammenstellungen bearbeiten. Die Studierenden bearbeiteten die Tests in ca. 90-minütigen Gruppensitzungen. In den Tests wurden die Studierenden unter anderem gebeten, möglichst schnell bestimmte Informationen in einem Text zu lokalisieren („scanning“, rezeptiv-heuristischer Kompetenzbereich); sie schätzten die Glaubwürdigkeit von Texten aus den Genres empirischer Originalartikel und populärwissenschaftlicher Text so schnell wie möglich ein (epistemisch-heuristischer Kompetenzbereich); und sie klassifizierten Textaussagen als Behauptung oder Begründung (epistemisch-systematischer Kompetenzbereich).

### Ergebnisse

Die Ergebnisse der Studie sprechen für zufriedenstellende bis sehr gute Reliabilitäten der Subtests (interne Konsistenz). Bis zum Vortrag werden weitere Ergebnisse zur kriterialen und faktoriellen Validität aus der ersten Erhebungsphase einer Längsschnittstudie vorliegen. Darüber hinaus deuten einige Befunde der Studie auf Defizite im Umgang mit wissenschaftlicher Literatur hin, die möglicherweise auf Fehlkonzepten beruhen. So scheinen viele Studierende ein sehr enges Verständnis wissenschaftlicher Originalliteratur aufzuweisen, das z.B. theoretische Beiträge ausschließt. Außerdem scheinen sie Schwierigkeiten zu haben, zwischen subjektiven Meinungen und argumentativ begründeten Behauptungen zu unterscheiden.

Bazerman, C. (1985). Physicists reading physics: Schema-laden purposes and purpose-laden schema. *Written Communication*, 2, 3–23.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Richter, T., Schroeder, S. & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, 96, 538–558.



Alexandra Winter (Universität Freiburg), Kristin Wäschle (Universität Freiburg), Jörg Wittwer (Universität Freiburg),  
Rainer Watermann (Freie Universität Berlin), Matthias Nückles (Universität Freiburg)

## **Wie vertraut sind Studierende der Bildungswissenschaften mit dem wissenschaftlichen Schreiben? Ein Test zu Genrewissen soll Aufschluss erbringen**

### Theoretischer Hintergrund und Fragestellung

Als fächerübergreifende generische Kompetenz ist der Erwerb von Fähigkeiten im wissenschaftlichen Schreiben erklärtes Ziel an deutschen Hochschulen, da das Schreiben im wissenschaftlichen Diskurs eine bedeutende Funktion einnimmt. Eine zentrale Voraussetzung für das wissenschaftliche Schreiben ist — in Abgrenzung zum inhaltsbezogenen Wissen — das sogenannte Genrewissen. Dieses kann das Verfassen wissenschaftlicher Texte erheblich erleichtern, indem es Autorinnen und Autoren Schemata zur Orientierung bei der Textproduktion bietet.

Unter Genrewissen soll hier sowohl Wissen um den generellen Aufbau als auch differenziertes Wissen über einzelne rhetorische Elemente eines Genres verstanden werden. Zudem besitzt jedes Genre eine übergeordnete kommunikative Funktion, deren rhetorische Ziele in den untergeordneten Abschnitten durch sogenannte moves realisiert werden (Goldman & Bisanz, 2002; Swales, 1990). So umfasst unser Modell von Genrewissen, das sich an einem von Becker-Mrotzek und Schindler (2007) entwickelten Schreibkompetenzmodell orientiert, bislang die Komponenten eines a) übergeordneten kommunikativen Ziels des jeweiligen Genres, b) dessen untergeordnete kommunikative Ziele auf Abschnittsebene sowie c) Stil und sprachliche Merkmale auf drei Wissens Ebenen, die sich als 1) explizierbares rhetorisches Wissen, 2) diagnostisch anwendbares rhetorisches Wissen und 3) rhetorisches Wissen in der Textproduktion bezeichnen lassen.

Da die Wissenschaft sich unterschiedlicher Genres bedient, konzentrierten wir unsere Testentwicklung auf das Genre des empirischen Forschungsartikels. Unter allen Genres kann der empirische Forschungsartikel als Schlüsselgenre betrachtet werden, weil er quantitativ in Hinblick auf die Anzahl der Veröffentlichungen dominierend ist. Andererseits hat der empirische Forschungsartikel auch qualitativ eine besondere Bedeutung, da er bezüglich seines Aufbaus mit weiteren Genres der Forschungsgemeinschaft wie beispielsweise Abstracts in enger Beziehung (Swales, 1990) steht.

Trotz Konventionen und zahlreichen Ratgebern, wie „gute“ wissenschaftliche Publikationen allgemein und ganz speziell das anspruchsvolle Genre des empirischen Forschungsartikels zu verfassen seien (siehe APA, 2009; Bem, 2003), werden die Kompetenzen von Studierenden beim wissenschaftlichen Schreiben relativ selten empirisch untersucht (Kelly & Takao, 2002; Myers, 1990; Roth et al., 1996). Bisher existieren zu ihrer Messung auch keine Testverfahren. Das Ziel unseres Forschungsprojekts besteht deshalb darin, ein auf dem Genrewissen basierendes Testverfahren für das wissenschaftliche Schreiben in den bildungswissenschaftlichen Studiengängen zu entwickeln und zu überprüfen.

### Methode

Bislang bearbeiteten 310 Studierende sowie 76 Promovierende der Bildungswissenschaften den vollständigen, alle Items umfassenden Test. In das Testverfahren gingen nach einer Pilotierungsphase 38 Items ein, wovon 25 Items inhaltlich zu über- und untergeordneten kommunikativen Zielen, elf Items zu Stil und Sprache wissenschaftlichen Schreibens, ein Item zu statistischem Wissen und ein Item zu Konventionen wissenschaftlichen Schreibens zugeordnet werden können. In die Analyse konnten N = 370 Fälle einbezogen werden. Unter der Annahme des Universitätsdimensionalen Rasch-Modells wurden die Testitems mit ConQuest auf ihre psychometrische Qualität geprüft.

### Ergebnisse und Diskussion

In den ersten IRT-Analysen zeigt sich, dass übergeordnete Kennwerte wie Varianz der Personenparameter und der ausgegebene Reliabilitätsparameter mit 0.848 als gut zu bewerten sind. Unter Berücksichtigung der gewichteten Modellanpassungswerte ( weighted Mean Squares ), akzeptabler Infit-Spannen (wMNSQ-Grenzen zwischen 0.8 und 1.2) und geeigneter Trennschärfen ( $\geq 3.0$ ) verbleiben 23 Items.

Während die hohe Tendenz zu Überfit-Werten noch genauer zu betrachten sein wird, zeichnen sich in einem vorläufigen Standard-Setting-Verfahren drei unterschiedliche Kompetenzniveaus ab: Niveau I ist durch wenig explizierbares rhetorisches Wissen zu Schema und Sektionen des empirischen Forschungsartikels gekennzeichnet. In Höhe dieses Niveaus, auf dem sich überwiegend Lehramtsstudierende befinden, werden lediglich Aufgabenanforderungen im Bereich des diagnostischen rhetorischen Wissens und der Rezeption von Forschungsartikeln besser bewältigt. Im Gegensatz dazu ist Niveau III, auf dem sich überwiegend Promovierende befinden, durch globales hohes explizierbares und diagnostisches rhetorisches Wissen insbesondere zu einzelnen Sektionen des empirischen Forschungsartikels gekennzeichnet. Niveau II reiht sich zwischen die eben genannten ein.

Demzufolge scheint das Testverfahren zur Differenzierung unterschiedlicher Kompetenzlevels geeignet. In weiterführenden Studien soll untersucht werden, inwiefern der Test auch prädiktiv hinsichtlich qualitativer Unterschiede bei der Produktion wissenschaftlicher Abstracts sein kann.

American Psychological Association (2009). Publication manual of the American Psychological Association (6th ed.). Washington, DC: APA.

Becker-Mrotzek, M. & Schindler, K. (2007). Schreibkompetenz modellieren. In M. Becker-Mrotzek & K. Schindler (Hrsg.), Texte Schreiben (S. 7-26). Köln: Gilles & Francke.

Bem, D. J. (2003). Writing the empirical journal article. In J. M. Darley, M. P. Zanna, H. L. Roediger III. (Eds.), The complete academic: A career guide (2nd ed.) (pp. 185-219). Washington, DC, US: American Psychological Association.

Goldman, S. R. & Bisanz, G. L. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León & A. C. Graesser (Eds.), The psychology of science text comprehension (pp. 417-436). Mahwah, NJ: Erlbaum.

Kelly, G. J. & Takao, A. (2002). Epistemic levels in argument. An analysis of University oceanography students' use of evidence in writing. Science Education, 86, 314-342.

Myers, G. (1990). Writing biology: Texts in the social construction of scientific knowledge. Madison, WI: University of Wisconsin Press. American Psychological Association.

Roth, W. M., McGinn, M. K. & Bowen, G. M. (1996). Applications of science and technology studies. Effecting change in science education. Science, Technology & Human Values, 21, 454-484

Swales, J. (1990). Genre analysis: English in academic and research settings. New York: Cambridge University Press.

GEBF TAGUNG  
2014

D14 b

Symposiums  
beitrag

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
IG-Farben-Haus  
Raum 251



## **Die Bedeutung des Selbst-Monitorings während des Schreibprozesses für die Entwicklung der wissenschaftlichen Schreibkompetenz**

### Theoretischer Hintergrund

Der Austausch über Forschung nimmt in den Bildungswissenschaften eine zentrale Position ein und soll Teil der Ausbildung sein (Wissenschaftsrat, 2000; KMK, 2005). Ein wichtiges Medium der Wissenschaftskommunikation ist der wissenschaftliche Artikel. Die Verschriftlichung des Forschungsvorgangs als ein wissenschaftlicher Artikel kann als die anspruchsvollste Form der Wissenschaftskommunikation angesehen werden, denn es werden dabei verschiedene Perspektiven zusammengeführt: die Botschaften der verwendeten Texte, die Intention des Autors und das Verständnis des Lesers (Kellogg, 2008).

Die Berücksichtigung dieser drei Perspektiven macht das Schreiben von wissenschaftlichen Texten anspruchsvoll und bedarf angemessener Unterstützung. Welche Art der Unterstützung bei der Entwicklung der Kompetenz des wissenschaftlichen Schreibens ist für Studierende geeignet? Auf der einen Seite geben zahlreiche Studien zum Erwerb der Schreibkompetenz auf Schulniveau Hinweise auf unterstützende Strategien, die sich im Grad der Anleitung unterscheiden (s. für einen Überblick Graham & Perin, 2007). Auf der anderen Seite handelt es sich bei Studierenden um erfahrene Lernende. Von ihnen wird erwartet, dass sie in ihrer Ausbildung zunehmend selbstreguliert lernen (Bereiter & Scardamalia, 1987; Graham & Harris, 2000; Hayes & Flowers, 1980; Zimmerman & Kitsantas, 1999).

Hayes und Flowers (1980) haben den Prozess des Schreibens unter dem Aspekt der Selbstregulation als Modell mit den Phasen des Planens, Umsetzens und Überarbeitens eines Textes konzipiert, auf die externe Faktoren wie die Textadressaten und interne Faktoren wie das Vorwissen der schreibenden Person einwirken. Die Fertigkeit zur Selbstregulation bedingt die Entwicklung der Schreibkompetenz (Graham & Harris, 2000). Von den Selbstregulationsstrategien gilt besonders das Selbst-Monitoring als wichtig für den Erwerb der Schreibkompetenz (Cresswell, 2000; Graham, Harris, & Mason, 2005; Winne & Hadwin, 1998; Xing, 2004). Wir definieren Selbst-Monitoring im Schreibprozess als das Auswerten des augenblicklich geschriebenen Textes: Was wurde geschrieben, was soll geschrieben werden, erreicht die Intention des Autors den Leser? (Hacker, 1994).

### Fragestellung

Die Untersuchung vergleicht verschiedene Formen der Unterstützung, um den Effekt auf die Kompetenzentwicklung durch Selbst-Monitoring bei kurzfristigen Interventionen mit Effekten durch oberflächenorientierte Strategien (Textaufbau, Textzusammenfassung) vergleichen zu können. Dabei wurden die Strategien zur Strukturierung des Textes und zur Textzusammenfassung als oberflächenorientierte und zur Anwendung des Selbst-Monitoring als tiefenorientierte Elaboration betrachtet.

### Methode

Wir führten eine experimentelle Interventionsstudie im Prätest-Posttest-Kontrollgruppen-Design durch, an der 60 Psychologiestudierende in drei gleichgroßen Gruppen als eine Experimental- und zwei Kontrollgruppen teilnahmen. Jede Gruppe wurde über die Anwendung des Aufbaus eines Forschungsartikels instruiert. Zusätzlich erhielt die Experimentalgruppe als Unterstützung eine Instruktion zur Anwendung einer Selbstregulierungsstrategie, die auf den Schreibprozess ausgerichtet ist, das Selbst-Monitoring. Eine Kontrollgruppe erhielt keine weitere Unterstützung. Die andere Kontrollgruppe erhielt als weitere Unterstützung eine Instruktion, wie Texte zusammengefasst werden können. Alle Teilnehmenden schrieben ein Abstract zu einem Forschungsartikel.

Die Abstracts wurden mit einer Inhaltsanalyse nach Mayring (2010) über ein hochinferentes Rating ausgewertet und quantifiziert. Für die Messung des Vorwissens und des Lernerfolgs haben wir selbst-entwickelte Skalen eingesetzt.

### Ergebnisse und Diskussion

Die Auswertung des Lernerfolgstests ergab, dass die Selbst-Monitoring-Gruppe beide Kontrollgruppen bei komplexen Schreibaufgaben, die die Anwendung von Textstrukturwissen erforderten, übertraf ( $p < .05$ ). Hingegen traten bei der Auswertung der Abstracts keine signifikanten Unterschiede zwischen den einzelnen Gruppen auf. Unsere Ergebnisse weisen darauf hin, dass es sinnvoll ist, die Studierenden bei der Anwendung von Selbst-Monitoring-Strategien zu unterstützen, um die Entwicklung der Kompetenz des wissenschaftlichen Schreibens zu fördern.

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, NJ: Erlbaum.

Cresswell, A. (2000). Self-monitoring in student writing: Developing learner responsibility. *ELT Journal*, 54(3), 235–244.

Graham, S., & Harris, K.R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist*, 35, 3–12.

Graham, S., & Perin, Dolores (2007): Writing next. effective strategies to improve writing of adolescents in middle and high schools. A report to Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.

Hacker, D.J. (1994). Comprehension monitoring as a writing process. In E.C. Butterfield (Ed.), *Children's writing: Toward a process theory of the development of skilled writing*. Greenwich, CN: JAI Press.

Hayes, J.R., & Flower, L.S. (1980). Identifying the organization of writing processes. In L.W. Gregg & R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.

Kellogg, R. (2008). Training writing skills: A cognitive developmental perspective. *Journal of writing research*, 1, 1-26.

Myers, G. (1991). Lexical cohesion and specialized knowledge in science and popular science texts. *Discourse Processes*, 14 (1), 1-26.



## **Training der Kompetenz angehender und aktiver Lehrkräfte zur Bewertung und argumentativen Nutzung von Evidenz**

### Theoretischer Hintergrund

Das Prinzip des evidenzbasierten Handelns ist nicht mehr ausschließlich Thema der Medizin, sondern spielt auch im Bildungswesen eine immer größere Rolle (Wiseman, 2010; Sackett 1996). Der Praxistransfer hat jedoch bisher in vielen Bereichen nicht, oder nur unzureichend, stattgefunden. Gerade Lehrkräfte an Schulen begründen ihr Handeln kaum auf Grundlage wissenschaftlicher Evidenz (Williams & Coles, 2007). Gleichwohl wird dies in den Standards für Lehrerbildung gefordert (KMK, 2004).

Im Anschluss an diese Entwicklungen werden in dem Projekt E4teach die Fragen adressiert, wie angehende und aktiver Lehrkräfte zur Bewertung und argumentativen Nutzung von Evidenz in der Lage sind und wie sich diese Kompetenz fördern lässt. Auf Basis eines Kompetenzmodells, das u.a. an handlungstheoretische Konzepte evidenzbasierte Praxis anknüpft (Ramos et al., 2003) wird hierzu ein Erhebungsverfahren entwickelt, mit dem man in der Lage ist die entsprechenden Kompetenzen der zukünftigen Lehrkräfte abzubilden. Zudem wird auf ein Training erstellt und dessen Wirksamkeit experimentell überprüft. Der vorliegende Beitrag beschäftigt sich mit dieser zweiten Zielsetzung. Bewährte Trainingskonzepte der evidenzbasierten Medizin können zu diesem Zwecke als Grundlage herangezogen werden. Aufgrund der Unterschiede beider Domänen ist es aber unabdingbar die Konzepte nicht nur zu übernehmen, sondern sie für den Kontext Bildungswissenschaft zu adaptieren (Finkel, 2003). Eine zentrale Ergänzung liegt darin, zusätzlich zum fachlichen Input, alltagsnahe Fallbeispiele aus der Bildungspraxis sowie Phasen praktischer Anwendung in Gruppenarbeiten zu implementieren.

Theoretische Grundlagen für das Training zum kompetenten Umgang mit Evidenz waren, neben den erwähnten handlungstheoretischen Konzepten, auch didaktische Designs wie beispielsweise das Vier-Komponenten-Instruktionsdesign-Modell (van Merriënboer et al, 2002). Dieses Training beginnt induktiv mit einem problemorientierten Einstieg anhand einer Situationsvignette. Darauf folgt eine deduktiv gestaltete Einführung in die Kernprinzipien der evidenzbasierten Praxis. Dieses Wissen wenden die Teilnehmenden nun selbständig unter abnehmender tutorieller Unterstützung anhand von Situationsvignetten, abwechselnd aus dem Gesundheits- und dem bildungswissenschaftlichen Bereich, an.

### Fragestellung

Dieser Beitrag geht der Frage nach, ob der kompetente Umgang von Lehramtsstudierenden mit bildungswissenschaftlicher Evidenz durch eine gezielte Intervention gefördert werden kann und welchen Einfluss Hintergrundvariablen wie wissenschaftstheoretisches Vorwissen, Fächerkombination, Überzeugungen zur Nutzung von Evidenz, sowie Ambiguitätstoleranz ausüben.

### Methode

Die Überprüfung der Wirksamkeit des Trainings erfolgt in einer aktuell laufenden Pilotstudie mit insgesamt N =60 Lehramtsstudierenden. Das Design umfasst erstens ein 2x2 Experiment (randomisierter Faktor Training vs. (Warte-)Kontrollgruppe, beobachteter Faktor Fächerkombination: natur- vs. geisteswissenschaftlich). Zweitens werden über Fragebögen und Interviews im Sinn von „Cognitive Labs“ Prozess- und Akzeptanzdaten erfasst, die eine zusätzliche Einschätzung der ablaufenden Lernprozesse und Wirkungen des Trainings erlauben. Als relevante Hintergrundvariablen werden epistemologische Überzeugungen, Überzeugungen zur Nutzung von Evidenz und die Ambiguitätstoleranz erfasst.

Die Messung der abhängigen Variablen erfolgt mit dem im Projekt entwickelten Erhebungsinstrument, das die beiden Kompetenzfacetten der Bewertung und argumentativen Nutzung von Evidenz erfasst. Die Lernenden bearbeiten hierzu bildungswissenschaftliche Problemfälle in Vignettenform, zu denen sie verschiedene Evidenzquellen erhalten, die sie bewerten und in eine schriftliche Lösung integrieren sollen.

### Ergebnisse

Die Durchführung der Studie läuft aktuell, so dass die Ergebnisse auf der Konferenz vorgestellt werden können. Auf Basis bisher verfügbarer Befunde (Finkel, 2003) erwarten wir, dass sich durch die Intervention nicht nur die Kompetenzen, sondern auch die Überzeugungen zum Umgang mit Evidenz positiv verändern. Zudem gehen wir davon aus, dass die Überzeugungen zur Nutzung von Evidenz, sowie die Ambiguitätstoleranz der Studierenden die Wirksamkeit des Trainings beeinflussen.

Finkel, M. L., Brown, H.-A., Gerber, L. M., & Supino, P. G. (2003). Teaching evidence-based medicine to medical students. *Medical Teacher*, 25(2), 202-204.

Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004). Bonn: KMK, verfügbar unter [http://www.kmk.org/fileadmin/ver-oeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/ver-oeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf)

Ramos, K. D., Schafer, S. & Tracz, S. M. (2003). Validation of the Fresno Test of competence in evidence based medicine. *BMJ*, 326, 319-321.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, (312), 71-72.

Van Merriënboer, J. J. G., Clark, R. E. & de Crook, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology Research and Development*, 50, 39-64.

Williams, D. & Coles, L. (2007). Evidence-based practice in teaching: an information perspective. *Journal of Documentation*, 63(6), 812-835.

Wiseman, A. W. (2010). The uses of evidence for educational policymaking: Global contexts and international trends. *Review of Research in Education*, 34, 1-24.





Katharina Wolf (DIPF und IDeA Zentrum), Marlis Abrie-Kuhn (DIPF und IDeA Zentrum), Silke Hertel (IDeA Zentrum und Universität Heidelberg), Andreas Eickhorst (Institut für Psychosomatische Kooperationsforschung und Familientherapie, DJI (Deutsches Jugendinstitut)), Manfred Cierpka (Institut für Psychosomatische Kooperationsforschung und Familientherapie)

## ***Feinfühlig Interaktionsgestaltung und Gestaltung von Lernumgebungen in der frühen Kindheit***

**Chair(s):** Katharina Wolf (DIPF und IDeA Zentrum)

Theoretischer Hintergrund: In den ersten Lebensjahren sind Eltern die wichtigsten Bezugspersonen für ihr Kind und nehmen in Erziehungs- und Bildungsprozessen eine zentrale Rolle ein (Wild & Lorenz, 2009; Hurrelmann, 2011). In Theorie und Forschung besteht weitestgehend Konsens darüber, dass die kognitive und die sozio-emotionale Entwicklung von Kindern von der Qualität der Lernumgebung und dem elterlichen feinfühligem Unterstützungsverhalten beeinflusst wird (Bradley & Caldwell, 1987; Foster, Lambert, Abbott-Shim, McCarty & France, 2005; Son & Morrison, 2010; Nicholson, Lucas, Berthelsen & Wake, 2010; Sylva, Melhuish, Sammons, Siraj-Blatchford & Taggart, 2012; Cierpka, 2012). Zur feinfühligem Lernbegleitung sind daher elterliche Kompetenzen auf unterschiedlichen Ebenen gefordert. Der aktuell auf politischer und wissenschaftlicher Ebene geführte Diskurs beschäftigt sich u.a. mit der Frage, wie Eltern hierbei unterstützt werden können. Im Fokus unserer Studie stehen daher Strategien zur Unterstützung (1) der kognitiven Entwicklung (Scaffolding) sowie (2) der feinfühligem Interaktion von Eltern und Kind in häuslichen Lernumgebungen (Ainsworth, Bell & Stayton, 1974; Van de Pol, Volmann & Beishuizen, 2010).

Fragestellungen: Unser Beitrag thematisiert folgende Fragestellungen unter Berücksichtigung quantitativer und qualitativer Forschungsperspektiven: (1) Können Eltern im Rahmen eines Basistrainings Scaffolding-Kompetenzen erwerben? (2) Können durch ein kombiniertes Training Strategien zur feinfühligem Beziehungsgestaltung gefördert werden? (3) Wird durch die Anreicherung des Basistrainings mit Strategien zur feinfühligem Beziehungsgestaltung ein Mehrwert erreicht? (4) Wie schätzen die Eltern die Qualität bzw. den Nutzen der Trainings ein?

Methode: Die Studie basiert auf einem quasi-experimentell angelegten Design mit randomisierter Zuteilung von 58 Eltern-Kind-Dyaden zu drei Treatmentgruppen: Wartekontrollgruppe, Basistraining (Scaffolding) und kombiniertes Training (Scaffolding und feinfühligem Beziehungsgestaltung). Zielgruppe waren Eltern von Kindern im Alter von 18-36 Monaten. Die Wirksamkeit der Trainings wurde auf Eltern- und Kinderebene durch standardisierte Fragebögen, Wissenstests, Fallvignetten und teilstandardisierte Videoaufzeichnungen erfasst. Zudem wurde eine Entwicklungsdiagnostik mit den „Bayley Scales of Infant Development II“ (Bayley, 1993) zur Erhebung der Eingangsvoraussetzungen der Kinder durchgeführt. Die Auswertung der Videoaufzeichnungen erfolgt mit den „Emotional Availability Scales“ (Biringen, 2008) und einem vom Projektteam entwickelten Ratingverfahren. Um die Wirksamkeit der Trainings und den Transfer der Strategien in den Familienalltag zu überprüfen werden teilstandardisierte Leitfadeninterviews durchgeführt.

Ergebnisse und Diskussion: Die Qualität des Trainings wurde von den Teilnehmerinnen und Teilnehmern als „hoch“ oder „recht hoch“ ( $M=1.71$ ,  $range=1-6$ ,  $SD=.73$ ) bewertet und es konnte ein Wissensanstieg im Hinblick auf die Gestaltung häuslicher Lernumgebung ( $M=3.39$ ,  $range=1-4$ ,  $SD=.66$ ) festgestellt werden. Darüber hinaus planen sie die Strategien auch in der Zukunft anzuwenden ( $M=3.36$ ,  $range=1-4$ ,  $SD=.65$ ). Die Analysen ergaben statistisch signifikante Zeit x Treatment Interaktionen für den „wahrgenommene Kompetenzzuwachs“ im Bereich Scaffolding. Teilnehmende des Basistrainings und des kombinierten Trainings berichten über einen signifikant stärkeren Kompetenzzuwachs von Prä- zu Posttest als Teilnehmende der Kontrollgruppe. Dieser Befund lässt darauf schließen, dass die Scaffolding-Kompetenzen in einem kombinierten Training genauso gefördert werden können wie in einem Scaffolding-Basistraining. Hinsichtlich der Strategien zur feinfühligem Beziehungsgestaltung weisen erste Ergebnisse darauf hin, dass die Teilnehmerinnen und Teilnehmer des kombinierten Trainings zwar einen kleinen Anstieg bezüglich ihrer Selbsteinschätzung zeigen, dieser lässt sich jedoch nicht statistisch absichern. Dieser unerwartete Befund lässt sich ggf. auf ein response-shift-bias Phänomen zurückführen (Howard, Schmeck, & Bray, 1979; Howard, 1980; Moore & Tananis, 2009). Die Auswertung der Videoaufzeichnungen und der Leitfadeninterviews soll hierzu mehr Aufschluss bringen.

Ainsworth, M. D. S., Bell, S. M. & Stayton, D. (1974). Infant-mother attachment and social development. In M. P. Richards (Ed.), *The introduction of the child into a social world* (pp. 99-135). London: Cambridge University Press.

Bayley, N. (1993). *Manual for the Bayley Scales of Infant Development 2nd Edition*. San Antonio, TX: The Psychological Corporation.

Bradley, R. H. & Caldwell, B. M. (1987). Early environment and cognitive competence: The Little Rock Study. *Early Child Development and Care*, 27, 307-341.

Biringen, Z. (2008). *Emotional Availability (EA) Scales Manual*, 4th Edition. Published by emotionalavailability.com, PO Box, 3525 Boulder, Colorado 80307.

Cierpka, M. (2012). *Frühe Kindheit 0-3 Jahre. Beratung und Psychotherapie für Eltern mit Säuglingen und Kleinkindern*. Berlin, Heidelberg, Springer.

Foster, M. A., Lambert, R., Abbott-Shim, M., McCarty, F. & France, S. (2005). A model of home learning environment and social risk factors in relation to children's emergent literacy and social outcomes. *Early Childhood Research Quarterly*, 20, 13-36.

Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review*, 4, 93-106.

Howard, G. S., Schmeck, R. R. & Bary, J. H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement*, 16, 129-135.

Hurrelmann, K. (2011). Familie und Bildungserfolg: Anmerkungen zur Expertise. In A. Sliwka & S. Frank (Eds.), *Familie schafft Chancen. Warum das Zuhause für den Bildungserfolg entscheidend ist und wie Eltern bei der Lernbegleitung ihrer Kinder gestärkt werden können* (pp. 20-25). Vodafone Stiftung Deutschland.

Moore, D. & Tananis, C. A. (2009). Measuring change in a short-term educational program using a retrospective pretest design. *American Journal of Evaluation*, 30 (2), 189-202.

Nicholson, J. M., Lucas, N., Berthelsen, D. & Wake, M. (2010). Socioeconomic inequality profiles in physical and developmental health from 0-7 years: Australian National Study. *Journal of Epidemiology and Community Health*, 66 (1), 81-87.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2012). *Effective pre-school, primary and secondary education 3-14 project (EPPSE 3-14). Final report from the key stage 3 phase: Influences on students' development from age 11 – 14*. London: Institute of Education.

Son, S.-H. & Morrison, F. J. (2010). The nature and impact of changes in home learning environment on development of language and academic skills in preschool children. *Developmental Psychology*, 46 (5), 1103-1118.

Wild, E. & Lorenz, F. (2009). Familie. In E. Wild & J. Möller (Eds.), *Pädagogische Psychologie* (pp. 235-260). Berlin: Springer.

Van de Pol, J., Volmann, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: a decade of research. *Education Psychological Review*, 22, 271-296.





## **Kompensatorische Effekte von frühkindlicher Bildung (FBBE) bzw. ganztägiger Bildung hinsichtlich geringer Förderung in der Familie?**

Die Nutzung der sekundären Sozialisationsinstanz Schule durch die Kinder hängt vom sozialen, kulturellen und ökonomischen Kapital der Familie ab (Bourdieu, 1983). Bereits bei Kindergarten- und Schuleintritt unterscheiden sich die Kinder in mathematischen Kompetenzbereichen, insbesondere weil sie in ihren Familien unterschiedliche Entwicklungsmöglichkeiten erfahren. Verschiedene Studien zeigen, dass es der Schule nicht im erhofften Umfang gelingt, die primären Disparitäten zu kompensieren (Ditton & Krüsken, 2009). Verbreitet wird deshalb gefordert, dass sich die Bildungsinstitutionen und die Familie an die veränderten gesellschaftlichen Rahmenbedingungen anpassen müssen, um die Kinder auch in Zukunft auf ihrem Weg des Erwachsenwerdens unterstützen zu können. In den letzten Jahren wurde eine stärkere Gewichtung neben der formalen auch der nicht-formalen und informellen Bildung postuliert. In diesem Kontext ist der Ausbau von FBBE und von Ganztagschulen bzw. Tagesschulen in der Schweiz in den letzten Jahren zu betrachten, in den diesbezüglich große Erwartungen gesetzt werden. Zentrale Anliegen sind dabei u.a. die Förderung von kognitiven Kompetenzen aller SchülerInnen sowie die Einlösung der Forderung nach Chancengleichheit im Bildungssystem für alle (Schüpbach et al., 2013).

Die meisten vorliegenden Studien zu den Wirkungen der FBBE untersuchen Kinder, die institutionell betreut werden, also Angebote im formellen Bereich. Für eine Förderung der kognitiven Kompetenzen ist sowohl die Qualität als auch die Quantität von Bedeutung. In diesem Paper wird auf die Quantität, die Zeitdauer fokussiert. Bei einem länger dauernden Besuch zeigen sich bei den meisten institutionellen Angebotsformen positive Effekte auf die kognitive Entwicklung und auf die Schulleistungen in Mathematik. Die bessere Entwicklung liegt während der Vorschulzeit und teilweise noch während der Schulzeit vor (vgl. Roßbach, 2005). Weitgehend positive Effekte findet man insbesondere auch für sozial benachteiligte Kinder in neueren Studien (Sammons et al., 2004). Hinsichtlich der Wirksamkeit der Ganztagschule zeigen die Ergebnisse der deutschen Studie StEG, dass erst unter Berücksichtigung der Intensität eine Wirkung auf die Schulnoten feststellbar ist (Kuhn & Fischer, 2011). Eine Vielzahl an Befunden aus US-amerikanischen Studien zeigt auf, dass Kinder, die Angebote nutzen bessere Ergebnisse in Leistungstests erzielen (Fredericks & Eccles, 2006). Im Weiteren zeigt sich, dass insbesondere Kinder aus Familien mit einem niedrigeren sozioökonomischen Status von einer Teilnahme an spezifischen Interventionsprogrammen profitieren und bessere Leistungen in Mathematik erzielen (Lauer et al., 2006).

Folgende Hypothesen lassen sich ableiten: (H1) Die FBBE, formeller und informeller Bereich, wirkt sich positiv auf die Mathematikleistungsentwicklung aus. (H2) Der FBBE gelingt es, im Hinblick auf die Mathematikleistungsentwicklung eine kompensatorische Wirkung hinsichtlich einer geringen familialen Entwicklungsförderung zu erzielen. (H3) Eine intensive Nutzung des Tagesschulangebots wirkt sich positiv auf die Mathematikleistungsentwicklung aus. (H4) Dem Tagesschulangebot gelingt es, im Hinblick auf die Mathematikleistungsentwicklung eine kompensatorische Wirkung hinsichtlich einer geringen familialen Entwicklungsförderung zu erzielen.

Die Hypothesen werden im Rahmen einer Längsschnittstudie geprüft. Die Stichprobe, anhand derer die Analysen durchgeführt wurden, umfasst N= 295 SchülerInnen, aus 43 Primarschulklassen in 35 Schulen aus 11 Kantonen. Die Hypothesen wurden mittels linearer latenter Wachstumsmodelle unter Berücksichtigung von drei Messzeitpunkten geprüft. Damit wurde der Entwicklungsstand am Ende des 1. Schuljahres (Intercept) sowie die Entwicklung bis zum 3. Schuljahr (Slope) modelliert.

Die Ergebnisse zeigen insgesamt, dass Kinder, die vor Schuleintritt über eine lange Zeitdauer FBBE nutzen, sich positiver in ihrer Schulleistung in Mathematik zwischen Ende 1. und Ende 3. Schuljahr entwickeln. Wird die FBBE über eine lange Zeitdauer genutzt, gelingt es dieser eine kompensatorische Wirkung hinsichtlich einer geringen familialen Entwicklungsförderung zu erzielen. Kinder, die ein Tagesschulangebot intensiv nutzen, entwickeln sich in ihrer Mathematikleistung besser als solche, die keines nutzen. Es gibt jedoch keine Annäherung der Tagesschulkinder mit geringer familialer Förderung an diejenigen mit höherer familialer Förderung. In diesem Sinne ist keine kompensatorische Wirkung der Tagesschule bei dieser Kindergruppe zu verzeichnen.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183-198). Göttingen: Schwartz.

Ditton, H. & Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *Journal für Bildungswissenschaft Online*, 1, 1, 33–61.

Fredericks, J. A. & Eccles, J. S. (2006). Is extracurricular participation associated with beneficial outcomes? *Concurrent and Longitudinal Relations. Developmental Psychology*, 42, 698-713.

Kuhn, H.-P. & Fischer, N. (2011). Entwicklung der Schulnoten in der Ganztagschule. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen* (S. 207–226). Weinheim: VS Verlag für Sozialwissenschaften.

Lauer, P., Akiba, M., Wilkerson, S. B., Aphorpe, H. S., Snow, D. & Martin-Glenn, M. L. (2006). Out-Of-School-Time Programs: Meta-Analysis of Effects for At-Risk Students. *Review of Educational Research*, 76 (2), 275–313.

Roßbach, H.-G. (2006). Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien. In Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Hrsg.), *Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren* (S. 55-174). München: Verlag Deutsches Jugendinstitut.

Sammons (2004). The effective provision of pre-school education (EPPE) Project. The continuing pre-school education at age 7 years. University of London, Institut of Education.

Schüpbach, M.; Herzog, W.; Ignaczewska, J. (2013). Entwicklung der Mathematikleistung von Ganztagschulkindern: Kompensatorische Wirkung von Ganztagschule? *Zeitschrift für Pädagogische Psychologie*, 27 (3), p. 157-168.



## Alltagsgespräche im Kindergarten – Gelegenheitsstrukturen für den Erwerb bildungssprachlicher Fähigkeiten

Der Besuch einer qualitativ guten vorschulischen Bildungseinrichtung kann einen Beitrag zur Verbesserung der Chancengleichheit leisten (Bos et al., 2003, Sammons et al., 2008). Über die Prozessqualität pädagogischer Interaktionen in vorschulischen Einrichtungen ist bisher aber wenig bekannt. Im Projekt "Prozesse der Sprachförderung im Kindergarten" werden Alltagsgespräche zwischen Lehrpersonen und Kindern gefilmt und sequenzanalytisch ausgewertet, um die Potenziale informeller Lehr-Lernsituationen für den Erwerb und die Förderung bildungssprachlicher Fähigkeiten zu erkunden. Die Gespräche werden als Ko-Konstruktionen bzw. als gemeinsamer, schrittweiser Handlungsvollzug (Schütz, 1932 [1974]) verstanden und untersucht. Ziel ist es, Grundlagen für eine "rationale" Sprachförderung zu erarbeiten, die Bildungsungleichheit nicht reproduziert, sondern zu ihrem Abbau beiträgt (Bourdieu & Passeron, 1971). Den soziologischen Rahmen bildet Pierre Bourdieus Theorie von Habitus und Feld (Bourdieu, 2001). Beim alltäglichen Handeln erwerben Kinder sprachliche Praktiken und Fähigkeiten mit spezifischen Merkmalen auf allen Ebenen des Sprachsystems und des Sprachgebrauchs. Im Feld der Schule erfahren gewisse Praktiken mehr Anerkennung als andere, wodurch Kinder, die mit der "schulischen Form" (Thévenaz-Christen, 2005) unvertraut sind, implizit benachteiligt werden (Käsler 2005). Aus linguistischer Sicht unterscheidet sich die schulische Bildungssprache in verschiedener Weise von der Alltagssprache: Sie wird in kollektiven Interaktionen mit "öffentlichem" Charakter verwendet (Erard & Schneuwly, 2005), ist in spezifische Handlungsmuster eingebettet (van Kleeck, 2011), konzeptionell schriftlich strukturiert (Koch & Österreicher, 1994) und nicht nur Medium, sondern auch Gegenstand des Unterrichts (Vygotzky, 1934[1986]). Lerntheoretisch wird der Erwerb bildungssprachlicher Fähigkeiten als Modelllernen beim gemeinsamen Handeln mit kompetenten Anderen konzipiert (Tomasello, 1999; Vygotzky, 1934[1986]). Dabei spielt die Herstellung und Aufrechterhaltung geteilter Aufmerksamkeit ("sustained shared thinking", Gutiérrez & Rogoff, 2003) eine Schlüsselrolle.

Die Analyse von Lehr-Lernprozessen im Hinblick auf die situierte, multimodale und sequenzielle Herstellung intersubjektiven Sinns bildet den methodologischen Kern des Projekts ProSpiK: Es geht darum, die interaktive Herstellung der Vollzugswirklichkeit (Garfinkel, 1967) bei Alltagsgesprächen im Kindergarten zu untersuchen. Dabei werden folgende Fragen bearbeitet:

1. Welche kommunikativen Formen eines "schulnahen" Sprachgebrauchs sind im Kindergarten anzutreffen?
2. Wie werden sie von den Lehrpersonen und Kindern interaktiv hergestellt?
3. (Wie) wird im Rahmen solcher Interaktionen soziale Ungleichheit von den Lehrpersonen und Kindern hergestellt, aufrecht erhalten oder abgebaut?

In Anbetracht der komplexen, noch wenig bekannten Wirklichkeit des Kindergartens wurde ein offener, verstehender und theoriegenerierender Zugang gewählt: Im Sinne der fokussierten Ethnografie bzw. Videografie (Knoblauch, 2006) wird der Alltag in 8 Klassen während je einer Woche mittels Videoaufnahmen und Feldnotizen dokumentiert. Zusätzlich werden Unterrichtskommentare der Lehrpersonen festgehalten sowie Informationen zu den Berufsbiografien der Lehrpersonen und familiären Bedingungen der Kinder erhoben. Diese Daten werden im Sinne einer offenen, zirkulären und schrittweise vertiefenden Forschungsstrategie (Strauss & Corbin, 1990 [1996]) analysiert: Auf der Basis eines Logbuchs werden einzelne Gesprächspassagen ausgewählt und in Datensitzungen sequenzanalytisch ausgewertet. Durch offene Kodierverfahren und unter Bezug ethnografischen Wissens werden kommunikative Formen und Merkmale von Herstellungsprozessen identifiziert, konzeptionell ausdifferenziert und verglichen. Durch den Einbezug unterschiedlicher (verbaler und visueller) Datensorten und das Aushandeln gemeinsamer Lesarten im Analyseteam werden die Ergebnisse trianguliert. Verfahren des Theoretical Sampling (Strauss, 1994) ermöglichen die Ausarbeitung und Validierung von Konzepten. Die bisherigen Auswertungen haben sich im Hinblick auf alle drei Fragestellungen als ergiebig erwiesen: An "schulnahen" kommunikativen Formen zeigen sich beispielsweise unterschiedliche Typen von kollektiven Interaktionen, Belehrungen und monologischen Sprachhandlungen. Interessante Befunde zu den Herstellungsprozessen betreffen u.a. den Schutz des gemeinsamen Projekts durch Lehrpersonen und Kinder, den Wechsel zwischen unterschiedlichen Referenzräumen und den Einsatz non- und paraverbalen Mittel zur Unterstützung der Ko-Konstruktion. Im Hinblick auf die Reproduktion von Ungleichheit sind die An- oder Aberkennung familiärer Erfahrungen, die unterschiedliche Verbindlichkeit beim Bearbeiten von Verstehensproblemen oder die Gewährung von privilegierten Rollen als bedeutsam einzustufen. Im Vortrag werden ausgewählte Konzepte anhand von Videodaten vorgestellt und diskutiert.

Bourdieu, P. & Passeron, J. (1971). Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs. Stuttgart: Klett.

Bos, W., Lankes, E., Schwippert, K., Valtin, R., Voss, A., Badel, I. & Plassmeier, N. (2003). Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E. Lankes & K. Schwippert (Hrsg.), Erste Ergebnisse aus IGLU (S. 69–142). Münster: Waxmann.

Erard, S. & Schneuwly, B. (2005). Le didactique de l'oral: savoirs ou compétences? In J.-P. Bronckart, E. Bulea M. Pouliot (Ed.), Repenser l'enseignement des langues: comment identifier et exploiter les compétences? (pp.69–97). Villeneuve d'Ascq cedex: Presses Universitaires de Septentrion.

Garfinkel, H. (1967). Studies in Ethnomethodology. Englewood Cliffs: Prentice Hall.

Gutiérrez, K. & Rogoff, B. (2003). Cultural Ways of Learning: Individual Traits or Repertoires of Practice. Educational Researcher, Vol. 32, No. 5, pp. 19–25.

Käsler, D. (2005). Sprachbarrieren im Bildungswesen. In Berger, P. & Kahlert, H. (Hrsg.). Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert (S. 130–154). Weinheim: Juventa.

van Kleeck, A. & Schwarz, A. (2011). Making academic talk explicit. Research directions for fostering classroom discourse skills in children from nonmainstream cultures. Schweizerische Zeitschrift für Bildungswissenschaften 1/2011, S. 29–46.

Koch, P. & Österreicher, W. (1994). Schriftlichkeit und Sprache. In: Günther, H. & Ludwig, O. (Hrsg.). Schrift und Schriftlichkeit, Band 1 (S. 587–604). Berlin: De Gruyter.

Knoblauch, H. (2006). Videography. Focused Ethnography and Video Analysis. In H. Knoblauch, B. Schnettler, J. Raab, H.-G. Soeffner (Ed.). Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology (pp. 69–83). Frankfurt a. M.: Peter Lang.

Sammons, P., Anders, Y., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Barreau, S. (2008). Children's Cognitive Attainment and Progress in English Primary Schools During Key Stage 2: Investigating the potential continuing influences of pre-school education. In H. Roßbach & H. Blossfeld (Hrsg.), Frühpädagogische Förderung in Institutionen (S. 179–197). Wiesbaden: VS Verlag für Sozialwissenschaften (Zeitschrift für Erziehungswissenschaft. Sonderheft 11 / 2008).

Schütz, A. (1932 [1974]). Der sinnhafte Aufbau der sozialen Welt. Frankfurt/Main: suhrkamp.

Strauss, A. (1994). Grundlagen qualitativer Sozialforschung. Datenanalyse und Theoriebildung in der empirischen und soziologischen Forschung. München: Fink.

Strauss, A. & Corbin, J. (1990 [1996]). Grounded Theory: Grundlagen Qualitativer Sozialforschung. Weinheim: Beltz.

Thévenaz-Christen, T. (2005). Les prémices de la forme scolaire. Etudes d'activités langagières orales à l'école enfantine genevoise. Genève: Université de Genève.

Tomasello, M. (1999). The Cultural Origins of Human Cognition. Cambridge: Harvard University Press.

Vygotzky, L. (1934, trad. 1986). Denken und Sprechen. Frankfurt/Main: Fischer Taschenbuch.



## ***Aufgaben und Funktionen von Kindertageseinrichtungen aus Sicht angehender frühpädagogischer Fachkräfte***

Hintergrund und Fragestellung: Die Ziele, die frühpädagogische Fachkräfte ihrer Tätigkeit und der Institution Kindertagesstätte zuschreiben, sind relevante Aspekte der Orientierungsqualität des pädagogischen Personals in Kindertageseinrichtungen. In dieser Studie werden daher die zentralen Aufgaben und Funktionen von Kindertageseinrichtungen (sowie dem dort arbeitenden Personal) aus Sicht angehender frühpädagogischer Fachkräfte untersucht. Dabei wurde von der im Sozialgesetzbuch sowie in nationalen Bildungsvereinbarungen üblichen Dreiteilung der Aufgaben und Funktionen in die Bereiche „Bildung“, „Erziehung“ und „Betreuung“ ausgegangen, die sich auch in anderen vorliegenden Studien als Strukturierung pädagogischer Einstellung bewährt hat (Tietze, 1998). Die Wahrnehmung dieser verschiedenen Aufgaben und Funktionen durch frühpädagogische Fachkräfte kann schließlich verstanden werden als Teil professioneller Identität, welche u.a. dazu beiträgt die Anforderungen des Berufs angemessen bewältigen zu können (Kunter, 2013). Diese bezieht sich zum einen auf ein handlungsleitendes Rollen- und Selbstverständnis, andererseits auf eine kritische Selbstreflexion und die persönliche Weiterentwicklung der pädagogischen Fachkraft. Typischerweise verändern sich diese pädagogischen Orientierungen sowohl über die Zeit der Ausbildung und Qualifikation (Dippelhofer-Stiem, 2006). Außerdem ist zu erwarten, dass Frühpädagogik-Studierende an Hochschulen z.T. andere Einstellungen und Überzeugungen aufweisen als Auszubildende an Fachschulen (Mischo et al., 2012, 2013). Insgesamt ergaben sich für die vorliegende Studie somit folgende Fragestellungen: (1) Welche dimensionale Struktur weisen die zentralen Aufgaben und Funktionen von Kindertageseinrichtungen aus Sicht angehender frühpädagogischer Fachkräfte auf? (2) Gibt es Unterschiede in der dimensionalen Struktur der perzipierten Aufgaben und Funktionen zwischen Ausbildungsbeginn und -ende sowie zwischen den Ausbildungsebenen (Fachschule vs. Hochschule)? (3) Gibt es Unterschiede im Niveau der perzipierten Aufgaben und Funktionen zwischen Ausbildungsbeginn und -ende sowie zwischen den Ausbildungsebenen?

Methode und Ergebnisse: Die Daten wurden mithilfe eines standardisierten Fragebogens erhoben, welche Skalen zur subjektiven Einschätzung der Relevanz von Aufgaben und Funktionen einer Kindertageseinrichtung (Items zu den Funktionen Bildung, Betreuung, Erziehung) sowie zu Erziehungszielen (das Kind als Erziehungsobjekt vs. Erziehungsgegenstand; vgl. Tietze, 1998) enthielten. Die Gesamtstichprobe umfasste 1571 Personen, wobei sowohl angehende Frühpädagogen befragt wurden, die sich am Beginn ihrer Ausbildung befanden, als auch solche, die ihre Ausbildung bereits so gut wie abgeschlossen hatten. Die TeilnehmerInnen schätzten dabei ein, wie wichtig ihnen persönlich diese Aufgaben und Funktionen erscheinen. Die Analyse der Daten mit Mplus umfasste schließlich insgesamt drei Schritte: Zunächst wurde eine konfirmatorische Prüfung verschiedener konkurrierender theoretischer Messmodelle durchgeführt. In einem zweiten Schritt wurde die Messinvarianz des Instruments bzgl. der einzelnen Subgruppen (Erst- vs. Sechstsemester bzw. Fach- vs. Hochschulstudierende) mittels multipler Gruppenvergleiche geprüft (Step-Up-Ansatz nach Brown, 2006). Zuletzt wurde schließlich ermittelt, wie sehr angehende frühpädagogische Fachkräfte die einzelnen Aufgaben und Funktionen wahrnehmen sowie der Effekt des Ausbildungszeitpunktes und des Institutionstyps im Rahmen einer Moderatoranalyse untersucht. Es zeigte sich, dass angehende frühpädagogische Fachkräfte entgegen der üblichen Verwendung der Begrifflichkeiten vorrangig zwei Hauptfunktionen einer Kindertagesstätte wahrnehmen: eine Bildungsfunktion, welche v.a. die kindliche Persönlichkeitsentwicklung in den Mittelpunkt stellt sowie eine Betreuungsfunktion. Diese Struktur ist dabei weitgehend unabhängig von Ausbildungszeitpunkt und Ausbildungsart (partielle Messinvarianz) – angehende frühpädagogische Fachkräfte unterschiedlichen Semesters bzw. unterschiedlichen Institutionstyps nehmen lediglich einzelne Indikatoren der jeweiligen Bereiche als unterschiedlich relevant bzw. als unterschiedlich stark mit dem jeweiligen Konstrukt verbunden war. Darüber hinaus schätzen angehende frühpädagogische Fachkräfte die Bildungsfunktion am Ende ihrer Ausbildung bzw. ihres Studiums als wichtiger ein als zu Beginn ihrer Ausbildung/ihrer Studiums. Gleichzeitig schätzten Fachschülerinnen die Betreuungsfunktion als wichtiger ein als Studierende. Unterschiedliche Interpretationsmöglichkeiten dieser Befunde und Implikationen für die Entwicklung einer spezifischen beruflichen Identität in den beiden Ausbildungssystemen (Fachschule, Hochschule) werden diskutiert und weitere Forschungsfragen werden abgeleitet.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. London, New York: Guilford Press.

Dippelhofer-Stiem, B. (2006). Berufliche Sozialisation von Erzieherinnen. In: Pädagogik der frühen Kindheit. Weinheim: Beltz.

Mischo, C., Hendl, J., Wahl, S. & Strohmmer, J. (2013). Kompetenzen angehender frühpädagogischer Fachkräfte an Fachschulen und Hochschulen. Empirische Pädagogik, 27 (1), 22-46.

Mischo, C., Wahl, S., Strohmmer, J. & Hendl, J. (2012). Knowledge orientations of prospective early childhood teachers. Journal of Early Childhood Teacher Education, 33, 144-162.

Kunter, M. (2013). Professionalisierung von Lehrkräften. In M. A. Wirtz (Hrsg.), Dorsch – Lexikon der Psychologie (16. Aufl., S. 1226). Bern: Verlag Hans Huber.

Tietze, W., Meischner, T., Gänsfuß, R., Grenner, K. Schuster, K.-M., Völkel, P. & Roßbach, H.-G. (1998). Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten. Neuwied: Luchterhand.

## "Professionalisierung von Fachkräften im Elementarbereich" PRIMEL

In diesem Beitrag soll das Forschungsprojekt „Professionalisierung von Fachkräften im Elementarbereich“ (PRIMEL) sowie erste Ergebnisse präsentiert werden. PRIMEL wird vom BMBF im Rahmen der Initiative „Ausweitung der Weiterbildungsinitiative Frühpädagogische Fachkräfte“ (AWiFF) gefördert. Beteiligte Kooperationspartner des PRIMEL Projekts sind: Goethe Universität Frankfurt (Prof. Dr. D. Kucharz, Dipl. Päd. M. Tournier), Leibniz Universität Hannover (Prof. Dr. K. Mackowiak, Dipl.-Psych. H. Wadepohl), Pädagogische Hochschule Weingarten (Prof. Dr. M. Dieck, Prof. Dr. E. Rathgeb-Schnierer, Prof. Dr. S. Ziroli, Dipl.- Soz.- Päd. M. Janßen, M.A. C. Hüttel, Dipl.-Sportwiss. U. Billmeier), Universität Koblenz-Landau (Prof. Dr. A. Kauertz, B.Ed K. Gierl) in Kooperation mit den Pädagogischen Hochschulen St. Gallen (M.A. S. Bosshart) und Schaffhausen (M.A. C. Lieger, Dipl. Psych. FH/ SBAP C. Burkhardt Bossi).

### Theoretischer Hintergrund

Durch die Ergebnisse (inter-)nationaler (Längsschnitt-)studien wie beispielsweise der OECD Studie Starting Strong (OECD, 2001) ist die Qualität der pädagogischen Arbeit in Kindergärten und damit die Qualität des pädagogischen Handelns von frühpädagogischen Fachkräften vermehrt in den Fokus gerückt. Des Weiteren weisen Langzeitstudien einen Zusammenhang zwischen der Kompetenz der pädagogischen Fachkräfte und der emotionalen, sozialen und kognitiven Entwicklung der Kinder nach (ECCE, 1997; EPPE, 2004). Dies führte zu einer Neuaufwertung des Auftrags der Kindertageseinrichtungen, bestehend aus der Trias „Bildung, Erziehung und Betreuung“, indem ihr Bildungsauftrag verstärkt in den Blick genommen wurde (Fthenakis, 2003; Rauschenbach, 2002; Stamm, 2010).

Diese Neuaufwertung hat auch Auswirkungen auf die Aus- und Weiterbildung der pädagogischen Fachkräfte, (OECD, 2001, 2004), die ihren Ausdruck in der bis heute andauernden Diskussion um die Akademisierung des Erzieher/innenberufs findet.

### Fragestellung

Im Forschungsprojekt PRIMEL wird die Qualität der Freispielbegleitung sowie Angebotsplanung und -gestaltung durch pädagogische Fachkräfte in Kindertageseinrichtungen analysiert. Dabei soll untersucht werden, welchen Einfluss folgende Faktoren auf das pädagogische Handeln haben:

- unterschiedliche Ausbildungsmodelle für pädagogische Fachkräfte im Elementarbereich in Deutschland (fachschulische und akademische Ausbildung) und der Schweiz (akademische Ausbildung).
- pädagogische und fachdidaktische Einstellungen, Fähigkeitsselbstkonzepte und Fachwissen der pädagogischen Fachkräfte (Selbsteinschätzung).
- die räumliche, materielle und personelle Ausstattung der Einrichtung.

### Design

Insgesamt wurden im PRIMEL-Projekt 90 pädagogische Fachkräfte (30 fachschulisch ausgebildete Fachkräfte/ 30 akademisch ausgebildete Fachkräfte in Deutschland/ 30 akademisch ausgebildete Kitalehrpersonen in der Schweiz) mittels Fragebögen befragt und an fünf Tagen videographiert (Freispielbegleitung sowie Angebotsdurchführung in den Bildungsbereichen Mathematik, Naturwissenschaften, Bewegung/ Sport und Kunst-Ästhetik).

Mittels der Fragebögen wurden die Strukturqualität (Ausbildung der Fachkräfte, Ressourcen der Kita in Form von Materialien, Personalschlüssel, Zahl der Gruppen, Größe der Einrichtung u.Ä.) sowie die Orientierungsqualität (Einstellung, Fähigkeitsselbstkonzept und Wissen der pädagogischen Fachkräfte) erfasst.

Über die Videographie von Freispiel- und Angebotssituationen wurde die Prozessqualität im Sinne der pädagogisch-psychologischen und didaktischen Handlungs-kompetenz der Fachkräfte erhoben. Die Auswertung erfolgt quantitativ mittels eines deduktiv entwickelten Kategoriensystems.

Es werden erste Ergebnisse aus der Fragebogen- und der Videoanalyse präsentiert.

ECCE Study Group (1997). European Child Care and Education Study: Cross National Analyses of the quality and effects of early childhood programmes on children's development. Free University Berlin.

Fthenakis, W. E. (Hrsg.). (2003). Elementarpädagogik nach PISA – Wie aus Kindertagesstätten Bildungseinrichtungen werden können. Freiburg, Basel, Wien: Herder.

Rauschenbach, T. (2002). Der Bildungsauftrag des Kindergartens. In: Ministerium für Frauen, Jugend, Familie und Gesundheit des Landes Nordrhein-Westfalen/ Sozialpädagogisches Institut des Landes Nordrhein-Westfalen (Hrsg.): Lebensort Kindertageseinrichtung. Bilden - Erziehen - Fördern. Frühkindliche Bildung im Kindergarten. Chancen und Möglichkeiten nach der PISA-Studie. Dokumentation.

[http://www.tageseinrichtungen.nrw.de/diskurs/doku\\_ws5.pdf](http://www.tageseinrichtungen.nrw.de/diskurs/doku_ws5.pdf), S. 13-21.

Stamm, Margrit (2010). Frühkindliche Bildung, Betreuung und Erziehung. Haupt: Bern, Stuttgart; Wien.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2004). The Effective Provision of Pre-School Education Project (EPPE) - Zu den Auswirkungen vorschulischer Einrichtungen in England. In: G. Faust, M. Götz, H. Hacker & H. Rossbach (Hrsg.). Anschlussfähige Bildungsprozesse im Elementar- und Primarbereich. Bad Heilbrunn: Klinkhardt (S. 139-153).

OECD (2001). Starting Strong. Early Childhood Education and Care. Paris: OECD.

OECD (2004). Starting Strong. Die Politik der frühkindlichen Betreuung, Bildung und Erziehung in der Bundesrepublik Deutschland. Paris.





## Effekte adaptierbarer Kooperationskripts auf den Erwerb von Kooperationsfertigkeiten

Chair(s): Ingo Kollar (Universität Augsburg)

### Theoretischer Hintergrund

Kooperationskripts sind Maßnahmen zur Unterstützung kooperativer Lernprozesse, die innerhalb von Kleingruppen Aktivitäten und Kooperationsrollen verteilen (Fischer, Kollar, Stegmann, & Wecker, 2013) und die sich zur Förderung von Kooperationsfertigkeiten als förderlich erwiesen haben (z.B. Stegmann, Wecker, Weinberger & Fischer, 2012). Kritisiert wird jedoch, dass Kooperationskripts eine starre Anleitung kooperativer Lernprozesse realisieren (Hesse, 2007). Dies wirft die Frage auf, wie der Kooperationskriptansatz flexibilisiert werden kann. Neben adaptiven Kooperationskripts, die etwa auf Basis von Intelligent-Tutoring-Ansätzen eine kontinuierliche Berechnung der Qualität der ablaufenden Kooperationsprozesse ermöglichen und das Skript kontinuierlich entsprechend anpassen (Diziol, Walker, Rummel, & Koedinger, 2010), kommen hierfür adaptierbare Kooperationskripts in Frage. Dabei können die Lernenden selbst entscheiden, welche Kooperationskriptvorgaben sie zur Strukturierung ihrer Kooperation erhalten möchten und welche nicht (Leutner, 2009). Es kann angenommen werden, dass die Möglichkeit, selbst über den Grad der Kooperationsstrukturierung zu entscheiden, bei den Lernenden insbesondere metakognitive Reflexionsprozesse über den eigenen Kooperationsprozess auslöst, die in einem positiven Zusammenhang mit dem Erwerb von Kooperationsfertigkeiten stehen.

Fragestellung: In einer experimentellen Studie wurde den folgenden Fragen nachgegangen:

(1) Was sind die Effekte eines adaptierbaren Kooperationskripts vs. eines starren Kooperationskripts vs. unstrukturierter Kooperation auf den Erwerb von Kooperationsfertigkeiten in einer computerunterstützten Lernumgebung?

(2) Inwiefern vermitteln metakognitive Prozesse die Effekte eines adaptierbaren Kooperationskripts auf den Erwerb von Kooperationsfertigkeiten?

### Methode

87 Studierende aus sozialwissenschaftlichen Studiengängen wurden zufällig zu Triaden zusammengeschlossen, die in einer asynchronen, textbasierten Lernumgebung an der Analyse von drei auf Basis der Attributionstheorie zu lösenden pädagogisch-psychologischen Problemfällen arbeiten sollten. Die Lernzeit betrug 70 Minuten. Die Triaden wurden zufällig den drei experimentellen Bedingungen zugewiesen. In der Bedingung „unstrukturierte Kooperation“ arbeiteten die Triaden ohne externe Unterstützung zusammen. In der Bedingung „starrs Kooperationskript“ wurden die Lernenden zur Einhaltung eines Peer-Review-Prozesses angehalten, demzufolge jeder Lernende für sich jeweils eine erste Analyse zu einem der drei Problemfälle formulieren und dann die Analysen seiner beiden Lernpartner kritisieren sollte. Danach sollte jeder Lernende auf die beiden Kritiken zu seiner initialen Analyse reagieren. Dieser Prozess wurde ein zweites Mal wiederholt und endete mit einer abschließenden Analyse für jeden Fall. Zu jedem Schritt wurden Textprompts präsentiert, die die Erstellung der Analysen bzw. Kritiken unterstützten (z.B. „Folgende Punkte sind mir noch unklar“ als Prompt für das Erstellen einer Kritik). Diese Prompts wurden über den gesamten Kooperationsprozess in unveränderter Form präsentiert. In der Bedingung „adaptierbares Kooperationskript“ hatten die Lernenden die Möglichkeit, einzelne Kooperationskriptvorgaben an- und auszuschalten und das Skript so an ihre subjektiven Bedürfnisse anzupassen. Die Kooperationsfertigkeiten wurde in einem Nachtest gemessen, in dem die Triaden einen weiteren Problemfall ohne Unterstützung analysieren sollten. Die Gruppendiskussionen wurden im Hinblick auf die Anwendung von Kooperationsfertigkeiten analysiert, die durch das Kooperationskript angeleitet worden waren (Cohen's Kappa  $\kappa = .71$ ). Das gleiche Kodierschema wurde zur Analyse eines Vortests verwendet, in dem die Lernenden individuell aufgefordert wurden, zu beschreiben, wie gute Kooperation ablaufen sollte. Zudem wurden die während des Lernprozesses ablaufenden metakognitiven Prozesse mit den Kategorien „Planung“, „Monitoring“ und „Reflexion“ mit hinreichender Interraterreliabilität kodiert (Cohen's Kappa  $\kappa = .85$ ).

### Ergebnisse und Diskussion

Mehrebenenanalysen zeigten, dass Studierende der Bedingung mit adaptierbarem Kooperationskript sowohl signifikant bessere Werte hinsichtlich der erworbenen Kooperationsfertigkeiten erreichten als Lernende ohne Kooperationskript ( $p < .05$ ) als auch als Lernende mit starrem Kooperationskript ( $p < .05$ ). Zwischen den Bedingungen ohne Kooperationskript und starrem Kooperationskript wurde kein signifikanter Unterschied bezüglich der erworbenen Kooperationsfertigkeiten festgestellt. Der positive Effekt des adaptiven Kooperationskripts wurde zudem durch die metakognitiven Prozesse „Monitoring“ (Sobel-Z = 2.57, SE = 0.05,  $p < 0.05$ ) und „Reflexion“ (Sobel-Z = 3.24, SE = 0.09,  $p < 0.01$ ) vermittelt. Adaptierbare Kooperationskripts scheinen also eine geeignete Möglichkeit darzustellen, Kooperationskripts flexibler zu gestalten. Durch die Aufforderung, das Kooperationskript aktiv anzupassen, wurden in vermehrtem Maße metakognitive Prozesse ausgelöst, die ihrerseits den Erwerb von Kooperationsfertigkeiten positiv beeinflussen.

Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using Intelligent Tutor Technology to Implement Adaptive Support for Student Collaboration. *Educational Psychology Review*, 22, 89–102. doi:10.1007/s10648-009-9116-9

Fischer, F., Kollar, I., Stegmann, K. & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56-66.

Hesse, F. W. (2007). Being told to do something or just being aware of something? An alternative approach to scripting CSCL. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning – cognitive, computational, and educational perspectives* (pp. 91-98). New York, NY: Springer.

Leutner, D. (2009). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In L. J. Issing & P. Klimsa (Eds.), *Online-Lernen: Handbuch für Wissenschaft und Praxis*. MUniversität: Oldenbourg.

Stegmann, K., Wecker, C., Weinberger, A. & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science*, 40(2), 297-323.

Dienstag,  
04.03.2014,  
14:15 - 16:00,  
PEG-Gebäude  
Raum 1.G102





## ***Effekte eines Trainings zur Förderung epistemologischer Überzeugungen und intellektueller Werte***

Bei der Verarbeitung konfligierender wissenschaftlicher Positionen (etwa zu Themen der nachhaltigen Entwicklung) ist es für die Entwicklung eines tieferen Verständnisses der Thematik notwendig, Fertigkeiten des argumentativen Denkens (etwa das Abwägen von Argumenten) anzuwenden. Für die Bereitschaft, diese Fertigkeiten anzuwenden, gibt es nach Kuhn, Cheney und Weinstock (2000) zwei zentrale Voraussetzungen: Zum einen dienen „reife“ (evaluativistische) epistemologische Überzeugungen als rationale Basis, argumentatives Denken als sinnvoll zu betrachten. Zum anderen reflektieren intellektuelle Werte als Ausmaß der Wertschätzung intellektuellen Engagements, ob argumentatives Denken auch als wünschenswert erachtet wird. Neben zeitintensiven interaktiven Diskussionen zur Förderung epistemologischer Überzeugungen und intellektueller Werte (etwa Kuhn, 2005; Kuhn & Udell, 2003) zeigten bisherige Kurzzeit-Interventionen zur Förderung epistemologischer Überzeugungen (Kienhues, Bromme, & Stahl, 2008; Pieschl, Bromme, Porsch, & Stahl, 2008) gemischte Befunde. Vor diesem Hintergrund war unser Ziel, eine kurze Trainingsintervention zur Förderung von evaluativistischen epistemologischen Überzeugungen und intellektuellen Werten zu entwickeln. Die bisherige Forschung identifizierte die Anregung von Selbsterklärungen durch Leitfragen – sogenannte Selbsterklärungsprompts – beim Lernen aus Beispielfideos als effektive Methode, um argumentative Fertigkeiten zu fördern (Busch, Renkl, & Schworm, 2008; Schworm & Renkl, 2007). Hauptbestandteil unserer computerbasierten ca. 50-minütigen Trainingsintervention waren daher zwei Beispielfideos, in denen jeweils zwei Personen über konfligierende Positionen zu authentischen Themen der nachhaltigen Entwicklung (Wiederansiedlung des Luchses, Klimawandel und Waldschäden) diskutierten. Eine der beiden gezeigten Personen modellierte evaluativistische epistemologische Überzeugungen und intellektuelle Werte. Die zweite Person übernahm schrittweise die epistemologischen Überzeugungen und intellektuelle Werte der ersten Person und machte deren Eigenschaften und Konsequenzen für die Bereitschaft zum argumentativen Denken deutlich. Zusätzlich wurden die Lernenden durch Selbsterklärungsprompts aufgefordert, sich die modellierten epistemologischen Überzeugungen und intellektuellen Werte und deren Konsequenzen für die Bereitschaft zum argumentativen Denken zu erklären. Der letzte Bestandteil der Trainingsintervention war eine selbstregulierte Argumentationsphase. Es wurden konfligierende Positionen zu einer neuen Thematik (Anbau gentechnisch veränderter Pflanzen im Freiland) präsentiert, und die Lernenden wurden aufgefordert, dazu selbstreguliert eine eigene Position zu entwickeln. Gemäß Empfehlungen von Harris, Alexander und Graham (2008) zur Festigung neuerer Strategien erhielten die Lernenden in dieser Phase die Möglichkeit, nun selbst evaluativistische epistemologischen Überzeugungen zu aktivieren. Vor diesem Hintergrund untersuchten wir folgende Fragestellung: Hat unsere Trainingsintervention positive Effekte auf epistemologische Überzeugungen, intellektuelle Werte sowie das konzeptuelle Wissen über beide Konstrukte?

Wir testeten unsere Trainingsintervention mit insgesamt 66 Schülerinnen und Schülern der gymnasialen Oberstufe in einem experimentellen Design mit zwei Bedingungen: (a) Trainingsintervention (Trainingsgruppe, n= 33) und (b) keine Trainingsintervention (Kontrollgruppe, n= 33). Unsere neu entwickelten Instrumente zur Erfassung epistemologischer Überzeugungen, intellektueller Werte sowie konzeptuellen Wissens über beide Konstrukte überprüften wir in einer Vorstudie mit einer eigenen Stichprobe (N= 68). Außerdem nutzten wir das von Teilnehmenden der Vorstudie gewonnene Feedback zur Verbesserung und Weiterentwicklung unserer Trainingsmaterialien. Wir erfassten die epistemologischen Überzeugungen, die intellektuellen Werte sowie das konzeptuelle Wissen über beide Konstrukte sowohl unmittelbar nach dem Experiment als auch in einem verzögerten Nachtest eine Woche später.

Die Ergebnisse zeigen, dass die Trainingsintervention einen positiven Effekt auf das konzeptuelle Wissen über epistemologische Überzeugungen und intellektuelle Werte hatte. Dieser Effekt war auch eine Woche nach dem Experiment noch nachweisbar und wurde durch die Qualität, mit der sich die Lernenden die im Video modellierten epistemologischen Überzeugungen und intellektuellen Werte und deren Konsequenzen für die Bereitschaft zum argumentativen Denken erklärten, mediiert. Weiterhin waren mehr Lernende der Trainingsgruppe als der Kontrollgruppe in der Lage, in der selbstregulierten Argumentationsphase evaluativistische epistemologische Überzeugungen zu aktivieren. Eine Woche nach dem Experiment wies die Trainingsgruppe höher entwickelte epistemologische Überzeugungen, höher ausgeprägte intellektuelle Werte sowie ein höheres konzeptuelles Wissen über beide Konstrukte auf als die Kontrollgruppe. Zusammenfassend ist unsere kurze Trainingsintervention ein effektiver erster Schritt zur Förderung evaluativistischer epistemologischer Überzeugungen, intellektueller Werte sowie konzeptuellem Wissen über beide Konstrukte.

- Busch, C., Renkl, A., & Schworm, S. (2008). Towards a generic self-explanation training intervention for example-based learning. In G. Kanselaar, V. Jonker, P. A. Kirschner & F. J. Prins (Eds.), *Proceedings of the 8th International Conference of the Learning Sciences*. Utrecht, Netherlands: ICLS.
- Harris, K. R., Alexander, P., & Graham, S. (2008). Michael Pressley's contributions to the history and future of strategies research. *Educational Psychologist*, 43(2), 86-96.
- Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology*, 78(4), 545-565. doi: 0.1348/000709907x268589
- Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, 15(3), 309-328. doi: 10.1016/s0885-2014(00)00030-7
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5), 1245-1260. doi: 10.1111/1467-8624.00605
- Pieschl, S., Bromme, R., Porsch, T., & Stahl, E. (2008). Epistemological sensitisation causes deeper elaboration during self-regulated learning. Paper presented at the Proceedings of the 8th international conference on international conference for the learning sciences - Volume 2, Utrecht, The Netherlands.
- Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99(2), 285-296. doi: 10.1037/0022-0663.99.2.285



## **Verbesserung der Schulnoten durch gemeinsame Hausaufgabenbearbeitung? Ergebnisse einer kombinierten Netzwerk- und Propensity Score Analyse**

### I Fragestellung

Auf der Basis zweier Panelwellen einer Netzwerkerhebung wird untersucht, ob eine gemeinsame Bearbeitung von Hausaufgaben zum Zeitpunkt t1 (Jg. 5) die Noten zu t2 (Jg. 6) verbessert. Derartige Analysen der Wirkung des sozialen Kapitals in Peernetzwerken stehen vor einem Endogenitätsproblem, weil die Selektion der Subjekte in das Treatment (hier: Hausaufgabennetzwerke) stark mit dem Outcome konfundiert sein kann. Gerade jene, die schlechte Noten haben, suchen Unterstützung. Gerade von jenen, die gute Note haben, wird Unterstützung gesucht.

### II Theoretischer Ansatz: Sozialkapital der Peergruppe

In Studien zur Migrations- und Integrationsforschung wurde das Sozialkapital als positive Ressource betrachtet (Haug & Pointer, 2007), obwohl auch dessen „negative Seite“ in Form von engen Bindungen zur ethnischen Community hervorgehoben wurde, die Freiheiten einschränken können (Portes, 1998).

Bislang gibt es jedoch nur wenige bildungssoziologische Studien, die soziales Kapital in Form von Beziehungen untersuchen, die Kinder zu ihrer Peergruppe haben (Windzio, 2013). Dies liegt an der eingegrenzten Perspektive auf den innerfamiliären Statusvermittlungsprozess. Doch auch Schulkinder können sich untereinander unterstützen, und zwar sowohl in der täglichen sozialen Interaktion als auch in Fragen des schulischen Lernens, z. B. bei Hausaufgaben. Es könnte somit wichtig sein, mit wem die Kinder im Klassenraum interagieren und mit welchen Kindern sie lernbezogene Unterstützungsnetzwerke aufbauen. Mit anderen Worten: Auch die „Qualität“ der Mitschüler und Freunde, deren Habitus und Kompetenzen, könnten Elemente des sozialen Kapitals darstellen. Aber welche Faktoren fördern oder verhindern lernbezogene Unterstützungsnetzwerke? Und verbessern Hausaufgabennetzwerke die Note, wenn man die besondere Selektivität in diese Netzwerke berücksichtigt?

In sozialen Netzwerken existiert ein beträchtliches Maß an sozialer, ethnischer, aber auch religiöser Homophilie (McPherson et al., 2001). Dies gilt insbesondere für Kontaktdimensionen, die ein höheres Niveau an Investition in soziale Kontakte verlangen, bspw. das gemeinsame Verbringen der Freizeit oder gegenseitige Besuche zu Hause (Windzio & Bicer, 2013). Die Kehrseite dieser Homophilie ist eine ethnische und soziale Segregation der Kontakte in Netzwerken. Damit ist aber offenkundig, dass der Zugang in die Hausaufgabennetzwerke vermutlich selektiv ist und diese Selektivität bei der kausalen Analyse berücksichtigt werden muss.

### III. Methoden

Auf der Basis von Dyaden wird im ersten Schritt ein Netzwerkmodell (Exponential Random Graph Model) der Hausaufgabenkontakte geschätzt, das ebenfalls Freundschaften der Kinder sowie Kontakte unter ihren Eltern berücksichtigt. Daraus wird ein Propensity Score gebildet, anhand dessen die Subjekte der Treatment- und der Kontrollgruppe mittels Matching-Verfahren balanciert werden, um einen kausalen Effekt der Hausaufgabennetzwerke auf die Notenverbesserung zu schätzen. Abschließend wird eine zentrale Annahme der Treatment-Effects Modells diskutiert, nämlich die stable Universitäts treatment value assumption (SUTVA), der zufolge keine Interaktionen zwischen Treatment- und Kontrollsubjekten erlaubt (Gangl, 2010). Diese Annahme ist in Netzwerken grundsätzlich nicht erfüllt, aber der Ansatz des Propensity Score Weighting (Guo & Fraser, 2010, p. 161) ist geeignet, um zumindest basale Muster sozialer Interaktionen (transitive und zyklische Triaden) statistisch zu kontrollieren.

### IV. Ergebnisse

Das Modell zur Vorhersage der Selektion zeigt, dass Hausaufgabennetzwerke an sich nicht durch ethnische oder religiöse Homophilie geprägt sind. Ethnische Homophilie finden wir aber in Freundschaftsnetzwerken, und in Elternnetzwerken kommt noch eine religiöse Homophilie hinzu. Kontakte in den drei Netzwerkdimensionen Hausaufgaben, Freundschaft und Elternkontakte sind stark voneinander abhängig. So schlägt die ethnische und religiöse Segregation in der Elterngeneration auch auf die Hausaufgabennetzwerke der Kinder durch. Es deutet sich zudem an, dass je schlechter ein Schüler ist, desto stärker sucht er Kontakte, und je besser ein Schüler ist, wird stärker sein Kontakt gesucht. Allerdings lässt sich in keiner der Analysevarianten ein positiver Effekt der Hausaufgabennetzwerke auf die Schulleistung bzw. auf deren Verbesserung feststellen.

Gangl, Markus (2010): Nichtparametrische Schätzung kausaler Effekte mittels Matchingverfahren. In: Christof Wolf und Henning Best (Hg.): Handbuch der sozialwissenschaftlichen Datenanalyse. 1. Aufl. Wiesbaden: VS Verl. für Sozialwiss., S. 931–961.

Guo, Shenyang; Fraser, Mark W. (2010): Propensity score analysis. Statistical methods and applications. Los Angeles, Calif: Sage.

Haug, Sonja; Pointer, Sonja (2007): Soziale Netzwerke, Migration und Integration. In: Axel Franzen und Franzen-Freitag (Hg.): Sozialkapital. Grundlagen und Anwendungen. Wiesbaden: VS Verl. für Sozialwiss. (Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft, 47), S. 367–396.

McPherson, J. Miller; Smith-Lowin, Lynn; Cook, James M. (2001): Birds of a feather: Homophily in social networks. In: Annual Review of Sociology 27, S. 415–444.

Portes, Alejandro (1998): Social capital: Its origins and applications in modern sociology. In: Annual Review of Sociology 24, S. 1–24.

Windzio, Michael (2013): Immigrant children's access to social capital in school-class networks. In: Michael Windzio (Hg.): Integration and Inequality in Educational Institutions. Dordrecht: Springer, S. 191–228.

Windzio, Michael; Bicer, Enis (2013): Are we just friends? Immigrant integration into high- and low-cost social networks. In: Rationality & Society 25, S. 123–145.



## **Musikalisches involviert sein - ein anderer Blick auf kulturelle Teilhabe von Grundschulkindern aus dem JeKi-Programm**

Theoretischer Hintergrund: Der Begriff der kulturellen Teilhabe

Kulturelle Teilhabe zu fördern ist zentraler Bestandteil vieler schulischer und außerschulischer Initiativen. Allerdings sind die Begriffe "kulturell" und "Teilhabe" unscharf (vgl. im Überblick Bockhorst, Reinwand-Weiss, & Zacharias, 2011). Problematisch erscheint dabei vor allem, dass nicht selten erstens je nach Definition sogenannte „hochkulturelle“ Aktivitäten zur Messlatte für das Gelingen der Förderung kultureller Teilhabe gemacht wird und dass zweitens die Frage aktiven Instrumentalspiels, wie es auch im JeKi-Programm gefördert wird, zur maßgeblichen Form von Teilhabe erklärt wird. Nicht immer werden die normativen Setzungen, die auf dichotomen Unterscheidungen beruhen, offengelegt (Lehmann-Wermser, in Druck; Lehmann-Wermser & Campos, 2013). Unklar bleibt angesichts der verschiedentlich beklagten Forschungslücken (Autorengruppe Bildungsbericht, 2012; Liebau et al., 2013), inwiefern die genannten Ansätze geeignet sind, kulturelle und speziell musikalische Teilhabe als Wege zur Musikkultur zu erfassen.

Fragestellung

Vor diesem Hintergrund stellen sich die folgenden Fragen:

1. Welche musikalischen Erfahrungen werden von Grundschulkindern „berichtet“ und wie lassen sich diese Erfahrungen systematisch beschreiben?
2. Welche musikalischen Erfahrungen werden aus dem familiären und schulischen Umfeld der Schülerinnen und Schüler „berichtet“ und in welchem Verhältnis stehen sie zum Selbstzeugnis der Kinder?
3. Lässt sich Teilhabe aus dem Datenmaterial heraus modellieren?

Methode

Im Kontext des Forschungsverbundes SIGrun wurde im Verlauf des zweiten und vierten Schuljahres Material erhoben. Es bestand aus „Malmappen“, in denen Kinder Erlebnisse mit Musik zeichnen sollten, Fotografien und Tagebücher aus den Elternhäusern sowie umfangreiche Interviews mit den Kindern und ihren Eltern sowie mit Lehrerinnen und Lehrern. Für diesen Vortrag wird auf den Korpus von 503 Kinderzeichnungen und 60 Fotos näher eingegangen.

Für die Analyse visueller Daten liegen inzwischen umfangreiche methodische und methodologische Überlegungen vor (vgl. im Überblick Friebertshäuser, 2007). Insbesondere der seriell-ikonografische Ansatz (Pilarczyk & Mietzner, 2005) ist dabei hilfreich, in einem an Mayring (2008) orientierten Verfahren induktiv Kategorien zu bilden.

Ergebnisse

Es konnte eine große Vielfalt an musikalischen Umgangsformen beschrieben werden, die für die Kinder bedeutsam sind. Aufschlussreich war, dass sie dabei Zugangsweisen nicht nur entlang der o.g. dichotomen Unterscheidungen ausprägten, sondern synchron unterschiedliche Formen. Dabei entwickelten sie domänenspezifische Expertise. Dass medial vermittelte Musik eine große Rolle spielt, überrascht nicht (Medienpädagogischer Forschungsverbund, 2003), wohl aber, dass dabei bereits Ansätze einer „Medienkompetenz“ (Baacke) erkennbar sind.

In der Konsequenz erscheint es sinnvoller, in Abgrenzung vom wenig hilfreichen Begriff kultureller Teilhabe von (musikalischem) "involviert sein" zu sprechen, wenn man versuchen will, kindliche und jugendliche Zugangsformen zu modellieren.

Autorengruppe Bildungsberichterstattung (Hrsg.) (2012). Bildung in Deutschland 2012. Bielefeld: Bertelsmann.

Bockhorst, H.; Reinwand-Weiss, V.; Zacharias, W. (Hrsg.) (2012). Handbuch Kulturelle Bildung. München: kopaed.

Friebertshäuser, B. (Hrsg.) (2007). Bild und Text. Methoden und Methodologien visueller Sozialforschung. Opladen: Budrich.

Lehmann-Wermser, A. (i. Dr.). Kulturelle Teilhabe. Über die Schwierigkeit, ein populäres Konstrukt empirisch zu erforschen. In: C. Khittl und St. Dressler-Zöllner (Hrsg.). Musik - pädagogisch - politisch. Festschrift Arnold Werner Jensen (S. 163–184). Heidelberg: Heidelberger Hochschulschriften,.

Lehmann-Wermser, A.; Jessel-Campos, C. (2013). Aneignung von Kultur. Wege zu kultureller Teilhabe und zur Musik. In: A. Hepp & A. Lehmann-Wermser (Hrsg.). Transformation des Kulturellen (S. 152–177). Wiesbaden: VS Verlag für Sozialwissenschaften (GWV).

Liebau, E. et al. (2013): Forschung zur Kulturellen Bildung in Deutschland. Bestand und Perspektiven. Ein Projektbericht. Friedrich-Alexander-Universität Erlangen-Nürnberg.

Mayring, Ph. (Hrsg.) (2008): Die Praxis der qualitativen Inhaltsanalyse. 2. Aufl. Weinheim, Basel: Beltz (Pädagogik). Medienpädagogischer Forschungsverbund Südwest (2003). KIM-Studie 2003. Kinder und Medien, Computer und Internet; online unter: <http://bildungsplus.forumbildung.de/templates/>.

Pilarczyk, U.; Mietzner, U. (2005). Das reflektierte Bild. Die seriell-ikonografische Fotoanalyse in den Erziehungs- und Sozialwissenschaften. Bad Heilbrunn: Klinkhardt.



Christian Brzinsky-Fay (Wissenschaftszentrum Berlin (WZB))

## ***Empirische Bildungsforschung kurzgefasst - Ein Science Slam***

Chair(s): Christian Brzinsky-Fay (Wissenschaftszentrum Berlin (WZB))

Diskutant/in: Jonna Blanck (Wissenschaftszentrum Berlin (WZB))

Das College for Interdisciplinary Educational Research (CIDER) stellt als vom BMBF und der Jacobs Foundation gefördertes Projekt eine neuartige Form der Nachwuchsförderung dar. Die Karriereförderung auf frühen PostDoc-Ebene erfordert ein großes Maß an Flexibilität, da in dieser Phase die Spezialisierung sowohl inhaltlich als auch hinsichtlich der Karriere sehr weit fortgeschritten ist. Das College zielt deshalb darauf ab, seinen 30 Fellows möglichst flexible Unterstützung für Auslandsaufenthalte, Mentoring, Mikrogruppentreffen, Publikationen, Konferenzbesuchen etc. zu gewähren, damit sie ihre Karriere ihrer individuellen Situation anpassen können.

Darüber hinaus hat das College inhaltlich das Ziel, interdisziplinäre Bildungsforschung zu fördern. Dazu wurden unter Einbeziehung der Disziplinen Soziologie, Erziehungswissenschaften, Psychologie und Ökonomie drei Hauptthemenfelder formuliert:

- educational inequalities and success against the odds,
- competence development as a cognitive and social process
- monetary and non-monetary returns to education.

Das CIDER-Symposium wird in einem kurzen Eingangsvortrag das College, die Fördermaßnahmen und seine inhaltlichen Schwerpunkte vorstellen. Anschließend werden eine Reihe von Fellows und Seniors in Speed Presentations unterschiedlicher Art ihre Arbeiten präsentieren, so dass der interdisziplinäre Charakter des College zur Geltung kommt.

Die Beiträge des Symposiums werden in Englisch und Deutsch vorgetragen.

Weitere Informationen über das College for Interdisciplinary Educational Research: [www.ciderweb.de](http://www.ciderweb.de)

**GEBF** TAGUNG  
2014

ÖF1

Öffentliches  
Forum

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
Casino-Gebäude  
Raum 823

2

**Empirische Bildungsforschung kurzgefasst - Ein Science Slam**

## **Bildung in Informationsumwelten**

Chair(s): Friedrich W. Hesse (Leibniz-Institut für Wissensmedien)

Diskutant/in: Rainer Bromme (Westfälische Wilhelms-Universität Münster)

Digitale Medien haben wesentlich dazu beigetragen, Bildungsprozesse zu transformieren. Das Internet bietet eine Fülle von Bildungsgelegenheiten. Beispielsweise ermöglichen neue Technologien, stärker mit Lerninhalten zu interagieren; darüber hinaus gibt es völlig neue Formen, sich mit anderen Lernenden weltweit auszutauschen. Die Veränderungen durch digitale Medien manifestieren sich nicht nur im Feld der formalen Bildung, sondern entfalten sich insbesondere bei informellen Bildungsprozessen im Kontext lebenslangen Lernens. Lernende beziehen bildungsrelevante Informationen aus verschiedenen Kontexten, die dann eine „idiosynkratische“ Informationsumwelt für jedes Individuum konstituieren. Solche Informationen können aus Offline- und Online-Kontexten stammen, die formaler oder informeller Natur sind. Digitale Medien tragen in besonderer Weise zu diesem Prozess bei, da sie nicht nur neue Kontexte erschaffen, sondern auch traditionelle Offline-Kontexte transformieren.

Der WissenschaftsCampus Tübingen „Bildung in Informationsumwelten“ untersucht empirisch die Frage, wie digitale Medien Bildungsprozesse formale und informelle Bildungsprozesse beeinflussen. Ein solcher Forschungsansatz ist notwendigerweise interdisziplinär, und so vereint die Forschung im WissenschaftsCampus Ansätze unter anderem aus der Psychologie, der Erziehungswissenschaft, der Informatik, der Soziologie und der Wirtschaftswissenschaft.

Das Symposium stellt aus den acht interdisziplinären Clustern des WissenschaftsCampus vier Bereiche vor, die die Vielfalt des Forschungsansatzes belegen. Der erste Beitrag (Buder) charakterisiert den Begriff der Informationsumwelt und gibt einen Überblick über das Forschungsthema. Es werden vier Kernbereiche identifiziert, die sich durch die Verfügbarkeit digitaler Medien ergeben: digitale Partizipation, digitale Autonomie, digitale Kompetenzförderung und digitale Intervention. Zum ersten Kernbereich werden empirische Ergebnisse berichtet. Der zweite Beitrag (Sassenberg et al.) zum Thema digitale Autonomie beschreibt am Beispiel der Gesundheitsbildung, wie die Vielfalt von Informationen im World Wide Web es Individuen ermöglicht, gesundheitsbezogene Informationen zu erschließen, und wie sich dies auf die Autonomie von Patienten und die Interaktion zwischen Ärzten und Patienten auswirkt.

Der dritte Beitrag (Nürk et al.) zum Thema digitale Kompetenzförderung zeigt auf, wie digitale Medien eingesetzt werden können, um basale Rechenkompetenzen im Sinne eines kollaborativen und spielbasierten Ansatzes zu trainieren.

Im vierten Beitrag (Scheiter et al.) zum Thema digitale Intervention geht es darum, wie digitale Medien Lernprozesse unterstützen können, indem sie Lernenden Möglichkeiten zur Adaptation der Lernumgebung anbieten oder gar adaptiv auf das Verhalten der Lernenden reagieren. Die Vielfalt der dargestellten Ansätze im Symposium wird abschließend kritisch eingeordnet und diskutiert (Bromme).



## **Rahmenkonzepte der Bildung in Informationsumwelten**

Dieser Beitrag charakterisiert den Begriff der Informationsumwelt und beschreibt dessen interdisziplinäre Genese. Es wird ein Rahmenmodell für Bildungsprozesse dargestellt, in welchem das lebenslang lernende Individuum im Mittelpunkt steht, und welches sowohl für Offline- als auch für Online-Kontexte konzipiert ist. Im Zentrum des Rahmenmodells stehen vier Begriffe, die den individuellen Bildungsprozess charakterisieren: Der Begriff der Informationsvielfalt beschreibt, dass sich lebenslang Lernende einer Fülle von zum Teil widersprüchlichen Informationen ausgesetzt sehen. Informationsvielfalt lässt sich sowohl auf einer zeitlichen Mikroebene (z.B. bei der Verarbeitung von gleichzeitig präsentierten Elementen auf der Seite eines Lehrbuchs) als auch auf einer Makroebene beschreiben (z.B. bei der Suche nach Informationen aus verschiedenen Webseiten). Die Begriffe der Selbstregulation und Fremdregulation beschreiben, wie Individuen mit Informationsvielfalt umgehen. Selbstregulation bedeutet, dass Lernende eine Eigenverantwortung bei der Auswahl an bildungsrelevanten Informationen haben, sei es im Mikrobereich (z.B. Rückgriff auf Lernstrategien) oder im Makrobereich (z.B. selbstbestimmte Erschließung von Themen im WWW). Gleichzeitig haben individuelle Bildungsprozesse auch fast immer fremdregulierte Anteile, sei es durch die Interaktion mit Lehrkräften, durch den Austausch mit Peers, oder durch den Einsatz adaptiver und/oder filternder Technologien. Ein vierter Begriff des Rahmenmodells (Diagnostik) verweist auf metakognitive Aspekte des Bildungsprozesses. Sowohl das lernende Individuum, als auch andere Beteiligte am Bildungsprozess (Lehrkräfte, adaptive Technologien) können dazu beitragen, den Bildungsprozess eines Individuums zu überwachen, Lernzustände zu diagnostizieren, um dann adäquat in das Spannungsfeld aus Selbstregulation und Fremdregulation eingreifen zu können. Ausgehend von diesem Rahmenkonzept werden vier empirisch adressierbare Kernbereiche identifiziert, die darlegen, wie die Verfügbarkeit von digitalen Medien den Bildungsprozess beeinflusst. Im Spannungsfeld zwischen Selbst- und Fremdregulation adressieren die ersten beiden Kernbereiche (digitale Partizipation, digitale Autonomie) eher den Aspekt der Selbstregulation; der dritte und vierte Kernbereich (digitale Kompetenzförderung, digitale Intervention) bezieht sich stärker auf fremdregulatorische Facetten des Bildungsprozesses. Digitale Partizipation umfasst die vielfältigen Möglichkeiten, die digitale Medien bieten, um an Bildungsprozessen aktiv teilzuhaben, also bildungsbezogene Informationen nicht nur zu rezipieren, sondern auch zu produzieren. Dies gilt insbesondere in Web 2.0-Kontexten (z.B. Wikipedia, Online-Foren). Digitale Autonomie adressiert den Umstand, dass es für informelle Bildung kein Curriculum gibt und geben kann – es obliegt der Autonomie von lebenslang Lernenden, sich ihre Informationsumwelt auszuformen. Dies lässt sich an Themen wie der Gesundheitsbildung illustrieren, in welcher eine Autonomie des Individuums aus medizinethischer Sicht ein sehr hohes und schützenswertes Gut darstellt. Digitale Kompetenzförderung zeigt Möglichkeiten auf, wie durch den gezielten Einsatz von digitalen Medien Lernumgebungen geschaffen werden können, die es ermöglichen, basale Kompetenzen (Lesen, Rechtschreiben, Rechnen) nicht nur zu diagnostizieren, sondern auch gezielt zu trainieren. Beim Themenbereich „Digitale Intervention“ schließlich steht das Problem im Vordergrund, dass gerade in informellen Bildungskontexten keine Lehrkraft zugegen ist, die einen Lernprozess begleiten kann. Diese Funktion der Fremdregulation kann aber teilweise durch den Einsatz adaptiver und adaptierbarer Technologien aufgefangen werden. Im abschließenden empirischen Abschnitt dieses Beitrags werden kurz Ergebnisse zum Themenbereich der digitalen Partizipation vorgestellt. Dabei geht es einerseits um Feldstudien zur zeitlichen Entwicklung von Wikipedia-Seiten; andererseits um die Frage nach der Gestaltung von Online-Diskussionsforen, die sowohl die Rezeption als auch die Produktion von Diskussionsbeiträgen fördern können.

Gesundheitsbildung hat sich seit der Entstehung und breiten Nutzung des World Wide Web (WWW) grundlegend verändert. Vor der Entstehung des WWW vermittelten vor allem Ärzte und Schulen Gesundheitswissen. Zusätzlich gab es in vielen Haushalten ein Nachschlagewerk, das Hausmittel und Diagnosen vergleichsweise kompakt zusammenfasste. Durch das WWW wird der gesamte Kanon des schul- und alternativmedizinischen Wissens der breiten Öffentlichkeit zugänglich. Damit kommen zu Informationsangeboten, deren Ziel die korrekte Information von Patienten und Gesunden ist, Informationsangebote, denen vor allem ein kommerzielles oder anderweitig fragwürdiges Interesse zugrunde liegt. Der Befund, dass die Ausbreitung von Grippeepidemien seit einigen Jahren verlässlich anhand der Lokalisierung von Google Suchbegriffeingaben verfolgt werden kann, ist nur ein Beleg für die weitreichende Nutzung des WWW zur Gesundheitsbildung.

Aus der Kombination der häufigen Nutzung des WWW zur Gesundheitsbildung und der heterogenen Qualität der zur Verfügung stehenden Inhalte ergibt sich ein zentrales Problem hinsichtlich der von Patienten zu treffenden Entscheidungen. Die Informationssuche im WWW erfolgt selbstreguliert. Bei der selbstregulierten Informationssuche haben die Ziele der Suchenden naturgemäß eine besonders starke Auswirkung auf die Selektion und Rezeption von Information. Bei Patienten steht das Ziel gesund zu werden naturgemäß im Vordergrund und ist möglicherweise stärker ausgeprägt als das Ziel, sich ein korrektes Bild vom State of the Art zu machen. Dementsprechend sollte mit zunehmender gesundheitlicher Bedrohung die Informationssuche im WWW zu einer stärkeren Rezeption von positiver, heilungsbezogener Information führen.

Diese Hypothese wurde in einer Reihe von Experimenten und in einer längsschnittlichen Studie geprüft. Dabei zeigte sich bei Bedrohung (verglichen mit unterschiedlichen sinnvollen Kontrollbedingungen) die erwartete Präferenz für positive und heilungsrelevante Information bei der Auswahl von Suchbegriffen und der Selektion von Links aus einer Liste von Suchergebnissen. Außerdem wurden bei Bedrohung auch mehr positive Informationen aus Texten extrahiert und als Ergebnisse einer WWW-Suche eher erinnert als in einer affektiv neutralen Kontrollbedingung. Schließlich konnte in einer Längsschnittstudie gezeigt werden, dass die häufige Verwendung des WWW zur Informationssuche im Gesundheitsbereich dazu führt, dass Patienten mit zunehmender Schwere der Erkrankung (und damit auch mit zunehmender Bedrohung) ein optimistischeres Bild von der eigenen Gesundheit hatten. Letzteres weist darauf hin, dass die Tendenz, unter Bedrohung mehr positive Information zu suchen, zu selektieren und zu erinnern, langfristig dazu führt, dass bei starker WWW-Nutzung (d.h. bei häufiger selbstgesteuerter Informationssuche) ein positiv verfälschtes Bild von der eigenen Situation entsteht. Passend dazu führt die häufige WWW-Nutzung zu gesundheitsrelevanten Themen auch zu einem höheren gesundheitsbezogenen Selbstbewusstsein bei Patienten.

Diese Befunde haben deutliche Implikationen für die Medizinethik und die Medizindidaktik. Die Autonomie des Patienten in Bezug auf medizinische Entscheidungen ist ein Grundpfeiler der Medizinethik. Ärzte sollten ihre Patienten so informieren, dass die Patienten im Rahmen ihrer Möglichkeiten selbst Entscheidungen über medizinische Maßnahmen treffen können. Durch das WWW können sich Patienten nicht nur unabhängig vom Arzt informieren (was zunächst einmal zur Autonomie beiträgt). Patienten, die das WWW zur Informationssuche nutzen, kommen mit großer Wahrscheinlichkeit – wie die Befunde nahelegen – auch zu einem zu optimistischen Bild ihrer Gesundheitssituation und sind gleichzeitig auch selbstbewusster. Somit besteht für den Arzt aus ethischen Gründen die Notwendigkeit, die Patienten ausführlicher zu informieren und falsche oder unvollständige Einschätzungen zu korrigieren.

Hans-Christoph Nürk (Eberhard Karls Universität Tübingen), Korbinian Möller (Leibniz-Institut für Wissensmedien),  
Jürgen Heller (Eberhard Karls Universität Tübingen), Torsten Grust (Eberhard Karls Universität Tübingen),  
Ann-Christine Ehlis (Universitätsklinik Tübingen)

## ***Aneignung grundlegender Kulturtechniken mit Hilfe von Medien am Beispiel mathematischer Fähigkeiten***

Grundsätzliches Ziel des vorgestellten Verbundprojektes ist es, die Differentialdiagnostik im Leistungsbereich nach der Grundschule zu verbessern und auf Basis dieser Diagnostik eine web-basierte interaktive Lernplattform zu konzipieren, implementieren und (auch neurokognitiv) zu evaluieren. Mit dieser Plattform sollen Schüler auch außerhalb formeller Lernumwelten spielerisch und adaptiv basale Kulturtechniken erlernen. Eine wesentliche diagnostische Voraussetzung ist die differentialpsychologische Untersuchung der Defizite im Rechnen und Rechtschreiben in der 5. und 6. Klasse, da in der Differentialdiagnostik die Altersbereiche nach der Grundschule bislang nur wenig erforscht wurden. Das Endprodukt des Clusters ist die web-basierte Realisierung einer Lerninteraktionsplattform.

Mathematische Fertigkeiten werden in einem interdisziplinären Forschungsverbund untersucht, in welchem verschiedene Teilprojekte inhaltlich eng miteinander zusammenarbeiten. Der Ausgangspunkt ist, dass mathematische Leistungen einer der wichtigsten Prädiktoren für späteren beruflichen Erfolg sind. Dennoch gibt es vor allem im Haupt- und Realschulbereich kaum eine gezielte systematische Differentialdiagnostik. Die Leistungsunterschiede zwischen den Schulformen scheinen in den meisten Studien in der Sekundarstufe immer größer zu werden (Schereneffekt). Allerdings gibt es speziell für die mathematischen Kompetenzen inkonsistente Befunde. Im Cluster wurden daher von 415 SchülerInnen der 5. und 6. Klasse an Haupt- und Realschulen deren arithmetische Basiskompetenzen in den vier Grundrechenarten erhoben. Realschüler schnitten grundsätzlich besser ab als Hauptschüler. Jedoch fanden wir differentielle Ergebnisse hinsichtlich der Leistungsentwicklung an den beiden Schulformen, die abhängig vom Aufgabentypus und der Aufgabenkomplexität waren. Während sich für einfache und schwierige Additionsaufgaben die Leistungsunterschiede zwischen Haupt- und Realschule von der 5. zur 6. Klasse verringerte, wurden nur für schwierige Divisionen eine Vergrößerung des Leistungsunterschieds beobachtet. Eine globale Aussage über die Leistungsentwicklung in Mathematik in der Sekundarstufe ist daher aus unserer Sicht nicht zulässig. Vielmehr hängt der Schereneffekt von der Aufgabenart ab und sogar innerhalb einer Aufgabenart von der Aufgabenkomplexität. Es scheint, dass die Hauptschüler bei einfacheren, schon aus der Grundschule bekannten Aufgaben(typen) aufholen, während sie bei komplexeren Aufgaben noch mehr den Anschluss verlieren.

Um mögliche Transfereffekte der zu implementierenden Lernspiele besser abschätzen zu können, untersuchten wir auch den Zusammenhang zwischen verschiedenen arithmetischen Fertigkeiten. Dabei zeigte sich insbesondere bei guten Schülern ein ausgeprägter Zusammenhang zwischen einfacher Multiplikation und einfacher Division (kleines  $1 \times 1$ ), während bei schlechteren Schülern dieser Zusammenhang weniger ausgeprägt war. Für schwierige Multiplikationen und Divisionen zeigte sich dieser Befund nicht. Dieses Datenmuster ist konsistent mit der Annahme, dass vor allem gute Schüler Division über Multiplikation lösen. Daher sollen für eine interaktive Lernplattform Multiplikation, Division und deren Beziehung zueinander besonders trainiert werden. Auf der Basis dieser Befunde wurde zusammen mit einem Teilprojekt aus der Informatik eine web-basierte interaktive Lernplattform für rechenschwache Schülerinnen und Schüler der 5. und 6. Klasse implementiert. Diese unterstützt die Entwicklung interaktiver Lernspiele, in denen die Schüler bezüglich ihrer spezifischen Schwäche gefordert werden. Die Plattform ist im World Wide Web erreichbar und stellt zu bereits mehrere Lernspiele zur Verfügung und wurde in einem Pilotprojekt bereits eingesetzt – Möglichkeiten und Probleme beim Einsatz der Lernplattform werden diskutiert.

Neben der Realisierung weiterer Lernspiele steht nun der Einsatz im Feld, die Adaption und die Datenanalyse im Mittelpunkt. Adaption nutzt Protokolle zurückliegender Spieldurchgänge und deren Online-Auswertung, um Spiele für Spieler individuell zu konfigurieren. In einer integrierten Zusammenarbeit aller Projekte von Informatikern, Psychologen und der Medizinischen Fakultät unter Einbeziehung von benachbarten Clustern sollen in Zukunft erste neurokognitive Online-Adaptionen von Lernspielen entwickelt werden, in welchen die Spieleigenschaften an den neurokognitiven Status des Lernalters angepasst werden.

Katharina Scheiter (Leibniz-Institut für Wissensmedien), Rolf Plötzner (Pädagogische Hochschule Freiburg), Alexander Renkl (Universität Freiburg), Anne Schüler (Leibniz-Institut für Wissensmedien), Gottfried Zimmermann (Hochschule der Medien Stuttgart)

## ***Adaptierbarkeit und Adaptivität beim Lernen mit Multimedia***

Sowohl in formellen als auch informellen Bildungskontexten werden Informationen gegenwärtig immer mehr in multimedialer Form präsentiert. In der Multimediaforschung wurden in den letzten Jahren verschiedene Theorien und Modelle zum multimedialen Lernen vorgeschlagen, die beschreiben, wie Lernende multimedialen Präsentationen Informationen entnehmen und diese schrittweise verarbeiten sollten, um ein Verständnis der präsentierten Sachverhalte zu erzielen. Gleichzeitig wurden verschiedene Merkmale von Lernenden identifiziert, die die Verarbeitung multimedialen Lernmaterials moderieren. Dennoch wurden bis heute kaum Multimediaumgebungen entwickelt, die in Abhängigkeit dieser Merkmale und Präferenzen von Lernenden bezüglich der multimedialen Gestaltung adaptierbar oder gar adaptiv sind. Von einem adaptierbaren System spricht man dann, wenn Lernende selbst über die Darbietungsweise (z.B. über Repräsentationsformate, Sequenz der Informationspräsentation) entscheiden können. Ein adaptives System nimmt Lernenden diese Entscheidungen ab, indem es zunächst lernrelevante Zustände und Vorgehensweisen des Lernenden festgestellt (z.B. Vorwissen, Angemessenheit der Informationsverarbeitung), um dann die Darbietungsweise an diese Zustände anzupassen, so dass ein möglichst optimiertes Passungsverhältnis zwischen Informationspräsentation und Möglichkeiten bzw. Fähigkeiten der Informationsrezeption resultiert. Die Gestaltung adaptierbarer und adaptiver Multimediasysteme bedarf dabei sowohl pädagogisch-psychologischen Wissens für die Konzeption instruktorischer Unterstützungsmaßnahmen als auch technischer Expertise für die Implementierung entsprechender Systeme. Dementsprechend kooperieren in dem vorgestellten Cluster Psychologinnen und Psychologen sowie Informatiker in vier Teilprojekten bei der Entwicklung eines multimedialen Systems, welches sowohl Aspekte von Adaptierbarkeit als auch von Adaptivität umfasst. Im Beitrag werden die Ergebnisse von zwei Teilprojekten vorgestellt, die verschiedene Aspekte der Effektivität des entwickelten Multimediasystems adressieren.

In Teilprojekt 1 (Plötzner & Ruf) wird untersucht, durch welche Gestaltungsmaßnahmen die Nutzungshäufigkeit kognitiver Unterstützungsangebote in multimedialen Lernumgebungen gefördert werden kann. Zur Untersuchung dieser Frage wurde zunächst eine gebrauchstaugliche Benutzerschnittstelle für die multimediale Lernumgebung entwickelt. Dazu wurde die Gebrauchstauglichkeit der Benutzerschnittstelle in drei aufeinander folgenden Studien formativ evaluiert und schrittweise überarbeitet. Anschließend wurde anhand dieser Benutzerschnittstelle in einer experimentellen Studie untersucht, wie sich (a) die Verfügbarkeit von Möglichkeiten der Selbstkontrolle und (b) unterschiedliche Arten der Darbietung kognitiver Unterstützungsangebote auf die Nutzungshäufigkeit der Unterstützungsangebote auswirken. Zur Erfassung der Nutzungshäufigkeit wurden unter anderem die Blickbewegungen der Lernenden erfasst. Es konnte gezeigt werden, dass eine offenkundige und gleichzeitig unaufdringliche Art der Darbietung die Nutzungshäufigkeit der Unterstützungsangebote signifikant steigert. Die angebotenen Möglichkeiten der Selbstkontrolle führten hingegen nicht zu einer häufigeren Nutzung. Teilprojekt 2 (Scheiter, Schüler & Schubert) beschäftigt sich mit der Lernwirksamkeit adaptiver Mechanismen des multimedialen Systems, mit deren Hilfe die Darbietung des Lernmaterials an die Informationsverarbeitung der Lernenden angepasst werden soll. In einer ersten Studie wurden zunächst Blickbewegungsdaten genutzt, um verschiedene Muster des Blickverhaltens und ihren Zusammenhang mit dem Lernerfolg zu identifizieren. Eine Clusteranalyse basierend auf sechs Blickbewegungsparametern (z.B. Text- und Bildbetrachtungsdauer, Text-Bild-Wechsel) ermöglichte die Identifikation von drei Gruppen, die sich bezüglich des spezifischen Zusammenspiels dieser Parameter und des Lernerfolgs unterschieden. Basierend auf diesen Ergebnissen wird zurzeit eine zweite Studie durchgeführt, in der untersucht wird, inwieweit eine Anpassung der Multimediaumgebung an ungünstige Verarbeitungsweisen den Lernerfolg steigert. In der Kontrollgruppe lernen Studierende selbstgesteuert mit dem multimedialen System. In der Experimentalgruppe werden die Blickbewegungen der Lernenden automatisch und online ausgewertet. Das System passt sich automatisch an das Lernverhalten an, indem z.B. bei zu kurzen Betrachtungszeiten des Textes bzw. des Bildes die jeweilige Repräsentation zum erneuten Lesen bzw. Betrachten hervorgehoben wird und ein Fortschreiten im Lernprozess erst möglich ist, wenn eine festgesetzte Betrachtungsdauer überschritten wurde. Bei der online-Registrierung und Auswertung von Blickbewegungen sowie Umsetzung des adaptiven Mechanismus handelt es sich um eine innovative Weiterentwicklung computerbasierter Lernumgebungen, die ohne eine enge Zusammenarbeit von Psychologie und Informatik nicht möglich wäre.

Petra Stanat (IQB, Humboldt-Universität zu Berlin), Poldi Kuhl (IQB, Humboldt-Universität zu Berlin),  
Hans Anand Pant (IQB, Humboldt-Universität zu Berlin)

## ***Effekte integrativen / inklusiven Unterrichts für Kinder mit sonderpädagogischem Förderbedarf: Konzeptuelle Grundlagen und empirische Befunde***

Chair(s): Petra Stanat (IQB, Humboldt-Universität zu Berlin)

Diskutant/in: Marcus Hasselhorn (DIPF Frankfurt)

Im Zuge der Umsetzung der UN-Behindertenrechtskonvention werden derzeit die Bemühungen intensiviert, Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) inklusiv bzw. integrativ an Regelschulen zu unterrichten. Dies wirft die Frage nach den Effekten auf, die für diese Kinder und Jugendlichen mit dem Besuch einer Regelschule verbunden sein könnten. Auf der einen Seite wird angenommen, dass Heranwachsende mit SPF in Regelschulen im Vergleich zu Förderschulen eine anregendere Lernumgebung vorfinden, die ihre Entwicklung eher begünstigt (vgl. z.B. Ruijs & Peetsma, 2009; Wocken, 2005). Auf der anderen Seite wird argumentiert, dass Regelschulen weniger gut den besonderen Bedürfnissen von Schülerinnen und Schülern mit SPF entsprechen können als Förderschulen, die z.B. kleinere Klassengrößen aufweisen und über speziell ausgebildete Lehrkräfte verfügen, was sich eher ungünstig auf die Entwicklung dieser Gruppe von Heranwachsenden auswirken sollte (vgl. z.B. Peetsma et al., 2001).

Vor dem Hintergrund dieser umstrittenen Annahmen ist das Ziel des Symposiums, die schulische Situation von Kindern und Jugendlichen mit SPF zu analysieren und erste Befunde zu Effekten inklusiven / integrativen Unterrichts zu präsentieren. Im ersten Beitrag des Symposiums (Dietze & Moser) geht es zunächst darum, das Konzept des sonderpädagogischen Förderbedarfs genauer zu bestimmen und die verschiedenen Ansätze seiner Operationalisierung zu systematisieren. Dabei wird auch auf den Zusammenhang zwischen Art der Diagnostik einerseits und Ansätzen der Förderung andererseits eingegangen. Der zweite Beitrag (Gresch & Piezunka) beschäftigt sich mit länderspezifischen Variationen in der Ausgestaltung sonderpädagogischer Förderung und der Identifikation von Schülerinnen und Schülern mit SPF im Regelschulsystem. Es wird diskutiert, welche methodischen Implikationen damit für Forschung zu Effekten inklusiven / integrativen Unterrichts verbunden sind. Im dritten und vierten Beitrag schließlich werden potenzielle Erträge untersucht, die für Kinder mit SPF mit dem Besuch einer Regelschule bzw. einer Förderschule verbunden sein können. Während der dritte Beitrag (Schwinger et al.) Effekte auf motivationale und affektive Merkmale untersucht, werden im vierten Beitrag (Kocaj et al.) neben den Auswirkungen inklusiver / integrativer Förderung auf das akademische Selbstkonzept der Schülerinnen und Schüler mit SPF auch Kompetenzen in den Fächern Deutsch und Mathematik analysiert. Insgesamt wird das Symposium aus der Perspektive verschiedener Disziplinen einen Einblick in den aktuellen Forschungsstand zur Identifikation und Förderung von Schülerinnen und Schülern mit SPF geben und Implikationen für die weitere Forschung diskutieren.

Peetsma, Thea, Margaretha Vergeer, Jaap Roeleveld, und Sjoerd Karsten. 2001. Inclusion in education: Comparing pupils' development in special and regular education. *Educational Review* 53: 125–135.

Ruijs, Nienke M., und Thea T.D. Peetsma. 2009. Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review* 4: 67–79.

Wocken, Hans. 2005. Andere Länder, andere Schüler? Vergleichende Untersuchung von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen.

<http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf> (Zugegriffen Mai 31, 2013).

**GEBF** TAGUNG  
2014

ÖF3

Öffentliches  
Forum

Dienstag,

04.03.2014,

16:30 - 18:00,

IG-Farben-Haus

Raum 411

2

Effekte integrativen / inklusiven Unterrichts für Kinder mit Förderbedarf



## ***Perspektiven sonderpädagogischer Unterstützung. Bereitstellung von Ressourcen aus nationaler und internationaler Sicht***

Dieser Beitrag wird die Kategorie „Sonderpädagogischer Förderbedarf“ in Deutschland analysieren. Zunächst wird dargestellt, welche aktuellen Regelungen (Schulgesetze, Verordnungen) für die Diagnostik in den Bundesländern bestehen (Instrumente, beteiligte Akteure) und welche potenziellen Folgen damit verbunden sind. In dieser Analyse wird ein besonderer Fokus auf die Funktion der Diagnostik (Status- vs. Förderdiagnostik) gelegt sowie auf den allgemeinen Zusammenhang von „labeling“ und „funding“. Darüber hinaus ist in diesem Kontext zu fragen, in welchen Bundesländern die Kategorie „Sonderpädagogischer Förderbedarf“ als ressourcenlegitimierendes Etikett in der jüngeren Vergangenheit bereits Veränderungen unterzogen worden ist (z.B. im Rahmen von pauschalisierten Mittelzuweisungen [Lehrer, Lehrerstunden, nicht-pädagogisches Personal]). Abschließend soll diskutiert werden, welche sonderpädagogischen Förderbedarfe künftig in den Ländern, die auf „klassische“ Feststellungsverfahren verzichten, möglicherweise unterstellt werden. In diesem Zusammenhang ist insbesondere das Prinzip der „Prävention“ zu diskutieren, bei dem (sonder-)pädagogische Förderungen weniger im Kontext von Förderbedarfen als vielmehr im Kontext curriculumbasierter Interventionen stattfinden – hier sollen auch internationale Erfahrungen, insbesondere die sonderpädagogische Förderung in den USA und Finnland, einbezogen werden. Diese Praxis birgt besondere Herausforderungen für die Ressourcenzuweisung, da hier weder eindeutige Professionsbezüge noch definierte sonderpädagogische Förderbedarfe steuerungswirksam werden. Inwieweit stattdessen systemische Zuweisungen unter Heranziehung alternativer Indikatoren wie z.B. Sozialindizes sinnvoll sind, wird diskutiert.

Cornelia Gresch (Wissenschaftszentrum Berlin für Sozialforschung (WZB)), Anne Piezunka (Wissenschaftszentrum Berlin für Sozialforschung (WZB))

## **Schüler/innen mit sonderpädagogischem Förderbedarf an Regelschulen: Eine Bestandsaufnahme und Anforderungen an die Forschung**

In Deutschland wird der Großteil der Schüler\_innen mit sonderpädagogischem Förderbedarf an Förderschulen unterrichtet. Allerdings zeichnet sich in den vergangenen Jahren ein zunehmender Trend zu Gunsten einer inklusiven/integrativen Beschulung ab. Die Umsetzung von Integration/Inklusion unterscheidet sich dabei zwischen Bundesländern, Schulen sowie innerhalb von Schulen. Dies beginnt bereits bei der Definition und Erfassung der betroffenen Schüler\_innen, dem Umfang, in dem Integration/Inklusion realisiert wird, den jeweiligen Rahmenbedingungen an den Schulen bis hin zu der konkreten Einbindung der Schüler\_innen mit sonderpädagogischem Förderbedarf in den gemeinsamen Unterricht.

Entsprechend dem Trend zur inklusiven Beschulung beschäftigt sich auch die Forschung zunehmend mit dieser Thematik. So wurden in den vergangenen Jahren viele Teilaspekte von integrativer Beschulung untersucht, wie beispielsweise Leistungszuwächse oder soziales Klima in Integrationsklassen. Diese Studien konzentrierten sich allerdings auf einzelne Bundesländer oder Schulen. Zum aktuellen Zeitpunkt liegen keine bundesweiten Studien vor, die sich übergreifend mit der Inklusionssituation und deren Konsequenzen auseinandersetzen. Auch existieren bislang keine länderübergreifenden Standards, auf deren Basis eine solche Untersuchung umgesetzt werden kann, wie etwa eine übergreifende Identifikationsgrundlage oder Erhebungsinstrumente, die zur Erfassung der Umsetzung von Integration/Inklusion verwendet werden können.

In dem Beitrag wird dargestellt, welche Herausforderungen mit einer bundesweiten Untersuchung von Schüler\_innen mit sonderpädagogischem Förderbedarf an Regelschulen einhergehen, und es werden aktuelle Lösungen präsentiert. Dabei werden zum einen die Herausforderungen bei Stichprobenziehungen thematisiert und gezeigt, welche Aspekte der Umsetzung von Inklusion bei den Erhebungsinstrumenten und Auswertungen berücksichtigt werden sollten. Für die Stichprobenziehung ist es aufgrund des geringen Anteils inklusiv/integrativ unterrichteter Schüler\_innen notwendig, vorab Schulen bzw. Klassen zu identifizieren, in denen inklusiv/integrativ unterrichtet wird. Diese Informationen sind in den meisten Bundesländern über die Kultusministerien verfügbar und können als Grundlage für die Stichprobenziehung verwendet werden. Allerdings handelt es sich bei dem Kriterium selbst (diagnostizierter sonderpädagogischer Förderbedarf) um kein länderübergreifend standardisiertes Kriterium. Zudem bedarf es einer sehr umfangreichen Ausgangsstichprobe, da mit einer erhöhten Nonresponserate der Zielgruppe zu rechnen ist.

In der Erhebung selbst sollten zudem verschiedene Aspekte integrativer/inklusive Beschulung erfasst werden, um die Inklusionssituation der einzelnen Schüler\_innen angemessen abbilden zu können. Auf der Basis des adaptierten Input-Output-Modells von Dubs (2004) werden relevante Input- und Prozessdimensionen vorgestellt, die sich etwa auf die Ressourcenausstattung, Zugangsbedingungen usw. beziehen. Darüber hinaus sollten Prozessdimensionen, wie beispielsweise Beschulungstypen oder Unterrichtsgestaltung, erhoben werden.

Anknüpfend an bereits durchgeführte bzw. laufende Forschungsprojekte wird der Beitrag wichtige Hinweise für die Umsetzung einer bundesweiten empirischen Untersuchung zur Umsetzung von Integration/Inklusion geben.

Dubs, Rolf. 2004. Qualitätsmanagement für Schulen, Bd. 13. 2. Aufl. (Nachdr.). St. Gallen: Institut für Wirtschaftspädagogik.

**GEBF** TAGUNG  
2014

ÖF3 b

Symposiums  
beitrag

Dienstag,

04.03.2014,

16:30 - 18:00,

IG-Farben-Haus

Raum 411

2

Effekte integrativen / inklusiven Unterrichts für Kinder mit Förderbedarf

Malte Schwinger (Justus-Liebig-Universität Gießen), Carola Grunschel (Universität Bielefeld), Sittipan Yotyodying (Universität Bielefeld), Rebecca Baumanns (Universität Bielefeld), Julia Gorges (Universität Bielefeld), Björn Serke (Universität Bielefeld)

## **Hat inklusive vs. exklusive Beschulung einen Einfluss auf Motivation, Selbstwert und Wohlbefinden von Kindern mit sonderpädagogischem Förderbedarf? Erste Befunde der Bielefelder Längsschnittstudie BiLieF**

Theoretischer Hintergrund: Neben der Vermittlung fachlicher Kompetenzen stehen Lehrkräfte vor der Aufgabe, SchülerInnen in ihrer Persönlichkeitsentwicklung und dem Aufbau einer positiven Einstellung zum Lernen zu begleiten. Dies gilt angesichts des erhöhten Risikos schulischen Misserfolgs im Besonderen für Kinder mit Lernbeeinträchtigungen, weshalb sowohl die Herstellung eines ausreichenden schulischen Wohlbefindens als auch die Stärkung des Selbstwertgefühls und der Motivation wiederholt als Ziele sonderpädagogischer Förderung formuliert worden sind (zsf. Bless & Mohr, 2007). Die Ausprägung motivationaler und affektiver Merkmale dürfte maßgeblich davon beeinflusst werden, ob lernbeeinträchtigte Kinder inklusiv, also gemeinsam mit nicht beeinträchtigten Kindern, oder exklusiv in eigenen Förderschulen unterrichtet werden. Einerseits sind bei inklusiver Beschulung negative Bezugsgruppeneffekte zu erwarten, die zu einer Verringerung des akademischen Selbstkonzepts und der Selbstwirksamkeit beitragen (Marsh, 2005; Sauer et al., 2007). Andererseits kann argumentiert werden, dass der Selbstwert und das Selbstkonzept von SchülerInnen an Förderschulen durch den Prozess der Überweisung und die stigmatisierenden Wirkungen der Institution in hohem Maße beschädigt werden (Schumann, 2009).

Fragestellung: Vor dem Hintergrund der durch die Ratifizierung der UN-Konvention anstehenden Umwälzungen in der Beschulung von SchülerInnen mit Lernbeeinträchtigungen wird im vorliegenden Beitrag der Frage nachgegangen, ob sich inklusiv vs. exklusiv beschulte DrittklässlerInnen mit sonderpädagogischem Förderbedarf Lernen (SFB-L) hinsichtlich verschiedener motivationaler Merkmale (Selbstkonzept, Interesse, Anstrengungsbereitschaft, Zielorientierungen) und affektiver Variablen (Selbstwert, schulisches Wohlbefinden, Stigmatisierung) unterscheiden.

Methode: Die Ergebnisse basieren auf Daten der ersten Erhebungswelle der Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements (BiLieF), bei der zwischen Oktober 2012 und März 2013 im gesamten Bundesland Nordrhein-Westfalen N = 421 Schüler/-innen mit SFB-L (Alter: M = 8.74 Jahre, SD = 0.67) einzeln befragt wurden. Von den untersuchten Kindern besuchten n = 183 eine Förderschule mit Förderschwerpunkt Lernen (FS), n = 181 eine Integrationsklasse bzw. gemeinsamen Unterricht (GU) und n = 57 eine Grundschule mit Unterstützung durch ein Kompetenzzentrum für sonderpädagogische Förderung (KsF). Jedes Kind wurde einzeln mit einem vierstufigen Antwortformat standardisiert befragt. Selbstkonzept und Interesse wurden mit jeweils 3 Items getrennt für die Bereiche Mathe, Lesen und Schreiben erfasst. Die Skalen zur Anstrengungsbereitschaft (5 Items) sowie zur Lern-, Annäherungs-, Leistungs- und Vermeidungs-Leistungszielorientierung (jeweils 4 Items) wurden hingegen mit Bezug zur Schule allgemein erhoben. Die Höhe des allgemeinen Selbstwerts wurde mit 5 Items erfasst, während zur Messung des schulischen Wohlbefindens zwei Subskalen affektives und körperliches schulisches Wohlbefinden konzipiert wurden. Die 7 Items der affektiven Subskala enthielten jeweils den Itemstamm „Warst Du in der letzten Woche oft...?“ gefolgt von verschiedenen positiven und negativen Emotionen wie Freude, Stolz oder Ängstlichkeit. Das körperliche Wohlbefinden in der Schule wurde mit insgesamt 4 Items wie z.B. „Wie häufig hast Du in der Schule Bauchschmerzen?“ erfragt. Die Skala Stigmatisierung enthielt 5 Items wie z.B. „Andere Kinder machen sich lustig darüber, dass ich auf dieser Schule bin.“ Zusätzlich wurden die Intelligenz (CFT 1) sowie die Lese- und Schreibkompetenz der Kinder getestet (ELFE 1-6; HSP).

Ergebnisse: Alle Skalen zeigten akzeptable bis gute interne Konsistenzen ( $.66 \leq \alpha \leq .90$ ). Varianzanalytische Mittelwertvergleiche ergaben eine signifikant höhere Selbstwirksamkeit sowie ein signifikant höheres Stigmatisierungsempfinden für FS (exklusiv) verglichen mit GU und KsF (inklusiv) Kindern. Des Weiteren zeigten sich signifikante Unterschiede in Intelligenz, Lese- und Schreibkompetenz zugunsten der inklusiv beschulten Kinder. Die übrigen motivationalen und affektiven Merkmale wiesen ähnliche Ausprägungen in den drei Schülergruppen auf. Insgesamt stützen diese ersten Befunde sowohl die Annahme von Bezugsgruppen- wie auch Stigmatisierungseffekten verschiedener Beschulungsformen. Allerdings sollten die Ergebnisse aufgrund des querschnittlichen Datenmaterials nur vorsichtig interpretiert werden. Eine tiefere Prüfung der Zusammenhänge im Längsschnitt ist dringend erforderlich.

Marsh, H.W. (2005). Big-fish-little-pond effect on academic self-concept: A reply to responses. Zeitschrift für Pädagogische Psychologie, 19, 141-144.

Bless, G. & Mohr, K. (2007). Die Effekte von Sonderschulunterricht und Gemeinsamen Unterricht auf die Entwicklung von Kindern mit Lernbehinderungen. In: J. Walter & F. Wember (Hrsg.), Sonderpädagogik des Lernens (S. 385- 392). Stuttgart: Hogrefe-Verlag.

Sauer, S. & Ide, S. & Borchert, J. (2007). Zum Selbstkonzept von Schülerinnen und Schülern an Förderschulen und in integrativer Beschulung: Eine Vergleichsuntersuchung. In: Heilpädagogische Forschung, 3.

Schumann, B. (2009). Inklusion statt Integration – eine Verpflichtung zum Systemwechsel. Deutsche Schulverhältnisse auf dem Prüfstand des Völkerrechts. Pädagogik, 2, 52-53.

GEBF TAGUNG  
2014

ÖF3 c

Symposiums  
beitrag

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
IG-Farben-Haus  
Raum 411

2

Effekte integrativen / inklusiven Unterrichts für Kinder mit Förderbedarf

## **Akademisches Selbstkonzept und schulische Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf an Förder- und Regelschulen**

Im Rahmen der Evaluation integrativer bzw. inklusiver Förderung von Kindern mit sonderpädagogischem Förderbedarf (SPF) ist es wichtig, sowohl schulische Kompetenzen als auch motivationale Aspekte der schulischen Entwicklung zu untersuchen. Das akademische Selbstkonzept (ASK) wird dabei als zentrale motivationale Komponente erachtet. Kinder mit einem geringen ASK weisen eine niedrigere Erwartungshaltung hinsichtlich zukünftiger schulischer Leistungen auf und sehen ihre eigenen Fähigkeiten eher als unveränderbar an. Es wird angenommen, dass diese Zusammenhänge gerade auch bei Kindern mit SPF zur Aufrechterhaltung und Verstärkung schulischer Misserfolge beitragen (Deci & Chandler, 1986). Bei der Entwicklung des ASK vergleichen Schülerinnen und Schüler ihre eigenen schulischen Leistungen mit denen ihrer Klassenkameraden (sozialer Referenzrahmen; Festinger, 1954; Marsh, 2005b), wobei im Hinblick auf den Einfluss der Art der Beschulung zwei entgegengesetzte Prozesse wirken können: Zum einen könnte sich eine Beschulung von Kindern mit SPF in Förderschulen aufgrund von Stigmatisierungsprozessen negativ auf das ASK auswirken (Dunn, 1968; Tracey & Marsh, 2000). Zum anderen kann ein hohes Leistungsniveau der Referenzgruppe in Regelschulen aufgrund des Big-Fish-Little-Pond-Effekts (BFLPE; Marsh, 1987) das ASK von Kindern mit SPF verringern (Renick & Harter, 1989). Hinsichtlich der Auswirkung der Art der Beschulung auf die schulischen Kompetenzen von Kindern mit SPF werden von Befürwortern einer Beschulung in Regelschulen Befunde zu differenziellen Lern- und Entwicklungsmilieus (Baumert, Stanat, & Watermann, 2006; Bos, Müller, & Stubbe, 2010) und leistungsheterogenen Lerngruppen (Slavin, 1996) als mögliche Erklärungsansätze herangezogen. Andererseits kann von der Perspektive einer geschützten Lernumgebung in Förderschulen argumentiert werden, dass diese den besonderen Bedürfnissen der Schülerinnen und Schüler mit SPF durch kleinere Klassengrößen, speziell ausgebildete Lehrkräften und mehr Kapazitäten für eine individualisierte Förderung besser gerecht werden.

Im vorliegenden Beitrag wird basierend auf den Daten des IQB-Ländervergleichs 2011 für die Primarstufe (Stanat, Pant, Böhme, & Richter, 2012) zunächst der Frage nachgegangen, wie die Beschulung von Kindern mit SPF in Regel- und Förderschulen mit dem ASK in Mathematik und Deutsch zusammenhängt. Darüber hinaus wird untersucht, ob sich die schulischen Kompetenzen dieser Schülergruppe in Abhängigkeit von der Art der Beschulung unterscheiden. In die Analysen wurde eine Teilstichprobe von Schülerinnen und Schülern mit einem diagnostizierten SPF einbezogen (N = 1195, 63.8% in Regelschulen, 66.3% Jungen, MAlter = 10.9 Jahre), der eindeutig einem der Förderschwerpunkte Lernen, Sprache oder soziale und emotionale Entwicklung zugeordnet werden konnte. Unter Verwendung von Mehrebenenmodellen wurde geprüft, welchen Einfluss das mittlere Leistungsniveau und die Art der Beschulung unter Konstanthaltung der individuellen Schulleistung auf das ASK von Kindern mit SPF haben. Die individuellen schulischen Leistungen wurden mit standardisierten Kompetenztests erfasst und auf Klassenebene aggregiert. Beim Vergleich der schulischen Kompetenzen wurden Propensity Score Matching-Verfahren verwendet, um für Gruppenunterschiede in relevanten Hintergrundvariablen zu kontrollieren.

Erste Ergebnisse deuten darauf hin, dass Kinder mit SPF in Förderschulen ein durchschnittlich höheres ASK in Mathematik und Deutsch aufweisen als Kinder mit SPF in Regelschulen. Während die individuelle Leistung positiv mit dem ASK von Kindern mit SPF verknüpft ist, besteht ein negativer Zusammenhang der durchschnittlichen Klassenleistung mit dem ASK. Unter Kontrolle der individuellen Leistung und der mittleren Klassenleistung verschwindet der positive Effekt der Beschulung in Förderschulen auf das ASK.

Bezüglich der zweiten Fragestellung zeigte sich, dass Schüler mit SPF an Grundschulen signifikant höhere Kompetenzen im Lesen, Zuhören und in Mathematik aufweisen als vergleichbare Schüler an Förderschulen. Die Effekte sind besonders ausgeprägt für Kinder mit Lernbehinderungen, weniger prägnant hingegen für Kinder des Förderschwerpunktes Sprache. Diese Resultate, die auf Daten einer großangelegten Schulleistungsstudie basieren, unterstützen den bislang wenig gesicherten empirischen Forschungsstand (Bear, Minke, & Manning, 2002; Ruijs & Peetsma, 2009). Mögliche zugrunde liegende Prozesse sowie potenzielle Unterschiede in Abhängigkeit vom Förderschwerpunkt der Kinder werden diskutiert.

Baumert, J., Stanat, P., & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.

Bos, W., Müller, S., & Stubbe, T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Eds.), *Bildungsverlierer. Neue Ungleichheiten* (pp. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.

Bear, G. G., Minke, K. M., & Manning, M. A. (2002). Self-Concept of Students with Learning Disabilities: A Meta-Analysis. *School Psychology Review*, 31, 405–427.

Chapman, J. W. (1988). Learning Disabled Children's Self-Concepts. *Review of Educational Research*, 58(3), 347–371.

Deci, E. L., & Chandler, C. L. (1986). The importance of motivation for the future of the LD field. *Journal of Learning Disabilities*, 19, 587–594.

Dunn, L. M. (1968). Special education for the mildly retarded – is much of it justifiable? *Exceptional Children*, 35, 5–22.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280–295.

Marsh, H. W. (2005a). Big fish little pond effect on academic self-concept : cross-cultural and cross-disciplinary generalizability. (pp. 1–21). Presented at the Australian Association for Research in Education 2005, Parramatta, N.S.W.: Australian Association for Research in Education.

Marsh, H. W. (2005b). Der "Big-Fish-Little-Pond"-Effekt und das akademische Selbstkonzept. *Zeitschrift für Pädagogische Psychologie*, 19, 119–127.

Renick, M. J., & Harter, S. (1989). Impact of social comparisons on the developing self- perceptions of learning disabled students. *Journal of Educational Psychology*, 81, 631–638.

Ruijs, N. M., & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4, 67–79.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21, 43–69.

Stanat, P., Pant, H. A., Böhme, K., & Richter, D. (Eds.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

Tracey, D. T., & Marsh, H. W. (2000). Self-Concepts of Primary Students with Mild Intellectual Disabilities: Issues of Measurement and Educational Placement. In H. W. Marsh & R. G. Craven (Eds.), *Self-concept theory, research, and practice: Advances for the new millennium* (pp. 419–425). Sydney, Australia: University of Western Australia.



Birgit Spinath (Universität Heidelberg)

## ***Gesellschaftlich relevante Fragen der Bildungsforschung und Bildungspraxis: Zugänge und Ergebnisse aus vier Forschungsdisziplinen***

Chair(s): Birgit Spinath (Universität Heidelberg)

Diskutant/in: Birgit Spinath (Universität Heidelberg) wirklich alles Spinath?

Dieses Symposium zeigt auf, dass und wie sich Bildungsforschung mit gesellschaftlich hoch relevanten Fragen auseinandersetzt und dabei wichtige Erkenntnisse für die Bildungspraxis liefert (vgl. Spinath, in Druck). Es wird aufgezeigt, mit welchen Zugängen sich verschiedene Disziplinen den Fragestellungen nähern und welche Bezüge es zwischen den unterschiedlichen Zugängen und Ergebnissen gibt. Zu diesem Zweck kommen Wissenschaftlerinnen und Wissenschaftler aus vier Disziplinen – Soziologie, Psychologie, Ökonomie und Erziehungswissenschaft – zusammen und beleuchten die Themen soziale Ungleichheit und Bildungsqualität von Bildungseinrichtungen. Zu Beginn stellt der Betrag von Martin Diewald die klassisch soziologische Frage, wie wir die Ursachen sozialer Ungleichheit verstehen können. Diewald berichtet unter anderem über ein neues DFG-Projekt, in dem Soziologen und Psychologen klassische Methoden der Sozialforschung mit verhaltensgenetischen Methoden kombinieren. Daran anknüpfend beleuchtet Ulrich Trautwein mit dem fachlichen Hintergrund der Pädagogischen Psychologie und empirischen Erziehungswissenschaft die Frage, ob und wie das gegliederte Schulsystem zur sozialen Ungleichheit beiträgt. Trautwein greift dabei auf eigene Studien zurück, die sich mit dem Übergang von der Grundschule in die nach Leistungen stratifizierte Sekundarstufe beschäftigen. Aus der Ökonomie kommend stellt C. Katharina Spieß die Frage nach den Kosten versäumter Bildungschancen. Spieß geht dabei insbesondere auf Studien im Bereich der frühkindlichen Bildung ein. Aus einer erziehungswissenschaftlichen Perspektive fragt Ewald Terhart, ob sich die Richtigen für ein Lehramtsstudium entscheiden. Terhart geht dabei unter anderem auf die Fragen ein, welches wünschenswerte Merkmale von Lehramtsstudierenden sind und ob sich im Lehrerberuf mehr soziale Aufsteiger finden als in anderen Berufen. Abschließend wird Birgit Spinath die Beiträge diskutieren und dabei insbesondere die Verschränkungen der Disziplinen beleuchten.

Spinath, B. (Hrsg.) (in Druck). Empirische Bildungsforschung: Aktuelle Themen der Bildungspraxis und Bildungsforschung. Heidelberg: Springer.

**GEBF** TAGUNG  
2014

ÖF4

Öffentliches  
Forum

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
Casino-Gebäude  
Raum 1.811

2

**Gesellschaftlich relevante Fragen der Bildungsforschung und Bildungspraxis**



## Wie können wir die Ursachen sozialer Ungleichheit verstehen?

Soziale Ungleichheit ist aufgrund der wachsenden Unterschiede zwischen Bevölkerungsgruppen als gesellschaftliches Problem aktueller denn je. Ungleichheit muss per se noch kein Problem sein. Nicht alle strukturierten Ungleichheiten werden als problematisch wahrgenommen. Ausgehend von strukturellen und kulturellen Kontexten auf der Ebene von Beziehungen, Organisationen und Gesellschaften können verschiedene Mechanismen der Genese sozialer Ungleichheiten untersucht werden, zum Beispiel Ausbeutung, soziale Schließung oder Diskriminierung. Diese Mechanismen zu entschlüsseln ist eine der klassischen Domänen der Soziologie (z. B. Diewald, 2010a; Diewald & Faist, 2011).

Mittlerweile hat sich die Soziologie – wenn auch noch sehr vorsichtig – in Richtung Verhaltensgenetik geöffnet (Diewald, 2010b). Dabei geht es nicht mehr um Anlage oder Umwelt, sondern darum, wie genetische und soziale Bedingungen über den Lebenslauf zusammenwirken. Psychologische Merkmale und Mechanismen spielen in diesem Wechselspiel eine große Rolle um zu verstehen, wie aus dieser Interaktion Ungleichheiten entstehen. Deshalb erweitert selbst dann, wenn man als Soziologe vorrangig an den Auswirkungen struktureller und kultureller Bedingungen interessiert ist, eine transdisziplinäre Zusammenarbeit mit Psychologie und Verhaltensgenetik deren Verständnis. Tatsächlich ist inzwischen nicht nur für wichtige psychische Merkmale wie Intelligenz, Persönlichkeit und Motivation eine erhebliche genetische Mitbedingtheit nachgewiesen worden, sondern auch für Bildungsabschlüsse, Sozialkapital, Einkommen und Vermögen.

Ein neues, langfristig angelegtes DFG-Forschungsprojekt mit dem Titel „Genetic and social causes of life chances“ (Leitung: Diewald, Riemann und F. M. Spinath) hat sich vorgenommen, die Ursachen sozialer Ungleichheit unter Berücksichtigung genetischer Informationen und psychologischer Prozesse transdisziplinär zu untersuchen. Der Beitrag stellt die Anlage des Projektes sowie die theoretischen und methodischen Grundlagen der interdisziplinären Zusammenarbeit dar, wobei insbesondere der Gewinn einer solchen Zusammenarbeit aus Sicht der bisherigen soziologischen Ungleichheitsforschung im Vordergrund stehen soll (vgl. Diewald & Mayer, 2009).

Diewald, M. & Faist, T. (2011). Von Heterogenitäten zu Ungleichheiten: Soziale Mechanismen als Erklärungsansatz der Genese sozialer Ungleichheiten. *Berliner Journal für Soziologie*, 21, 91-114.

Diewald, M. (2010a). Ungleiche Verteilungen und ungleiche Chancen. Zur Entwicklung sozialer Ungleichheiten in der Bundesrepublik. In F. Faulbaum & C. Wolf. (Hrsg.), *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*. Schriftenreihe der ASI-Arbeitsgemeinschaft Sozialwissenschaftlicher Institute (S. 11–38). Wiesbaden: VS Verlag für Sozialwissenschaften.

Diewald, M. (2010b). Zur Bedeutung genetischer Variation für die soziologische Ungleichheitsforschung. *Zeitschrift für Soziologie*, 39, 4–21.

Diewald, M. & Mayer, K.U. (2009). The sociology of the life course and life span psychology: Integrated paradigm or complementing pathways? *Advances in Life Course Research*, 14, 5-14.

## **Wie stark trägt das gegliederte Schulsystem zur sozialen Ungleichheit bei?**

Einer der überraschendsten und für viele auch erschreckendsten Befunde der ersten PISA-Erhebung 2000 war, dass in keinem anderen Land der Zusammenhang zwischen schulischen Leistungen der Kinder und dem soziökonomischen Status des Elternhauses so eng war wie in Deutschland. Wenngleich die Assoziation von soziökonomischem Status und Schulleistung nicht per se ein Anzeichen von Ungerechtigkeit im Bildungssystem ist, so gilt dies doch für einen außerordentlich starken Zusammenhang. In der Folge galt es zu klären, was die Ursachen für diesen starken Zusammenhang waren. Es liegt nahe, dazu die Besonderheiten des deutschen Schulsystems, wie die Gegliedertheit nach Schultypen, die frühe Trennung der Schülerschaft nach der 4. Klasse sowie die Art der Übergangentscheidungen als Ursachen in den Blick zu nehmen.

Der vorliegende Beitrag fasst den aktuellen Forschungsstand zu der Frage zusammen, inwiefern das gegliederte Schulsystem zur sozialen Ungleichheit beiträgt. Dabei greift der Autor auf eigene Studien zurück, die sich mit den Effekten der Gliederung des Schulsystems in verschiedene Schultypen befasst (z. B. Becker et al., 2012), Effekten der Zusammensetzung der Schülerschaft (z. B. Dumont et al., 2013) sowie dem Beitrag der Übergangentscheidungen zur sozialen Ungerechtigkeit (Baeriswyl et al., 2006).

Baeriswyl, F., Wandeler, C., Trautwein, U. & Oswald, K. (2006). Leistungstest, Offenheit von Bildungsgängen und obligatorische Beratung der Eltern: Reduziert das Deutschfreiburger Übergangsmodell die Effekte des sozialen Hintergrunds bei Übergangentscheidungen? Zeitschrift für Erziehungswissenschaft, 9, 373–392.

Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking: Do academic-track schools make students smarter? Journal of Educational Psychology, 104, 682–699.

Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. Psychologie in Erziehung und Unterricht, 60, 163–183.

Dienstag,

04.03.2014,

16:30 - 18:00,

Casino-Gebäude

Raum 1.811



## ***Was sind die Kosten versäumter Bildungschancen?***

Bildung ist eine der wichtigsten Möglichkeiten zur Sicherung von individuellem und gesellschaftlichem Wohlstand. Die Bildungsökonomie beschäftigt sich unter anderem mit der Frage der Effizienz von Investitionen in Bildung. Welche Erträge fallen für den Einzelnen und die Gesellschaft an? Grundsätzlich geht die ökonomische Theorie der „Skill Formation“ davon aus, dass Investitionen am effizientesten sind, wenn sie bereits frühkindlich getätigt werden. Allerdings setzt dies eine hohe Qualität der Bildungsangebote voraus. Die grundsätzlich hohe Rendite frühkindlicher Bildungsinvestitionen trifft insbesondere für bildungsbenachteiligte Gruppen zu. Umso erstaunlicher ist es, dass Deutschland im internationalen Vergleich relativ wenig öffentliche Ressourcen in die frühkindliche Bildung investiert, auch wenn hier in den letzten Jahren größere Ausgabensteigerungen zu verzeichnen sind. Darüber hinaus sind in Deutschland insbesondere im sehr frühen Bildungsbereich sozioökonomische Unterschiede in der Nutzung dieser Angebote festzumachen - nicht alle Kinder sind in Kindertageseinrichtungen vertreten. Diese Unterschiede in der Nutzung sind nicht nur im Bereich der formalen Bildung, sondern auch bei non-formalen Bildungsangeboten zu beobachten. Dies kann Bildungsungleichheiten noch verstärken. Der vorliegende Beitrag greift diese und weitere Themen und Fragen auf. Die Autorin stellt ihre eigenen Studien zur Inanspruchnahme von frühkindlichen Bildungs- und Betreuungsangeboten (z. B. Schober & Spieß, 2012) vor und geht auf Gesichtspunkte der Effektivität und Effizienz von Investitionen in frühkindliche Bildung ein (z. B. Spieß, 2010; 2013).

Schober, P. & Spieß, C. K. (2012). Frühe Förderung und Betreuung von Kindern: Bedeutende Unterschiede bei der Inanspruchnahme besonders in den ersten Lebensjahren. DIW Wochenbericht, 43, 17–31.

Spieß, C. K. (2010). Sieben Ansatzpunkte für ein effektiveres und effizienteres System der frühkindlichen Bildung in Deutschland. In T. Apolte & U. Vollmer (Hrsg.), Bildungsökonomik und Soziale Marktwirtschaft (S. 3–18). Stuttgart: Lucius & Lucius.

Spieß, C. K. (2013). Effizienzanalysen frühkindlicher Bildungs- und Betreuungsprogramme: Das Beispiel von Kosten-Nutzen-Analysen. Zeitschrift für Erziehungswissenschaft, 16, 333–354.

## ***Entscheiden sich die Richtigen für ein Lehramtsstudium – und wer sind die Richtigen?***

„Auf die Lehrerinnen und Lehrer kommt es an“ ist ein Credo, das auch durch die empirische Bildungsforschung gut gestützt wird. Genauso verbreitet ist die Wahrnehmung, dass sich in Deutschland nicht immer die Richtigen für ein Lehramtsstudium entscheiden und somit viele Lehrerinnen und Lehrer für ihren Beruf nicht oder nicht optimal geeignet sind. Ist das tatsächlich so? Wer entscheidet sich für ein Lehramtsstudium und warum? Ist der Lehrerberuf ein Beruf für soziale Aufsteiger? Sollte es eine stärkere Auswahl zu Studienbeginn geben und wenn ja, wie könnte diese aussehen? Kann zu diesem frühen Zeitpunkt – das Abitur wird nach der Verkürzung der Schulzeit heute schon im Alter von 17 Jahren erworben – schon festgestellt werden, wer die Eignung zum Lehrerberuf besitzt und wer nicht? Wie muss die Lehrerbildung reformiert werden, um Studierende auf die zukünftigen Anforderungen vorzubereiten? Warum ist die Lehrerbildung eine Dauerbaustelle?

In Bezug auf diese Fragen fasst der vorliegende Beitrag den Stand der empirischen Forschung im Bereich der Lehrerbildung und des Lehrerberufs zusammen (z. B. Terhart, 2012a,b; Terhart, Bennewitz & Rothland, 2011, Terhart et al., 2013). Ein besonderer Schwerpunkt liegt dabei auf dem Veränderungsbedarf für die Praxis der Lehrerbildung, für die der Autor Empfehlungen ableitet (z. B. Terhart, 2013).

Terhart, E. (2012a). Vom „pädagogischen Begleitstudium“ zu den „Bildungswissenschaften in der Lehrerbildung“: Themen und Trends. *Beiträge zur Lehrerbildung*, 30, 49–61.

Terhart, E. (2012b). Attraktivität des Lehrerberufs: Möglichkeiten und Grenzen der Beeinflussung. *Zeitschrift für Bildungsverwaltung*, 28 (2), 39 – 51.

Terhart, E. (2013). *Erziehungswissenschaft und Lehrerbildung*. Münster: Waxmann.

Terhart, E., Bennewitz, H. & Rothland, M. (Hrsg.) (2011). *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann.

Terhart, E., Schulze-Stocker, F., Holzberger, D. & KUniversität-Habenicht, O. (2013). Bildungswissenschaften in der Universitätslehrerbildung: Unterscheiden sich Absolventinnen und Absolventen mit Erstem Staatsexamen von Absolventinnen und Absolventen mit einem Master of Education-Abschluss? In A. Gehrman, B. Kranz, S. Pelzmann B. & A. Reinartz (Hrsg.), *Formation und Transformation der Lehrerbildung* (S. 224 – 238). Bad Heilbrunn: Klinkhardt.



Susanne Kuger (DIPF)

## ***Soziale Disparitäten in der frühkindlichen Bildung und Erziehung in England und Deutschland – ein interdisziplinäres Bild***

Chair(s): Susanne Kuger (DIPF)

Diskutant/in: Birgit Becker (Goethe-Universität Frankfurt)

Die Bildung, Erziehung und Betreuung in der frühen Kindheit erleben seit Jahren in vielen Industrienationen einen quantitativen Ausbau. Dies ist teilweise eine Folge von international replizierten Studienergebnissen, welche die Möglichkeit positiver Effekte eines Besuchs einer Kindertageseinrichtung auf die kindliche kognitive und sozial-emotionale Entwicklung beschreiben (Anders, 2013). Besonders Kindern, die ein Risiko für ungünstige Bildungskarrieren tragen, möchte man frühzeitig Unterstützung und Förderung zukommen lassen, um späteren Disparitäten entgegenzuwirken (McCartney et al., 2007). Jedoch weisen diese Studien zuweilen auf geringe oder ausbleibende Effekte. Dabei fällt insbesondere in der angloamerikanischen Literatur auf, dass die Heterogenität der Ergebnisse zuweilen mit Hintergrundmerkmalen der Kinder kovariiert (NICHD & Duncan, 2003). Das Symposium geht diesen Assoziationen aus interdisziplinärer Sicht nach. Es werden ökonomische, soziologische, erziehungswissenschaftliche und psychologische Perspektiven vereint, die gemeinsam einen Blick auf soziale Disparitäten in der Nutzung, der erfahrenen Qualität, der differenziellen Wirksamkeit und Grenzen der Interventionsmöglichkeiten werfen.

Der erste Beitrag schafft durch die Betrachtung der englischen Situation eine Verbindung zur internationalen Literaturlage, die gerne – zuweilen unreflektiert – als Begründung für die Einführung erweiterter Maßnahmen in Deutschland herangezogen wird. Die erste Studie analysiert dabei den Zugang und das Eintrittsalter von Kindern unterschiedlicher sozialer Hintergründe in frühkindliche Betreuung auf Basis der Millennium Cohort Study. Da die Forschung mittlerweile gut belegen kann, dass nicht der Besuch an sich, sondern vor allem die in den Einrichtungen vorgefundene pädagogische Qualität entwicklungsfördernd ist, schließt der zweite Beitrag mit einer Analyse unterschiedlicher Zugangswahrscheinlichkeiten zu guter Qualität von Kindern verschiedener sozialer Herkunft in einer deutschen Studie unmittelbar an. Sie prüfen zusätzlich die subgruppenspezifische Wirksamkeit von Kindergartenqualität. Inwiefern die Hoffnung, gute Qualität in den Einrichtungen könnte dabei ungünstige Anreizbedingungen in den Familien kompensieren, aufrecht erhalten werden kann, untersucht die dritte Studie. Hier werden mittelfristig kindliche Leseleistungsergebnisse in der Grundschule mit den frühkindlich erlebten Lernumwelten in Familie und Kindergarten in Verbindung gebracht. Zum Abschluss wird von den Auswirkungen eines lokalen frühpädagogischen Interventionsprogramms berichtet, welches die Anreizqualität speziell in bildungsfernen Familien erhöhen und diese animieren soll, ihre Kinder frühzeitig in eine Kindertagesstätte zu geben.

Das Symposium bietet einen interdisziplinären Blick auf soziale Disparitäten in der Nutzung und in den Ergebnissen frühkindlicher Bildung und Erziehung. In der Diskussion wird gezielt die angesprochen, inwiefern die Ergebnisse mit Befunden zu sozialen Disparitäten in anderen Altersgruppen übereinstimmen und welchen Mehrwert die interdisziplinäre Bearbeitung des Themas birgt.

Anders, Y. (2013): Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. In: Zeitschrift für Erziehungswissenschaft, 16 (2), 237-275.

McCartney, K., Dearing, E., Taylor, B. A. & Bub, K. L. (2007). Quality child care supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology*, 28 (5-6), 411-426.

NICHD Early Child Care Research Network & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development*, 74(5), 1454-1475.

**GEBF** TAGUNG  
2014

ÖF5

Öffentliches  
Forum

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
PEG-Gebäude  
Raum 1.G191

2

Soziale Disparitäten in der frühkindlichen Bildung und Erziehung



## ***Soziale Ungleichheit beim Eintritt in frühkindliche Betreuung und kindliche Entwicklung – Analysen auf der Basis der Millennium Cohort Study***

### Theoretischer Hintergrund

Obwohl in den letzten Jahren die Beteiligung in frühkindlicher Bildung und Erziehung in fast allen OECD Ländern gestiegen ist, besuchen Kinder aus benachteiligten Familien nach wie vor weniger häufig eine Kindertageseinrichtung als gleichaltrige Kinder aus privilegierteren Familien. Studien zu frühkindlicher Bildung und Betreuung zeigen, dass der Besuch einer Kindertageseinrichtung, insbesondere bei hoher Qualität, einen positiven Einfluss auf die Entwicklung und Leistungen von Kindern aus benachteiligten Elternhäusern haben kann. Die Politik hat auf die Ergebnisse der frühkindlichen Bildungsforschung reagiert und Reformen verabschiedet, die den Zugang zu Kinderbetreuung erleichtern sollen. Dennoch bestehen weiterhin Unterschiede in der kindlichen Entwicklung vergleicht man die kognitive und sozial-emotionale Entwicklung für unterschiedliche Betreuungssettings. Darüber hinaus sind das Eintrittsalter, die Betreuungsdauer und die Qualität entscheidend für die kindliche Entwicklung.

### Fragestellung und Methode

Der vorliegende Beitrag untersucht den Einfluss des Eintrittsalters auf die sozial-emotionale Entwicklung von Kindern der britischen Millennium Cohort Study. Wir vergleichen Kinder, deren Eintrittsalter in frühkindliche Betreuung variiert. Basierend auf einer großen und repräsentativen Stichprobe (N=7.904) analysieren wir mit einem quasi-experimentellen Schätzverfahren den Effekt des Eintrittsalters in eine Kindertageseinrichtung auf die sozial-emotionale Entwicklung. Wir instrumentieren hierzu den Besuch einer Kindertageseinrichtung im Alter von drei Jahren mit dem Anteil der Kinder in England, die einen Anspruch auf kostenlose Kindertagesbetreuung im Alter von drei Jahren haben.

### Ergebnisse und Diskussion

Für alle Kinder der Stichprobe zeigen sich in den Ergebnissen, dass ein Eintritt nach dem dritten Lebensjahr in die Kindertagesbetreuung die sozial-emotionalen Probleme erhöht im Vergleich zu Gleichaltrigen, die zu einem früheren Zeitpunkt in die Kindertageseinrichtung eintreten. Für Kinder, die in Niedrigeinkommensfamilien aufwachsen, verstärken sich die gefundenen Effekte des späten Eintrittsalters noch. Unsere Ergebnisse ergänzen die bisherige Literatur zum einen durch robustere Effekte des verwendeten Instrumentalvariablenansatzes. Und zum andern untersuchen wir den Einfluss des Eintrittszeitpunktes auf die kindliche Entwicklung. Unsere Ergebnisse zeigen, dass politische Entscheidungsträger die soziale Ungleichheit beim Eintritt in frühkindliche Betreuung berücksichtigen und einen früheren Zugang als bisher ermöglichen sollten.

## **Zugang zu und Effekte der Kindergartenqualität: Die Bedeutung des sozialen Hintergrundes**

### Theoretischer Hintergrund

Während andere Länder schon lange auf die Ausweitung frühkindlicher Bildungsangebote setzen um z.B. armutsbedingte Kompetenzunterschiede abzuschwächen (z.B. Head Start in den USA seit 1965), ist erst in den letzten Jahren auch in Deutschland das mögliche Potenzial vorschulischer, außerfamiliärer Betreuungseinrichtungen wieder in den Fokus des wissenschaftlichen und gesellschaftlichen Interesses gerückt. Obwohl fast alle Kinder in Deutschland zumindest im letzten Jahr vor der Einschulung einen Kindergarten besuchen, ist bislang unklar, ob alle Kinder die gleichen Chancen haben, einen Kindergarten mit guter Qualität zu besuchen.

### Fragestellung und Datengrundlage

Der vorliegende Beitrag geht daher der Frage nach, ob der Zugang zu einem Kindergarten höherer Qualität nach sozialer Herkunft selektiv ist und ob sich die Kindergartenqualität je nach sozialer Herkunft als differentiell bedeutsam für die Entwicklung mathematischer Kompetenzen erweist. Dazu wurden Daten der BiKS-Studie, welche 547 Kinder in 97 Kindergärten untersucht, genutzt. Die Kindergartenqualität wurde durch geschulte Beobachter in einer vierstündigen Beobachtungssituation mit Hilfe zweier Beobachtungsinstrumente (KES-R und KES-E) gemessen. Zusätzlich wurden Merkmale der sozialen Herkunft der Familie über einen standardisierten Fragebogen bzw. Interview erfasst. Zur Messung der mathematischen Kompetenzen der Kinder im Alter von 3,4 und 5 Jahren wurde die Skala „Arithmetik“ der „Kaufman Assessment Battery for Children (K-ABC)“ genutzt.

### Ergebnisse und Diskussion

Im Rahmen logistischer Regressionsanalysen wird gezeigt, dass Kinder mit Migrationshintergrund eine geringere Wahrscheinlichkeit aufweisen, einen Kindergarten höherer Qualität zu besuchen – der sozioökonomischen Status (SES) der Familie sowie der mütterliche Bildungshintergrund zeigen keine Effekte. Darüber hinaus wurden latente Wachstumskurvenmodelle spezifiziert, die die mathematische Entwicklung im Alter zwischen 3 und 5 Jahren in einem Multigruppensdesign je nach sozialer Herkunft untersuchen. Die Ergebnisse zeigen einen positiven Effekt der Kindergartenqualität auf den Zuwachs der mathematischen Kompetenzen für Kinder ohne Migrationshintergrund und für Kinder deren Eltern einen höheren Bildungsabschluss erlangten und einen höheren sozioökonomischen Status aufweisen. Der Beitrag unterstreicht die Bedeutsamkeit qualitativ hochwertiger Kindergartenbetreuung für die kindliche Kompetenzentwicklung und diskutiert vor dem Hintergrund des Gedankens der Kompensation, inwiefern dies auch für Kinder aus benachteiligten Verhältnissen zutrifft.

## **Wechselwirkungen vorschulischer Erfahrungen in Kindergarten und Familie und ihre Bedeutung für das Lesen im Grundschulalter**

### Theoretischer Hintergrund

Eine breite empirische Forschungsbasis verweist mittlerweile auf die Bedeutung von „Vorläuferfertigkeiten“ für den späteren Leseerwerb (Ennemoser et al., 2012). In Diskussionen um die Anregung und Förderung dieser Vorläuferfertigkeiten rücken daher die beiden wichtigsten Lernumwelten vor der Schulzeit, Familie und Kindergarten, in den Fokus der Wissenschaft, Administration und Bildungspraxis. Für beide Anregungssettings liegen Belege für die Bedeutung qualitativ hochwertiger Anregung für die Leistungsentwicklung des Kindes vor (Anders, 2013; Melhuish et al., 2008; Sénéchal, 2006; Tietze, Roßbach, & Grenner, 2005). Dabei berücksichtigen diese Studien die beiden Lernumwelten meist als getrennte Entitäten. Ein solches Vorgehen berücksichtigt dabei nicht, dass zwischen ihnen direkt oder indirekt Wechselwirkungen stattfinden können. Eine solche gegenseitige Beeinflussung der beiden Lernumwelten und denkbaren Kombinationsmöglichkeiten ihrer Anregung des Kindes ist jedoch zentral. Zum einen wäre möglicherweise erkennbar, ob und inwiefern sich die institutionelle und die familiäre Anregungsumwelten aufeinander einstellen können und zum anderen könnten die Ergebnisse zeigen, ob Kindergärten schlechte Anregungsbedingungen zu Hause kompensieren können, ob sich die Effekte der beiden Lernumwelten addieren, oder ob sich speziell Kinder mit guten Anregungsbedingungen in beiden Lernumwelten besonders positiv entwickeln.

### Fragestellung und Methode

Der Beitrag geht eben dieser Frage nach, wie die Literacyanregung der beiden frühkindlichen Lernumwelten Kindergarten und Familie während der drei Jahre des Kindergartenbesuchs eines Kindes miteinander in Beziehung stehen und welche eigenständige und gemeinsame Bedeutung sie für die Prädiktion des späteren Leseverständnisses des Kindes haben. Dazu wurden Daten von 121 Kindern aus 51 Kindergärten, ihren Familien und den entsprechenden Einrichtungen der Längsschnittstudie BiKS-3-10 analysiert. Die Analyse berücksichtigt Daten zur formellen und informellen Literacyanregung (Sénéchal, LeFevre, Thomas, & Daley, 1998; Conner, Morrison, & Slominski, 2006) jeweils in beiden Lernumwelten Kindergarten und Familie im ersten und im letzten Kindergartenjahr sowie des Leseverständnisses der Kinder in der 2. Klasse (im Alter von etwa acht Jahren). Zum Einsatz kommen Strukturgleichungsmodelle und multiple lineare Regressionsanalysen.

### Ergebnisse

Die Ergebnisse zeigen bei moderater Stabilität der Qualitätsindikatoren weder quer- noch längsschnittlich relevante Beziehungen der Anregungsniveaus der Lernumwelten miteinander. Die beiden Lernumwelten scheinen damit eher neben- denn miteinander zu existieren und zeigen keine wechselseitigen Bezüge. Bezüglich der Leistungsentwicklung der Kinder steht ein später besseres Leseverständnis insbesondere in Beziehung mit früher hochwertiger institutioneller Literacyanregung, weniger mit der späteren Qualität. Keine Beziehung zu den Leseleistungen in der zweiten Klasse beweisen dabei die familialen Lernbedingungen während der Kindergartenzeit. Es zeigen sich differenzielle Ergebnisse, die im Sinne eines Matthäuseffekts interpretiert werden müssen, da Kinder mit eher besseren Bedingungen zuhause und eher besseren Bedingungen im Kindergarten von dieser Anregung ganz besonders in ihrer Leistungsentwicklung zu profitieren scheinen.

### Diskussion

Der Beitrag illustriert damit die derzeit trotz üblicherweise längerfristiger Auseinandersetzung mit einem Kind, seiner Entwicklung und seinen familialen Lebensverhältnissen eingeschränkten Möglichkeiten frühpädagogischer Einrichtungen auf die familiäre Lernumwelt zu reagieren. Es ist anscheinend kaum möglich, fehlende familiäre Anregungen zu kompensieren oder zumindest gleiche Entwicklungsgeschwindigkeiten auf insgesamt unterschiedlichem Niveau zu ermöglichen. Der Beitrag betont damit die Notwendigkeit des Ausbaus der Zusammenarbeit zwischen frühpädagogischen Einrichtungen und Familien. Zugleich deutet er an, dass die derzeitigen Anstrengungen zur Kompensation sozialer Disparitäten in der kindlichen Entwicklung durch die institutionelle Förderung noch deutlich verbessert und ausgebaut werden müssen um in bedeutsamem Umfang entwicklungswirksam zu sein.

Anders, Y. (2013): Stichwort: Auswirkungen frühkindlicher, institutioneller Bildung und Betreuung. Zeitschrift für Erziehungswissenschaft, 16, 2, 237-275.

Connor, C. M./ Morrison, F. J./ Slominski, L. (2006): Preschool Instruction and Children's Emergent Literacy Growth. Journal of Educational Psychology, 98, 4, S. 665-689.

Ennemoser, M./ Marx, P./ Weber, J./ Schneider, W. (2012): Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens – Evidenz aus zwei Längsschnittstudien vom Kindergarten bis zur 4. Klasse. Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie, 44, 2, S. 53-67.

Melhuish, E. C./ Phan, M. B./ Sylva, K./ Sammons, P./ Siraj-Blatchford, I./ Taggart, B. (2008): Effects of home learning environment and preschool centre experience upon literacy and numeracy development in early primary school. Journal of Social Issues, 64, 1, S. 95-114.

Sénéchal, M. (2006): Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. Journal for the Scientific Study of Reading, 10, 1, S. 59-87.

Sénéchal, M./ LeFevre, J.-A. / Thomas, E./ Daley, K. (1998): Differential effects of home literacy experiences on the development of oral and written language. Reading Research Quarterly, 33, 1, S. 96-116.

Tietze, W./ Roßbach, H.-G./ Grenner, K. (2005): Kinder von 4 bis 8 Jahren. Zur Qualität der Erziehung und Bildung in Kindergarten, Grundschule und Familie. – Weinheim.

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
PEG-Gebäude  
Raum 1.G191



## **Familiäre Anregungsqualität und Erziehungsvorstellungen als Zielvariablen frühpädagogischer Modellprogramme: Differenzielle Effekte in Abhängigkeit vom sozialen Hintergrund?**

### Theoretischer Hintergrund

Kinder unterscheiden sich bereits zu Beginn der Grundschulzeit stark in ihren kognitiven Kompetenzen (z.B. Magnuson et al., 2004). Diese Unterschiede werden oftmals über die weiteren Schuljahre fortgetragen (Sammons et al., 2008). Zudem hängt der schulische Bildungserfolg von Kindern in Deutschland stärker als in anderen Ländern von ihrer sozialen und kulturellen Herkunft ab (OECD 2004a, 2004b). Darüber hinaus weisen viele Kinder aus sozial benachteiligten Familien eine geringe Anregungsqualität in der Familie auf (Melhuish et al., 2008). Vor diesem Hintergrund wird das Potenzial vorschulischer Bildungsprogramme diskutiert, insbesondere auch zur Kompensation der Benachteiligung von Kindern aus sozial schwachen Familien oder Kindern mit Migrationshintergrund. In den USA wurden Modellprogramme wie das Perry Preschool Project (Schweinhart, 1993) mit dieser Zielsetzung konzipiert und in zum Teil umfassenden Evaluationsstudien auf ihre Wirkungen hin analysiert. Die Resultate sind beeindruckend: Gerade vorschulische Programme, in denen auch die Familien intensiv einbezogen wurden, zeigen noch im Erwachsenenalter Vorteile (z.B. Schweinhart et al., 2005).

Auch in Deutschland wurden Ansätze konzipiert und erprobt. Bislang ist allerdings noch ein großer Forschungsbedarf im Hinblick auf die Auswirkungen solcher Modellprojekte festzustellen. Ein Ansatz, der in Deutschland entwickelt wurde, ist das Modellprojekt „Chancenreich“, welches ein Angebot an alle Familien zur Unterstützung bei der Pflege und Erziehung ihrer Neugeborenen und Kleinkinder in der Stadt Herford darstellt. Insbesondere sollen bildungsferne Familien erreicht werden. Im Rahmen des Modellprojekts werden verschiedene Module umgesetzt (z.B. Familienbesuche, die Bereitstellung eines Elternhandbuchs, Elterntraining). Ferner wird ein frühzeitiger Kindergartenbesuch angestrebt. Die Eltern erhalten einen Bonus von 500 EUR, wenn sie verschiedene Module abdecken und ihr Kind bis zum 3. Lebensjahr in einer Kindertagesstätte anmelden. Durch den intensiven und frühen Zugang zu den Familien bietet das Modellprojekt theoretisch sehr gute Voraussetzungen, die Erziehungskompetenz der Eltern zu steigern und über unterschiedliche Zugangswege zu einer positiven Entwicklung der Kinder beizutragen.

### Fragestellung

Das Projekt AQua Fam („Ansätze zur Erhöhung der familialen Anregungsqualität“) untersucht die Auswirkungen der Teilnahme am Projekt „Chancenreich“ auf das Erleben und Erziehungsverhalten der Eltern und auf die soziale und sprachliche Entwicklung der Kinder. In der vorgestellten Studie wird auf folgende Fragestellungen fokussiert (1) Welche soziostrukturellen Merkmale weisen Familien auf, die am Projekt „Chancenreich“ teilnehmen? (2) Welcher Zusammenhang besteht zwischen dem sozialen Hintergrund, verschiedenen Aspekten der familialen Anregungsqualität und Erziehungsstilen? (3) Welche differenziellen Zusammenhänge zeigen sich zwischen Erziehungsstilen, familialer Anregungsqualität und dem sozialen Entwicklungsstand der Kinder? (4) Welchen Einfluss hat die Teilnahme am Projekt „Chancenreich“ auf Erziehungsstile und den sozialen Entwicklungsstand der Kinder?

### Methode

Die Studie ist als Querschnittstudie mit zwei Untersuchungsgruppen – einer Interventionsgruppe und einer Kontrollgruppe – angelegt. Die Stichprobe der Interventionsgruppe umfasst 240 3-jährige Kinder und Familien, die am Modellprojekt seit der Geburt teilnehmen. Zusätzlich wird eine Kontrollgruppe von 80 3-Jährigen Kindern und Familien, die nicht am Programm teilgenommen haben, untersucht. Die Testungen und Befragungen der Kinder und Familien finden bei den Familien zu Hause statt. Diese werden durch die dafür geschulten Datenerheberinnen durchgeführt. Die Zielindikatoren in dieser Studie (sozialer Hintergrund der Familie, Erziehungsstile, familiäre Anregungsqualität, sozialer Entwicklungsstand der Kinder) beruhen auf Beobachtungsdaten, einer Befragung der Eltern sowie einer Einschätzung des Entwicklungsstandes der Kinder durch die Erzieherin. Die Fragestellungen werden mit Hilfe von regressionsanalytischen Verfahren und Strukturgleichungsmodellen untersucht.

### Ergebnisse

Die Datenerhebung wird Ende des Jahres abgeschlossen, so dass noch keine finalen Resultate vorliegen. Erste Analysen der Stichprobe weisen jedoch darauf hin, dass das Projekt Chancenreich in hohem Maße auch von nicht bildungsbenachteiligten Familien in Anspruch genommen wird. Die Zusammenhänge zwischen sozialem Hintergrund, familialer Anregungsqualität, Erziehungsstilen und sozialem Entwicklungsstand werden diesbezüglich einer vertieften Analyse unterzogen. Insbesondere wird diskutiert welche differenziellen Potenziale familienorientierte Modellprojekte in Abhängigkeit vom sozialen Status haben können.

Magnuson, K., Meyers, M., Ruhm, R. & Waldfogel, J. (2004). Inequality in pre-school education and school readiness. *American Educational Research Journal*, 41, 115-157.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., & Phan, M. (2008). Effects of home learning environment and preschool center experience upon literacy and numeracy in early primary school. *Journal of Social Issues*, 64, 95-114.

OECD (2004a). *Learning for tomorrow's world. First results from PISA 2003*. Paris, France: OECD.

OECD (2004b). *Messages from PISA 2000*. Retrieved from

<http://www.pisa.oecd.org/dataoecd/31/19/34107978.pdf>

Sammons, P., Anders, Y., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Barreau, S. (2008). Children's cognitive attainment and progress in English primary schools during Key Stage 2.: Investigating the potential continuing influences of pre-school education. *Zeitschrift für Erziehungswissenschaft, Sonderheft 11*, 179-198.

Schweinhart, L. J., Barnes, H. V. & Weikart, D. P. (1993). *Significant Benefits. The High/Scope Perry Preschool Study Through Age 27*. Ypsilanti, MI: High/Scope Press.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R. & Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Ypsilanti, MI: High/Scope Press.



Marko Neumann (DIPF), Kai Maaz (DIPF)

## ***Der Einfluss von schulstrukturellen Rahmenbedingungen auf schulische Leistungen und die psychosoziale Entwicklung***

Chair(s): Marko Neumann (DIPF), Kai Maaz (DIPF)

Diskutant/in: Rainer Watermann (Freie Universität Berlin)

Die Bedeutung schulstruktureller Rahmenbedingungen für die Leistungs- und psychosoziale Entwicklung von Schülerinnen und Schülern ist seit langem Gegenstand der Bildungsforschung. Eine zentrale Rolle spielt dabei die Frage nach den Auswirkungen unterschiedlicher Formen der schulischen Leistungsdifferenzierung. Für das deutsche Schulsystem sind hierbei unter anderem die Gliederung der Sekundarstufe I in verschiedene Schulformen, aber auch die Schaffung homogener Lerngruppen innerhalb von Schulen – etwa in Form von Kursen auf unterschiedlichem Anforderungsniveau in der gymnasialen Oberstufe – zu nennen. Bezüglich beider Formen der Leistungsdifferenzierung haben sich in den vergangenen Jahren substantielle Veränderungen vollzogen. So hat der überwiegende Teil der Bundesländer inzwischen in der Sekundarstufe I auf ein zweigliedriges System bzw. ein Zwei-Säulen-Modell umgestellt, das neben dem Gymnasium im Wesentlichen nur noch eine weiterführende Schulform vorsieht, die zum Haupt- und mittleren Schulabschluss führt, in mehreren Bundesländern aber auch den direkten Weg zum Abitur ermöglicht. Auch mit Blick auf die inhaltlich-organisatorische Ausgestaltung der gymnasialen Oberstufe kam es in den letzten Jahren zu deutlichen Veränderungen, was sich unter anderem darin manifestiert, dass in vielen Bundesländern die Differenzierung in Grund- und Leistungskurse für die Kernfächer Deutsch, Mathematik und Fremdsprache aufgehoben wurde.

Die Beiträge des Symposiums untersuchen die Auswirkungen (veränderter) schulstruktureller Rahmenbedingungen auf die schulischen Leistungen und die psychosoziale Entwicklung der Schülerinnen und Schüler. Friederike Zimmermann, Kerstin Schütte und Olaf Köller gehen in ihrem Beitrag der Frage nach, inwieweit die Umstellung von der Dreigliedrigkeit auf die Zweigliedrigkeit in der Sekundarstufe I im Bundesland Bremen zu Veränderungen in den Lernständen und in der Leistungsentwicklung in Mathematik und im Leseverständnis führte. Der Beitrag von Bettina Sonnenberg, Wolfgang Wagner, Jochen Kramer, Karin Berendes und Ulrich Trautwein untersucht die Auswirkungen der Reform der gymnasialen Oberstufe im Bundesland Thüringen auf die Fachleistungen der Abiturienten in Mathematik, Englisch, Physik und Biologie. Im Zentrum des Beitrags von Henrike Knoppick, Michael Becker, Marko Neumann und Kai Maaz steht die Entwicklung des schulischen und allgemeinen Wohlbefindens der Schülerinnen und Schüler beim Übergang von der Grundschule in die neu strukturierte (zweigliedrige) Sekundarstufe in Berlin. Der Fokus liegt auf der moderierenden Rolle der besuchten Schulform (Gymnasium vs. Integrierte Sekundarschule) für die Wohlbefindensentwicklung. Hanna Dumont, Anna K. Chmielewski und Ulrich Trautwein nehmen in ihrem Beitrag eine international vergleichende Perspektive ein und untersuchen die Auswirkung unterschiedlicher Formen der Leistungsgruppierung auf das akademische Selbstkonzept. Die Beiträge werden von Rainer Watermann diskutiert.

**GEBF** TAGUNG  
2014

ÖF6

Öffentliches  
Forum

Dienstag,  
04.03.2014,  
16:30 - 18:00,  
PEG-Gebäude  
Raum 1.G168

Der Einfluss von schulstrukturellen Rahmenbedingungen

2



## **Strukturreformen im Bildungssystem: Allheilmittel für Schülerinnen und Schüler?**

Die Zeit nach PISA 2000 ist in vielen Ländern der Bundesrepublik Deutschland genutzt worden, um die Schulstrukturen des allgemeinbildenden Systems in der Sekundarstufe I zu ändern. Im Kern haben mittlerweile fast alle Länder auf ein Zwei-Säulen-Modell umgestellt, in dem neben dem Gymnasium lediglich eine weitere Schulform existiert, die vielfach auch zum Abitur führt. Die Fusionierung des nichtgymnasialen Bereiches geschah wenigstens aus zwei Gründen, (1) den sinkenden Schülerzahlen und (2) dem sehr schwachen Abschneiden deutschen Hauptschülerinnen und -schüler in Schulleistungsstudien. Inwieweit solche Schulstrukturreformen die intendierten Folgen haben (z. B. höhere schulische Leistungen in den neuen Strukturen), ist eine offene, gleichzeitig nur empirisch zu beantwortende Frage.

Folgt man der Argumentation, dass vor allem sehr schwache Schülerinnen und Schüler davon leistungsmäßig profitieren sollten, wenn sie mit leistungsstärkeren Klassenkameraden gemeinsam unterrichtet werden, so lassen sich folgende Hypothesen ableiten:

1.) In einem zweigliedrigen System, das neben dem Gymnasium die Haupt- und die Realschule zu einer gemeinsamen Schulform fusioniert, sollten sich (nach Kontrolle der kognitiven Grundfähigkeiten, des sozioökonomischen Status und des Migrationshintergrundes) höhere Schulleistungen gegen Ende der Sekundarstufe I zeigen als im klassischen dreigliedrigen System.

2.) Diese Vorteile sollten im Wesentlichen durch einen Anstieg der Leistungen bei den schwachen Schülerinnen und Schülern erklärbar sein. Zur Überprüfung dieser Hypothesen nutzen wir in der vorliegenden Studie die Daten einer Mehrkohorten-Längsschnittstudie in Bremen, in die Schülerinnen und Schüler aus 11 Schulzentren eingebunden waren. Während zwei Kohorten (N1 = 1476; N2 = 1342) noch im dreigliedrigen System mit Gymnasium, Real- und Hauptschule unterrichtet wurden, wurde die dritte Kohorte (N = 1001) in einem zweigliedrigen System mit Sekundarschule und Gymnasium geführt. In allen drei Kohorten lag der Gymnasialanteil bei rund einem Viertel. Für die drei Kohorten liegen querschnittliche (Kohorte 1) und längsschnittliche (Kohorten 2 und 3) Leistungsdaten für die Fächer Deutsch (Lesen) und Mathematik vor. Darüber hinaus liegen auch Werte zu den kognitiven Grundfähigkeiten, der sozialen Herkunft und dem Migrationsstatus der Schülerinnen und Schüler vor. Die Analysen in der 9. Jahrgangsstufe zeigen zunächst, dass sich die drei berücksichtigten Kohorten nicht signifikant hinsichtlich ihrer kognitiven Grundfähigkeiten, ihres sozioökonomischen Status und ihres Migrationshintergrundes unterscheiden. Für die Leistungsstände im Lesen und in Mathematik ergeben sich insgesamt ernüchternde Befunde. Weder für die Lesekompetenz noch die Mathematikleistungen zeigen sich Unterschiede zwischen den drei Kohorten. Auch profitiert der untere Bereich der Leistungsverteilung nicht von der Beschulung im zweigliedrigen System. Längsschnittliche Analysen der Kohorten 2 und 3 (von der 7. zur 9. Jahrgangsstufe) ergeben weiterhin, dass auch die Zuwächse über zwei Schuljahre unabhängig von der Schulstruktur (Zwei- vs. Dreigliedrigkeit) sind. Vor dem Hintergrund dieser Befunde und der rezenten Literatur (z. B. Hattie, 2009) werden die Grenzen von Strukturreformen diskutiert.

Bettina Sonnenberg (Eberhard Karls Universität Tübingen), Wolfgang Wagner (Eberhard Karls Universität Tübingen),  
Jochen Kramer (Eberhard Karls Universität Tübingen), Karin Berendes (Eberhard Karls Universität Tübingen),  
Ulrich Trautwein (Eberhard Karls Universität Tübingen)

## **Effekte der Oberstufenreform in Thüringen**

In den vergangenen Jahrzehnten sind in fast allen Bundesländern die gymnasialen Oberstufen reformiert worden (Trautwein et al., 2010; Wagner et al., 2011). Die wesentliche Zielsetzung dieser Reformen war die Sicherstellung der Studierfähigkeit (und Ausbildungsfähigkeit) der Schülerinnen und Schüler in Hinblick auf vertiefte mathematisch-naturwissenschaftliche sowie deutsch- und fremdsprachliche Kompetenzen. Mit der Einführung von Kernfächern mit einheitlichen Unterrichtszeiten einerseits sowie mit Belegungs- und Prüfungspflichten von naturwissenschaftlichen und fremdsprachlichen Fächern andererseits war die Rückorientierung zu einem Kerncurriculum wesentliches Element der Reformen, das vertiefte Kompetenzen aller Schülerinnen und Schüler gewährleisten sollte.

Empirische Untersuchungen zu den Effekten der Reformen auf die Schülerleistungen, im Besonderen im naturwissenschaftlichen Bereich, sind bisher nur eingeschränkt vorhanden und zeichnen ein uneinheitliches Bild. Einige der wenigen Studien verweisen auf eine hohe Stabilität der Leistungen und Leistungsvarianz von Schülerinnen und Schülern, während einige Ergebnisse auf eine Leistungsverbesserung sowie eine Verringerung der Leistungsstreuung als Folge der Einschränkung von Wahlmöglichkeiten in der Oberstufe hindeuten (Jonkmann et al., 2010; Nagy et al., 2010; Neumann et al., 2010; Wagner et al., in Vorbereitung). Es ist demnach weiterhin eine offene Frage, inwieweit die Reduktion der Kurswahl- und Prüfungsfachmöglichkeiten zugunsten einer stärkeren Fokussierung auf ein Kerncurriculum die zentralen Ziele der Reform einlösen konnte und zu einer Erhöhung der Leistungen auf breiter Basis und gleichzeitig zu einer Angleichung der Leistungen geführt hat.

Der Beitrag greift diese Forschungslücke auf, indem der Frage nachgegangen wird, wie sich die Verringerung der Wahlmöglichkeiten in der gymnasialen Oberstufenreform in Thüringen im Jahr 2009 auf die Schülerleistungen von Zwölfklässlern in den Fächern Mathematik, Englisch, Physik und Biologie auswirkt. Die Reform beinhaltete einerseits die Einführung von einheitlich vierstündigen Kernfächern, die in der Stundenanzahl eher dem bisherigen Grundkursniveau (drei oder vier Wochenstunden) als dem Leistungskursniveau (sechs Wochenstunden) entsprechen. Andererseits implementierte die Reform fünf statt vier Prüfungsfächer und weitete die Belegungspflicht von fremdsprachlichen und naturwissenschaftlichen Fächern aus, womit die Möglichkeit zur Spezialisierung durch Abwahl von Kursen eingeschränkt wurde (Thüringer Kultusministerium, 2008).

Auf Basis der Zusatzstudie des Nationalen Bildungspanels zur Oberstufenreform in Thüringen (Raab, Skopek, & Zielonka, 2012) und der Verwendung eines Kohorten-Kontroll-Designs werden die Kompetenzen der letzten Kohorte vor der Reform im Abschlussjahrgang 2010 und der ersten Kohorte nach der Reform im Abschlussjahrgang 2011 untersucht. Dabei werden die mittleren Leistungen der Schülerinnen und Schüler sowie die Leistungsstreuung insgesamt und differenziert für das obere und untere Verteilungsspektrum der kognitiven Grundfähigkeiten vor und nach der Reform verglichen.

In Übereinstimmung mit den Zielen der Reform zeigen die Daten zunächst eine Vereinheitlichung der Unterrichtszeit sowie signifikant geringere Abwahlquoten für Fächer wie Englisch und Physik nach der Reform. Trotz dieser zum Teil hochsignifikanten Unterschiede offenbaren die Ergebnisse für die mittleren Leistungen eine hohe Persistenz der Leistungen in Mathematik, Biologie und Physik. Während sich beispielsweise die Unterrichtszeit in Mathematik signifikant verringert hat, unterscheiden sich die Leistungen vor und nach der Reform nicht signifikant. Auch die Leistungen in Biologie und Physik unterscheiden sich nicht signifikant, obgleich der Anteil der Abwähler nach der Reform deutlich geringer ist. Für Englisch verweisen die Ergebnisse hingegen auf leicht signifikant höhere Leistungen nach der Reform bei Jungen. Dies deutet auf eine Erhöhung der Leistungen als Folge eines Anstiegs der Unterrichtszeit kombiniert mit einer häufigeren Belegung von Englisch als Prüfungsfach hin. Interessanterweise zeigen sich bei der Leistungsstreuung in den vier Fächern trotz der Reduktion der Wahlmöglichkeiten insgesamt keine signifikanten Angleichungen der Leistungen. Vor dem Hintergrund der Befunde wird die Reichweite von Reformen im Bildungswesen diskutiert.

Jonkmann, K., Trautwein, U., Nagy, G., & Köller, O. (2010). Fremdsprachenkenntnisse in Englisch vor und nach der Neuordnung der gymnasialen Oberstufe in Baden-Württemberg. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke & K. Maaz (Hrsg.), Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand (S. 181-214). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Nagy, G., Neumann, M., Trautwein, U., & Lüdtke, O. (2010). Voruniversitätsreife Mathematikleistungen vor und nach der Neuordnung der gymnasialen Oberstufe in Baden-Württemberg. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke & K. Maaz (Hrsg.), Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand (S. 147-180). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Neumann, M., Trautwein, U., Nagy, G., Lüdtke, O., & Maaz, K. (2010). Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand: Ein Zwischenresümee. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke & K. Maaz (Hrsg.), Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand (S. 267-280). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Thüringer Kultusministerium (Hrsg.) (2008). Informationen zur neuen gymnasialen Oberstufe. Weimar: Gutenberg Druckerei.

Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (2010). Institutionelle Reform und individuelle Entwicklung: Hintergrund und Fragestellungen der Studie TOSCA-Repeat. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke & K. Maaz (Hrsg.), Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand (S. 15-36). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Wagner, W., Kramer, J., Trautwein, U., Lüdtke, O., Nagy, G., Jonkmann, K., Maaz, K., Meixner, S. & Schilling, J. (2011). Upper secondary education in academic school tracks and the transition from school to postsecondary education and the job market. Zeitschrift für Erziehungswissenschaft, 14 (Sonderheft 14), 233-249.

Wagner, W., Norman, R., Dicke, A.L. & Rosenberger, F. (in Vorbereitung). Alle alles lehren: Schulleistungen und Englisch, Mathematik und den Naturwissenschaften vor uns nach der Neuordnung der gymnasialen Oberstufe in Sachsen.

Zielonka, M., Skopek, J. & Raab, M. (2012). Curricular Reform Study in Thuringia (TH) SUF Version 1.0.0. Data Manual. Research Data Papers. Bamberg: NEPS Data Center.



## ***Der Einfluss des Übergangs in differenzielle Lernumwelten auf das allgemeine und schulische Wohlbefinden von Kindern***

Das Wohlbefinden im Kindes- und Jugendalter unterliegt Veränderungsprozessen, die unter anderem mit Veränderungen der schulischen Lernumwelt in Verbindung gebracht werden. Vielfach wird der Übergang an die weiterführende Schule als kritisches Lebensereignis diskutiert, der mit Einbußen im Wohlbefinden und damit assoziierten Variablen des psychischen Erlebens verbunden sein kann (z.B. Chung, Elias, & Schneider, 1998). Allerdings konnten andere Autorinnen und Autoren keinen oder sogar einen positiven Effekt des Übergangs nachweisen (z.B. Proctor & Choi, 1994). Die Schulform der weiterführenden Schule wird dabei als Einflussfaktor auf die Bewältigung der Übergangserfahrung betrachtet. Hier sind sowohl Unterschiede in der Bezugsgruppe als auch kontextuelle Merkmale der neuen Lernumwelt (z.B. organisatorische oder curriculare Bedingungen) von Bedeutung. Bezüglich schulnaher Facetten des Wohlbefindens (z.B. Valtin & Wagner, 2004) hat sich das Gymnasium als eher nachteilige Lernumwelt erwiesen. Erklären lässt sich dies unter anderem durch Bezugsgruppeneffekte des leistungsstarken Lernmilieus. Van Ophuysen (2008) fand für die Schulfreude keinen negativen Effekt der veränderten Lernumwelt und begründet dies damit, dass der Wechsel an eine neue Schule mit positiven Erwartungen und einem damit assoziierten gesteigerten Wohlbefinden verbunden sein kann, so dass es an der neuen Schule auch zu einer Steigerung des Wohlbefindens kommen kann. Andere Autoren diskutieren außerdem Assimilationseffekte einer Integration in eine leistungsstarke Lerngruppe (z.B. Preckel & Brüll, 2010).

Die vorliegende Arbeit setzt an dieser Stelle an und untersucht anhand einer breiten Konzeptualisierung die psychosoziale Entwicklung im Übergangsprozess im neu strukturierten Berliner Sekundarschulwesen, das für den Übergang nach der Grundschule nur noch zwei weiterführende Schulformen vorsieht: das Gymnasium sowie die neu implementierte Integrierte Sekundarschule. Mit dem Ziel, ein möglichst umfassendes Bild zu erhalten, wird die Entwicklung und Veränderung des schulischen und des allgemeinen Wohlbefindens sowie des akademischen Selbstkonzepts und der allgemeinen Selbstwirksamkeitserwartung im Verlauf des Übergangs von der 6. zur 7. Klasse untersucht. Dabei wurden einerseits stärkere Effekte des Übergangs auf die schulnahen (i.e. schulisches Wohlbefinden) im Gegensatz zu den schuldistanalen (i.e. allgemeines Wohlbefinden) Variablen erwartet. Mit Blick auf die weiterführende Schulform wurde vor allem für die schulnahen Faktoren eine negativere Entwicklung am Gymnasium erwartet. Der vorliegende Beitrag nutzt dazu die Angaben aus Befragungen von N = 3403 Schülerinnen und Schülern und deren Eltern sowie Daten aus den Schülerakten, die im Rahmen der BERLIN-Studie zur Untersuchung der Auswirkungen der Schulstrukturreform (Maaz, Baumert, Neumann, Becker, & Dumont, 2013) erhoben worden waren.

Deskriptive Mittelwertvergleiche der Schülerschaften beider Schulformen zeigten, dass die zukünftigen Gymnasiastinnen und Gymnasiasten bereits auf der Grundschule ein höheres Wohlbefinden, ein höheres akademisches Selbstkonzept und eine höhere Selbstwirksamkeitserwartung aufwiesen. Allerdings vollzogen sich bei dieser Schülergruppe die stärksten negativen Veränderungen des psychischen Erlebens nach dem Übergang. In einem zweiten Schritt wurden Propensity Score Matching-Analysen vorgenommen, um den Effekt der Schulform auf Veränderungen in den psychosozialen Merkmalen von vorher bestehenden Ausgangsunterschieden zu trennen. Die Vergleiche der gematchten Stichproben ergaben, dass Schülerinnen und Schüler, die auf ein Gymnasium übergegangen sind, auch nach Berücksichtigung von Ausgangsunterschieden stärkere Einbußen in einzelnen Facetten des allgemeinen und des schulischen Wohlbefindens sowie des akademischen Selbstkonzepts erlebten als Schülerinnen und Schüler, die auf eine Integrierte Sekundarschule wechselten. Einzige Ausnahme bildete die Schulzufriedenheit. Hier zeigten die Sekundarschülerinnen und –schüler im Durchschnitt signifikant geringere Ausprägungen nach der Berücksichtigung von Ausgangsunterschieden. Die Ergebnisse bestätigen die Befunde zur Wirkung des Übergangs auf das Wohlbefinden und damit assoziierte Variablen des psychischen Erlebens. Vor dem Hintergrund dieser Befunde wird die Bedeutung einer multidimensionalen und –faktoriellen Betrachtungsweise des Übergangs und der Effekte von Schulformen weitergehend diskutiert.

Chung, H., Elias, M. & Schneider, K. (1998). Patterns of individual adjustment changes during middle school transition. *Journal of School Psychology*, 36, 83-101.

Maaz, K., Baumert, J., Neumann, M., Becker, M. & Dumont, H. (2013) (Hrsg.). *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführende Schulen*. Münster: Waxmann Verlag.

Preckel, F., & Brüll, M. (2010). The benefit of being a big fish in a big pond: Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences*, 20, 522–531.

Proctor, T. & Choi, H.-S. (1994). Effects of transition from elementary school to jUniversitator high school on early adolescents' self-esteem and perceived competence. *Psychology in the Schools*, 31, 319-327.

Valtin, R. & Wagner, C. (2004). Der Übergang in die Sekundarstufe I: Psychische Kosten der externen Leistungsdifferenzierung. *Psychologie in Erziehung und Unterricht*, 51, 52-68.

Van Ophuysen, S. (2008). Zur Veränderung der Schulfreude von Klasse 4 bis 7. Eine Längsschnittanalyse schulformspezifischer Effekte von Ferien und Grundschulübergang. *Zeitschrift für Pädagogische Psychologie*, 22, 293-306.



## ***Effekte verschiedener Formen der Leistungsgruppierung auf das akademische Selbstkonzept: Ein internationaler Vergleich***

In den meisten Bildungssystemen stellen Leistungsgruppierungen, d.h. das Gruppieren von Schülerinnen und Schülern entsprechend ihrer Leistungen und Fähigkeiten in unterschiedliche Schulformen, Schulzweige oder Kurse (im Englischen unter dem Stichwort tracking bekannt), ein zentrales Element schulstruktureller Rahmenbedingungen dar. Dabei ist die Annahme, dass durch das Schaffen von verschiedenen Leistungsgruppen (tracks) Schülerinnen und Schüler entsprechend ihrer Fähigkeiten optimal gefördert werden können. Trotz dieses vielversprechenden Zieles, hat es in den letzten Jahren sowohl in der Bildungspolitik als auch in der Bildungsforschung eine starke Diskussion über die Sinnhaftigkeit von Leistungsgruppierungen insbesondere im Hinblick auf Chancengerechtigkeit gegeben. Während der Besuch eines low tracks in der Tat häufig negative Konsequenzen für den Kompetenzerwerb und die Bildungsbeteiligung hat (Brunello & Checchi, 2007), ist die Befundlage bezüglich der Effekte auf das akademische Selbstkonzept von Schülerinnen und Schülern weniger eindeutig. So argumentieren einige Autoren, dass Schülerinnen und Schüler in low tracks niedrigere Fähigkeitsselbstschätzungen entwickeln (Oakes, 1985). Andere finden hingegen positive Effekte auf das akademische Selbstkonzept, wenn man einer leistungsschwächeren Gruppe angehört (Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007). Die bisherigen Studien zu Effekten der Leistungsgruppierung auf das akademische Selbstkonzept wurden jedoch in einer Vielzahl von Bildungssystemen durchgeführt und es wurde nicht zwischen verschiedenen Formen der Leistungsgruppierung unterschieden.

An dieser Stelle setzt der vorliegende Beitrag an, indem er drei unterschiedliche Formen der Leistungsgruppierung – die Aufteilung von Schülerinnen und Schüler in verschiedene Schulformen (between-school streaming), Schulzweige (within-school streaming) und Kurse (course-by-course tracking) – im Hinblick auf Effekte auf das akademische Selbstkonzept miteinander vergleicht. Dabei wurde angenommen, dass sich die Referenzgruppen von Schülerinnen und Schülern je nach Form der Leistungsgruppierung voneinander unterscheiden, was zu unterschiedlichen Fähigkeitsselbstkonzepten führen sollte. Untersucht wurde diese Fragestellung mittels Mehrebenenmodellen anhand von Daten aus PISA 2003 von ca. 99000 Schülerinnen und Schülern aus 3400 Schulen in 20 Ländern für die Domäne Mathematik. Zunächst wurde der Zusammenhang zwischen track level und mathematischem Selbstkonzept für die verschiedenen Formen der Leistungsgruppierung untersucht (Modell 1). Anschließend wurden sukzessive weitere Variablen in das Modell aufgenommen: die individuelle Leistung der Schülerinnen und Schüler (Modell 2), die mittlere Leistung des tracks (Modell 3) sowie Noten, Geschlecht, SES und Ethnizität als Kontrollvariablen (Modell 4). Die Modelle wurden als 3-Ebenen-Modelle über alle Länder und als 2-Ebenen-Modelle für jedes Land getrennt berechnet. Sowohl im Rahmen der 3-Ebenen-Modelle als auch der 2-Ebenen-Modelle zeigte sich, dass Schülerinnen und Schüler in high tracks in Systemen mit course-by-course tracking ein höheres mathematisches Selbstkonzept aufwiesen als Schülerinnen und Schüler gleichen Leistungsniveaus in low tracks. Das umgekehrte Bild wurde in Systemen mit within-school und between-school streaming festgestellt: Hier hatten die Schülerinnen und Schüler in low tracks höhere Fähigkeitsselbstkonzepte als Schülerinnen und Schüler in high tracks. Dieses Befundmuster blieb selbst nach Kontrolle des mittleren Leistungsniveaus des tracks als auch weiterer Kontrollvariablen bestehen und zeigte sich gleichermaßen in den 3-Ebenen-Modellen und 2-Ebenen-Modellen. Die Ergebnisse bestätigen demnach die Annahme, dass die Effekte von Leistungsgruppierungen auf das akademische Selbstkonzept je nach Form der Leistungsgruppierung unterschiedlich ausfallen. Das Befundmuster spricht dafür, dass Schülerinnen und Schüler in Systemen mit course-by-course tracking andere Referenzgruppen zur Einschätzung ihrer eigenen Fähigkeiten heranziehen als Schülerinnen und Schüler in Systemen mit within-school und between school-streaming: Schülerinnen und Schüler, die nur für einen bestimmten Kurs nach ihren Leistungen gruppiert werden, vergleichen sich mit allen Schülerinnen und Schülern und sind sich über die Zugehörigkeit zu ihrer Leistungsgruppe bewusst; Schülerinnen und Schüler, die verschiedene Zweige innerhalb einer Schule oder gar verschiedene Schulformen besuchen, vergleichen sich hingegen vorrangig mit Schülerinnen und Schülern des gleichen tracks und die Zugehörigkeit zu ihrer Leistungsgruppe verliert an Bedeutung.

Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy*, 22, 781–861.

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, 44, 631–669.

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

Chair(s): Marcus Hasselhorn (DIPF, IDeA), Ulrike Hartmann (DIPF, IDeA), Andreas Gold (Goethe-Universität Frankfurt, IDeA)

Diskutant/in: Claudia Mähler (Universität Hildesheim)

IDeA (Individual Development and Adaptive Education of Children at Risk) ist ein gemeinsames Forschungszentrum des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF), der Goethe-Universität Frankfurt (GU) und des Frankfurter Sigmund-Freud-Instituts (SFI). Seit 2008 verbindet das Zentrum vielfältige Forschungsperspektiven bei der empirischen Untersuchung von individuellen Entwicklungsprozessen und von Ansätzen zur Lernförderung bei Kindern in den ersten zwölf Lebensjahren. Dabei wird davon ausgegangen, dass Risikofaktoren für erfolgreiches Lernen sowohl individuelle als auch kontextuelle Charakteristika umfassen und dass solche Risiken Veränderungen über die Zeit unterliegen. Die Forschungsprojekte am IDeA-Zentrum beschäftigen sich mit unterschiedlichen Facetten kindlicher Entwicklung, vor allem wenn sie aufgrund vorhandener Risikofaktoren nicht optimal verläuft. Darauf aufbauend werden Einflüsse adaptiver Lernumgebungen und der Einsatz von Lehrmethoden untersucht. Die Beiträge beleuchten aus verschiedenen Blickwinkeln Risikofaktoren für kindliches Lernen und die Möglichkeiten, ihnen entgegenzuwirken.

Der erste Beitrag (Fischmann et al.) betrachtet die sozio-emotionalen Kompetenzen von Kindern im Kindergartenalter aus psychoanalytischer Perspektive. Verglichen wird die Wirksamkeit zweier Präventionsangebote, die in Kindertagesstätten mit erhöhter sozialer Problemlage zum Einsatz gekommen sind. FAUSTLOS ist ein bewährtes gewaltpräventiv konzipiertes Programm. Das aufwendigere Programm FRÜHE SCHRITTE ist von den Autorinnen auf der Grundlage der psychoanalytischen Theorie entwickelt worden. In sorgfältig randomisierten Stichproben einer großen Grundgesamtheit wird eine vergleichende Wirksamkeitsprüfung der beiden Vorgehensweisen vorgenommen.

Der zweite Beitrag (Müller et al.) thematisiert die Sprachförderung von Kindern im Elementar- und Primarbereich. Dass die professionellen Kompetenzen pädagogischer Fachkräfte eine entscheidende Gelingensbedingung erfolgreicher Sprachförderung darstellen, wird zwar häufig behauptet, ist aber selten empirisch untersucht worden. Die Autorinnen untersuchen das Niveau des linguistischen Grundlagenwissens und das praxisrelevante Wissen der pädagogischen Fachkräfte in Krippen, Kindergärten und Grundschulen. Sie stellen fest, dass es entsprechende Professionalisierungsbedarfe gibt.

Der dritte Beitrag (Becker, Tuppat & Lange) beschäftigt sich mit dem Übergang vom Kindergarten in die Grundschule und untersucht aus soziologischer Perspektive die institutionellen Faktoren, die bei der Einschulung bzw. Rückstellung von Kindern mit Migrationshintergrund eine Rolle spielen. Diskutiert wird, ob die Einrichtung neuer Formen der Schuleingangsstufe ethnisch bedingte Bildungsungleichheiten vermindern helfen kann. Die Autorinnen greifen damit eine besonders aktuelle Thematik auf, weil die mit dem Migrationsstatus einhergehenden Disparitäten in Bezug auf Einschulung/Rückstellung besonders evident sind.

Im vierten Beitrag (Decristan et al.) wird die Perspektive der Schuleffektivitätsforschung mit einer Forschungstradition verknüpft, die sich mit der Identifikation von Merkmalen guten Unterrichts beschäftigt. Grundlage ist eine quasi-experimentelle Interventionsstudie zur Wirksamkeit unterschiedlicher Instruktionmethoden im naturwissenschaftlichen Grundschulunterricht. Die Autorengruppe berichtet von Effekten der kognitiven Aktivierung und eines unterstützenden Lernklimas auf die Wirksamkeit des formativen Assessments.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
Casino-Gebäude  
Raum 823





Tamara Fischmann (Sigmund-Freud-Institut Frankfurt, IDeA), Verena Neubert (Sigmund-Freud-Institut, IDeA), Lorena Hartmann (Sigmund-Freud-Institut, IDeA), Peter Ackermann (Sigmund-Freud-Institut, IDeA), Maria Schreiber (Sigmund-Freud-Institut, IDeA), Marianne Leuzinger-Bohleber (Sigmund-Freud-Institut, IDeA)

## ***Evaluation zweier Frühpräventionsprojekte in Kindertagesstätten mit erhöhter sozialer Problemlage***

Das Projekt EVA befasst sich mit der Evaluation zweier Frühpräventionsprogramme in Kindertagesstätten. Wie viele Studien aus dem Bereich der empirischen Säuglingsforschung, der Bindungsforschung, der Psychoanalyse und der Frühpädagogik belegen, sind tragende emotionale Beziehungserfahrungen in den ersten Lebensjahren die beste Voraussetzung für eine gelingende psychische, kognitive und psychosoziale Entwicklung einschließlich des Spracherwerbs. Zwar greifen vielfältige Präventionsprogramme diese empirische Befundlage auf und bemühen sich um eine frühe Förderung von Kindern schon im Kindergartenalter. Doch nur selten wird die Wirksamkeit der Programme wissenschaftlich untersucht.

Vor diesem Hintergrund überprüft das Projekt EVA die differentielle Wirksamkeit zweier bewährter Präventionsprogramme – FAUSTLOS und FRÜHE SCHRITTE. Ausgehend von einer Basiserhebung mit 5300 Kindern und einer darauf basierenden Clusteranalyse wurden die beiden Präventionsprojekte randomisiert in je sieben ausgewählten Kindertagesstätten durchgeführt (insgesamt wurden 240 Kinder in die Studie eingeschlossen). Ziel des Projekts ist es zu untersuchen, ob sich der Mehraufwand des psychoanalytisch orientierten Präventionsangebot 1 "FRÜHE SCHRITTE", im Vergleich mit dem standardisierten Präventionsangebot 2 "FAUSTLOS" kurz- und langfristig lohnt bzw. welches die Vor- und Nachteile beider Präventionsangebote sind. „FRÜHE SCHRITTE“ besteht aus verschiedenen psychoanalytisch begründeten Modulen: Einer Supervision des Erzieherinnenteams, der wöchentlichen Anwesenheit von erfahrenen Kindertherapeuten in den Einrichtungen für Beratungen und Therapien vor Ort, einer intensiven Elternarbeit, einem Einsatz des Programms FAUSTLOS und der Begleitung des Übergangs vom Kindergarten in die Grundschule. FAUSTLOS ist ein standardisiertes Gewaltpräventionsprogramm, das von den ErzieherInnen wöchentlich angeboten wird.

Eine erste dreijährige Projektphase wurde 2011 abgeschlossen. Eine Replikation der EVA-Studie wird z.Zt. durchgeführt.

Im Kongressbeitrag werden erste Ergebnisse präsentiert. Ein auffallendes Ergebnis war, dass zu Beginn der Studie über ein Drittel der Kinder die problematischen Bindungstypen C (unsicher ambivalent gebunden) und D (desorganisiert gebunden) aufwiesen, die bekanntlich eine ungünstige Prognose aufweisen (gemessen mit dem Manchester Child Attachment Story Task). Erste Ergebnisse zeigen, dass es gelungen ist, einem relativ großen Anteil dieser Kinder zu einem weniger problematischen Bindungstyp zu verhelfen und dadurch ihre weitere psychosoziale und kognitive Entwicklung nachhaltig positiv zu beeinflussen.

Weiter konnte gezeigt werden, FRÜHE SCHRITTE die selbstreflexiven Kompetenzen der Erzieherinnen stärkte und dass die Eltern bezüglich ihrer elterlichen Kompetenzen unterstützt werden konnten. Kurz wird auf das sogenannte KIGRU Projekt (Begleitung gefährdeter Kinder beim Übergang Grundschule-Kindergarten) eingegangen.

Leuzinger-Bohleber, M.; Läzer, K.L.; Pfenning-Meerkötter, N.; Fischmann, T.; Wolff, A.; Green, J. (2011): Psychoanalytic treatment of ADHD children in the frame of two extraclinical studies: The Frankfurt Prevention Study and the EVA Study. *Journal of Infant, Child, and Adolescent Psychotherapy* 10): 32-50

Pfenning-Meerkötter, N.; Läzer, K.L.; Schiller, B.; Hartmann, L.K.; Leuzinger-Bohleber, M. (2012): No intermediate space for children at risk. In: Fonagy, P.; Kächele, H.; Leuzinger-Bohleber, M.; Taylor, D. (eds.): *The significance of dreams: Bridging clinical and extraclinical research in psychoanalysis*. London: Karnac Books, 245-268

Läzer, K. L., Leuzinger-Bohleber, M., Rüger, B., Fischmann, T. (2013): Evaluation of two prevention programs 'Early Steps' and 'Faustlos' in daycare centers with children at risk: the study protocol of a cluster randomized controlled trial. In: *Trials* 2013, 14:268 doi:10.1186/1745-6215-14-268

<http://www.trialsjournal.com/content/14/1/268/abstract>



Anja Müller (Goethe Universität Frankfurt, IDEa), Barbara Geist (Goethe-Universität Frankfurt),  
Sabrina Geyer (Goethe-Universität Frankfurt, IDEa), Katinka Smits (Goethe-Universität Frankfurt, IDEa),  
Petra Schulz (Goethe-Universität Frankfurt, IDEa)

## ***Sprachförderkompetenzen von pädagogischen Fachkräften im Elementar- und Primarbereich***

Diese Studie untersucht das Fachwissen über Sprachförderung bei verschiedenen Gruppen pädagogischer Fachkräfte und leitet aus den Ergebnissen ab, dass die Professionalisierungsbedürfnisse fachgruppenspezifisch sind.

Die Sprachförderkompetenz der pädagogischen Fachkraft (PF) ist mitentscheidend für das Gelingen von Sprachförderung. In der Sprachförderung soll sich die PF an die sprachlichen Fähigkeiten des individuellen Kindes anpassen und sich dabei am natürlichen Spracherwerb orientieren (vgl. Hopp et al. 2010). Dafür benötigen PF u.a. spezifisches Fachwissen über linguistische und praxisrelevante Grundlagen von Sprachförderung (vgl. Hopp et al. 2010). Bislang liegen nur wenige Studien zum Fachwissen von PF vor. Fried (2007) fand in einer Fragebogenstudie, dass 75% der befragten PF Diagnostik als genuinen Bestandteil der Sprachförderung ansehen, jedoch nur 50% sich die Durchführung einer Sprachdiagnostik zutrauen. Im Rahmen der Pilotierung eines Fragebogens zur Erhebung der Sprachförderkompetenz von PF untersuchten Thoma et al. (2011) das Fachwissen von PF aus Kitas und von Studierenden eines frühpädagogischen Bachelorstudiengangs. Die Ergebnisse zeigten, dass das Wissen zwischen beiden Gruppen und innerhalb der Gruppen stark variiert. Rothweiler et al. (2009) konnten nachweisen, dass das Fachwissen von PF durch Weiterbildungsmaßnahmen nachhaltig erweitert werden kann.

Bislang fehlen Untersuchungen, die das Fachwissen von PF in verschiedenen Bereichen der Sprachförderung wie linguistisches Grundlagenwissen und praxisrelevantes Wissen erfassen und dabei die unterschiedlichen Institutionen berücksichtigen, in denen Sprachförderung durchgeführt wird. Da in der Sprachförderung verschiedenen Fachkräfte-Gruppen wie z.B. Kita und Grundschule agieren, sind Unterschiede zwischen den verschiedenen Fachkräfte-Gruppen zu erwarten, die auf verschiedene Professionalisierungsbedürfnisse hinweisen. Ziel dieser Studie ist es daher, das linguistische und praxisrelevante Fachwissen über Sprachförderung von PF aus verschiedenen Fachgruppen zu untersuchen und zu vergleichen. Folgende Fragen wurden untersucht:

Über welches linguistische und praxisrelevante Grundlagenwissen verfügen PF?

Unterscheidet sich das Fachwissen zwischen den Fachgruppen (Kita, U3, Grundschule)?

56 PF (54 Frauen; 2 Männer) nahmen an der Untersuchung teil; 31 Kita-Fachkräfte (Durchschnittsalter=40;1), 13 U3-Fachkräfte (Durchschnittsalter=38;6), 10 Grundschullehrkräfte (Durchschnittsalter=42;5). Zur Erhebung des Fachwissens wurde die paper-and-pencil Version des SprachKOPF-Wissenstests (Thoma/Tracy 2013) verwendet, der das Fachwissen von PF aus dem Elementar- und Primarbereich standardisiert erfasst. Der Test beinhaltet 55 Items (54 Mehrfachwahlaufgaben und 1 Zuordnungsaufgabe) in zwei Wissensbereichen: Basiswissen über die verschiedenen linguistischen Ebenen (N=22) und praxisrelevantes Wissen (N=33), das Fragen zum Spracherwerb (inklusive Mehrsprachigkeit) und Sprachdiagnostik/-förderung umfasst. Aufgrund der in der Testkonstruktion berücksichtigten Ratekorrektur erhält ein Proband bei einer vollständig zufälligen Beantwortung aller Fragen ein Gesamtergebnis von 0. Die 56 Probanden erreichten einen Durchschnittswert von .38 (SD=.23; Bereich: .00-.92). Die Ergebnisse liegen damit in dem von Thoma/Tracy (2013) angegebenen Bereich der Vergleichsstichprobe (Mittelwert=.36; SD=.22; Bereich: -.09-.95). Die Kita-Fachkräfte erzielten einen Durchschnittswert von .28 (SD=.19; Bereich: .00-.62), die U3-Fachkräfte von .42 (SD=.23; Bereich: .02-.74) und die Lehrkräfte von .61 (SD=.27; Bereich: .25-.92). Die Unterschiede zwischen den Gruppen sind signifikant ( $F(2,53)=11.41$ ;  $p < .001$ ). Paarvergleiche zeigen signifikante Unterschiede zwischen Kita-Fachkräften und U3-Fachkräften ( $t(df=44)=2.38$ ;  $p < .05$ ), zwischen Kita-Fachkräften und Lehrkräften ( $t(df=39)=4.74$ ;  $p < .001$ ) sowie zwischen Lehrkräften und U3-Fachkräften ( $t(df=23)=2.19$ ;  $p < .05$ ). Die Analyse der Items hinsichtlich linguistischem (LW) und praxisrelevantem Wissen (PW) ergab für die Kita-Fachkräfte einen signifikanten Unterschied (LW: Mittelwert=.26; SD=.20; PW: Mittelwert=.33; SD=.21;  $t(df=30)=2.38$ ;  $p < .05$ ). Die Werte der U3-Fachkräfte und der Lehrkräfte unterscheiden sich in diesen zwei Testskalen jeweils nicht signifikant (U3-Fachkräfte: LW: Mittelwert=.43; SD=.21; PW: Mittelwert=.40; SD=.24;  $t(df=14)=.63$ ;  $p=.541$ ; Lehrkräfte: LW: Mittelwert=.63; SD=.24; PW: Mittelwert=.58; SD=.20;  $t(df=9)=1.13$ ;  $p=.287$ ). Zusammenfassend zeigen die Ergebnisse Unterschiede im linguistischen und praxisrelevanten Fachwissen zwischen den verschiedenen Fachkräfte-Gruppen: Lehrkräfte haben ein umfangreicheres Fachwissen über Sprachförderung als U3- und Kita-Fachkräfte; U3-Fachkräfte haben wiederum ein umfangreicheres Fachwissen als Kita-Fachkräfte. Aus diesen Ergebnissen leitet sich ein unterschiedlicher Professionalisierungsbedarf für die Fachkräfte-Gruppen ab. Erklärungsansätze für die Ergebnisse der einzelnen Fachkräfte-Gruppen sowie gruppenspezifischen Professionalisierungsbedürfnisse werden im Vortrag diskutiert.

Fried, L. (2007): Sprachförderkompetenz von ErzieherInnen. In: Sozial Extra 15, H. 13, S. 26–28.

Hopp, H./Thoma, D./Tracy, R. (2010): Sprachförderkompetenz pädagogischer Fachkräfte. Ein sprachwissenschaftliches Modell. Zeitschrift für Erziehungswissenschaft 13, 4, S. 609-629.

Rothweiler, M./Ruberg, T./Utecht, D. (2009): Praktische Kompetenz ohne theoretisches Wissen? Zur Rolle von Sprachwissenschaft und Spracherwerbtheorie in der Ausbildung von ErzieherInnen und GrundschullehrerInnen. In: Wenzel, D./Koepfel, G./Carle, U. (Hrsg.): Kooperation im Elementarbereich. Eine gemeinsame Ausbildung für Kindergarten und Grundschule. – Baltmannsweiler, S. 111-122.

Thoma, D./Tracy, R. (2013): SprachKOPFv06.1. Instrument zur standardisierten Erhebung der Sprachförderkompetenz pädagogischer Fachkräfte, Testbogen Teilkompetenz Wissen. – Mannheim: Universität Mannheim.

Thoma, D./Ofner, D./Seybel, C./Tracy, R. (2011): Professionalisierung in der Frühpädagogik: Eine Pilotstudie zur Sprachförderkompetenz. Frühe Bildung, 0, S. 31-36.



Birgit Becker (Goethe-Universität Frankfurt, IDeA), Julia Tuppatt (Goethe-Universität Frankfurt, IDeA), Matthias Lange (Goethe-Universität Frankfurt)

## ***Institutionelle Rahmenbedingungen an der Schnittstelle Einschulung – vermindert die Eingangsstufe ethnische Bildungsungleichheit?***

Dieser Beitrag untersucht Risiken für Kinder mit Migrationshintergrund beim Übergang in die Grundschule aus einer soziologischen Perspektive. Deskriptiv zeigt sich, dass Kinder mit Migrationshintergrund überproportional häufig von der Einschulung zurückgestellt werden (Autorengruppe Bildungsberichterstattung 2010). Der Schwerpunkt dieses Beitrags liegt auf einer Betrachtung institutioneller Faktoren für dieses Rückstellungsrisiko.

Die Einschulungsentscheidung obliegt in den meisten deutschen Bundesländern der Schulleitung der aufnehmenden Schule auf Grundlage der Schuleingangsuntersuchung, in der (vor der Einschulung) die Schulfähigkeit eines Kindes überprüft wird. Kinder, die für (noch) nicht schulfähig befunden werden, können verpflichtend für ein Jahr von der Einschulung zurückgestellt werden.

Als theoretischer Hintergrund dieses Beitrags dient die Theorie der institutionellen Diskriminierung (Gomolla & Radtke 2009). Diese verortet die Produktion und Reproduktion ethnischer Bildungsungleichheiten in der Institution Schule. Eine Kernidee ist hierbei, dass institutionelle Rahmenbedingungen dazu führen können, dass bestimmte Gruppen von Kindern (wie z.B. Kinder mit Migrationshintergrund) systematisch benachteiligt werden.

Auf Basis dieser theoretischen Überlegungen gehen wir in diesem Beitrag der Frage nach, ob das Risiko einer Rückstellung für Kinder mit Migrationshintergrund systematisch nach verschiedenen institutionellen Rahmenbedingungen bei der Einschulung variiert. Konkret untersuchen wir, ob die Eingangsstufe nach dem hessischen Modell das erhöhte Rückstellungsrisiko für Kinder mit Migrationshintergrund gegenüber dem regulären Verfahren der Einschulung in die Regelgrundschule vermindert.

Im Modell der Eingangsstufe, in dem das erste Schuljahr auf zwei Jahre ausgedehnt wird, sollen alle Fünfjährigen ungeachtet ihrer Schulfähigkeit eingeschult werden, während bei der Einschulung in die Regelgrundschule prinzipiell alle Kinder zurückgestellt werden können, die „nicht den für den Schulbesuch erforderlichen körperlichen, geistigen und seelischen Entwicklungsstand haben“ (Hessisches Schulgesetz, §58). Es ist davon auszugehen, dass im Eingangsstufenmodell Rückstellungen nur noch im Fall erheblicher gesundheitlicher Probleme oder kognitiver Entwicklungsdefizite erfolgen. Nach der Theorie der institutionellen Diskriminierung ist der „Spielraum“ der Entscheidung hierbei deutlich geringer als im Regelverfahren. Dies sollte die Überrepräsentation von Kindern mit Migrationshintergrund bei Rückstellungen vom Schulbesuch verringern.

Um diese Hypothese zu überprüfen, verwenden wir Daten der Schuleingangsuntersuchungen der Jahre 2008 bis 2012 aus sieben hessischen Landkreisen bzw. kreisfreien Städten (N=41.321). Für jedes Kind im Datensatz ist angegeben, ob der Schulbezirk am Wohnort eine Regelgrundschule oder eine Schule mit Eingangsstufe vorhält. Die abhängige Variable der Analysen stellt die Empfehlung der Schulärztin bzw. des Schularztes dar, ob das Kind regulär eingeschult werden sollte oder ob eine Rückstellung empfohlen wird. Der Entwicklungsstand des Kindes in verschiedenen Bereichen sowie einige weitere familiäre Merkmale sind in diesen Daten ebenfalls enthalten und können in den multivariaten Analysen statistisch kontrolliert werden.

Als Auswertungsmethode werden logistische Regressionen angewendet, um das Rückstellungsrisiko für Kinder mit Migrationshintergrund im Vergleich zu Kindern ohne Migrationshintergrund in Abhängigkeit von der Existenz einer Eingangsstufe zu untersuchen. Konkret wird untersucht, ob es einen Interaktionseffekt zwischen dem Migrationshintergrund und der institutionellen Rahmenbedingung Eingangsstufe auf das Rückstellungsrisiko gibt.

Unsere Befunde deuten darauf hin, dass die Eingangsstufe die Nachteile kompensieren kann, die für Kinder mit Migrationshintergrund bei der Einschulung in die Regelgrundschule bestehen: Während die Rückstellungsquote von Kindern mit Migrationshintergrund in der Regelschule mit 12,5% etwa doppelt so hoch ausfällt wie die Rückstellungsquote von Kindern ohne Migrationshintergrund (6,8%), verschwindet dieser Unterschied in der Eingangsstufe fast komplett (5,7% zu 5,4%). Auch besteht ein auf dem 99%-Niveau signifikanter negativer Interaktionseffekt zwischen der Existenz einer Eingangsstufe und einem Migrationshintergrund auf die Rückstellungsempfehlung. Der Nachteil von Migrantenkindern variiert also stark je nach institutionellem Setting. Die Eingangsstufe könnte einen wirksamen Lösungsansatz darstellen, ethnische Bildungsungleichheit bei der Einschulung zu verringern.

Autorengruppe Bildungsberichterstattung (2010): Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demografischen Wandel. Berlin: Bundesministerium für Bildung und Forschung.

Gomolla, M. & Radtke, O. (2009): Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule. 3. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.

**GEBF** TAGUNG  
2014

E01 c

Symposiums  
beitrag

Mittwoch,  
05.03.2014,

09:00 - 10:45,

Casino-Gebäude

Raum 823

2

Entwicklungsrisiken und individuelle Förderung im Elementar- und Primarbereich

Jasmin Decristan (DIPF), Eckhard Klieme (DIPF), Mareike Kunter (Goethe-Universität Frankfurt), Gerhard Büttner (Goethe-Universität Frankfurt), Benjamin Fauth (Goethe-Universität Frankfurt), Lena Hondrich (DIPF), Silke Hertel (Universität Heidelberg), Svenja Rieser (Goethe-Universität Frankfurt), Ilonca Hardy (Goethe-Universität Frankfurt)

## **Interaktionen zwischen Formativem Assessment und Unterrichtsqualität in der Grundschule**

### Hintergrund

Eine der gegenwärtig prominentesten und effektivsten Lehr-Lern-Methoden ist „Formatives Assessment“ (Black & Wiliam, 2009), bei der der jeweilige Wissensstand der Schülerinnen und Schüler lernbegleitend diagnostiziert und rückgemeldet wird. Die Qualität der Umsetzung (z.B. des Feedbacks) wird als entscheidend für die Wirksamkeit einer Methode und als positiver Moderator für dessen Effektivität angesehen (z.B. Landenberger & Lipsey, 2005). Parallel zur Effektivitätsforschung und weitgehend unabhängig davon wurden keine methodenspezifischen, sondern allgemeine Qualitätsmerkmale von Unterricht identifiziert, die mit schulischem Lernen in Beziehung stehen: kognitive Aktivierung, unterstützendes Klima und strukturierte Klassenführung (vgl. Klieme, Pauli & Reusser, 2009). In einem kognitiv aktivierenden Unterricht knüpfen Lehrkräfte an Vorwissen an, vergleichen Ideen und Konzepte miteinander und stellen zum Denken herausfordernde Aufgabenstellungen. Ein unterstützendes Klima bezieht sich auf positive Lehrer-Schüler-Interaktionen und konstruktiv-wohlwollende Lernerunterstützung. Für eine strukturierte Klassenführung mit einem störungsarmen Unterricht setzen Lehrkräfte (präventive) Strategien und Regeln ein.

In diesem Beitrag wird die Perspektive der Effektivitätsforschung mit der Perspektive allgemeiner Unterrichtsqualitätsmerkmale verbunden und deren Wechselspiel in der Förderung des naturwissenschaftlichen Verständnisses von Grundschulkindern geprüft.

### Hypothesen

(1) Die drei Basisdimensionen von Unterrichtsqualität (kognitive Aktivierung, unterstützendes Klima, strukturierte Klassenführung) fördern das konzeptuelle Verständnis von Grundschulkindern.

(2) Die drei Unterrichtsqualitätsdimensionen moderieren die Effektivität von Formativem Assessment auf das konzeptuelle Verständnis von Grundschulkindern.

### Methode

Ausgangspunkt ist das Projekt IGEL (Individuelle Förderung und adaptive Lern-Gelegenheiten in der Grundschule), eine quasi-experimentelle Interventionsstudie zur Wirksamkeit von drei verschiedenen Fördermethoden im naturwissenschaftlichen Grundschulunterricht. In vorangegangenen Analysen hat sich insbesondere das Formative Assessment im Vergleich zu einer Treatment-Kontrollgruppe als lernförderlich erwiesen (vgl. Decristan et al., 2013). Die Analysen dieses Beitrags knüpfen hieran an und beziehen sich auf die Formative Assessment- und die Kontrollgruppe der Studie mit 28 Klassen und 551 Kindern. Nach der Intervention wurde das konzeptuelle Verständnis (13 Items; EAP/PV Reliabilität = .70) erfasst. Schülerangaben zu den drei Basisdimensionen von Unterrichtsqualität wurden klassenweise zusammengefasst (Cronbachs alpha > .79, ICC2 > .73; vgl. Fauth et al., 2014). Die Hypothesen wurden in Mehrebenen-Regressionsanalysen mithilfe von Mplus 7 geprüft, wobei jeweils für kognitive Grundfähigkeiten, Vorwissen zum Thema Schwimmen und Sinken, naturwissenschaftliche und Sprachkompetenz der Schülerinnen und Schüler vor Beginn der Intervention kontrolliert wurde.

### Ergebnisse

Aufbauend auf bisherigen Ergebnissen des Projekts, dass Formatives Assessment das konzeptuelle Verständnis von Grundschulkindern fördert (vgl. Decristan et al., 2013), zeigten sich erwartungsgemäß auch positive Effekte von kognitiver Aktivierung und einem unterstützendem Klima auf Schülerleistungen (Hypothese 1). Beide Unterrichtsqualitätsdimensionen erwiesen sich jeweils als positive Moderatoren für die Effektivität des Formativen Assessments (Hypothese 2). Wider Erwarten ließen sich diese positiven Wirkungen für eine strukturierte Klassenführung nicht bestätigen.

### Diskussion

Die Ergebnisse untermauern gegenwärtige theoretische Modelle, in denen die Qualität der Lehrer-Schüler-Interaktionen für das Lernen betont wird. Zugleich stützen sie empirische Ergebnisse zur Relevanz von Merkmalen der Unterrichtsqualität für die Effektivität von Fördermethoden. Für die schulische Aus- und Weiterbildung sowie die Schulpraxis zeigen die Ergebnisse, dass es wichtig ist, bei der Umsetzung von Fördermethoden allgemeine Unterrichtsqualitätsmerkmale zu berücksichtigen. Das jeweils soziespezifische Zusammenspiel zwischen den drei allgemeinen Unterrichtsqualitätsdimensionen und weiteren Fördermethoden bleibt in zukünftigen Studien empirisch zu prüfen.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. doi:10.1007/s11092-008-9068-5

Decristan, J., Hondrich, A. L., Büttner, G., Hertel, S., Klieme, E., Kunter, M., Lühken, A., Adl-Amini, K., Djakovic, S.-K., Mannel, S., Naumann, A. & Hardy, I. (eingereicht). Impact of Additional Guidance in Science Education on Primary Students' Conceptual Understanding.

Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school: dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom* (pp. 137–160). Münster: Waxmann.

Landenberger, N. A. & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1, 451–476. doi:10.1007/s11292-005-3541-7



Ilka Wolter (University of Sussex), Bettina Hannover (Freie Universität Berlin)

## ***Unterschiede in der schulischen Entwicklung von Mädchen und Jungen: Hat die Lehrkraft einen Einfluss?***

**GEBF** TAGUNG  
2014

**E02**

**Symposium**

Chair(s): Ilka Wolter (University of Sussex), Bettina Hannover (Freie Universität Berlin)

Diskutant/in: Oliver Dickhäuser (Universität Mannheim)

Insbesondere vor dem Hintergrund, dass Frauen in vorschulischer und schulischer Bildung überrepräsentiert sind, hat in den letzten Jahren die Frage viel Aufmerksamkeit erfahren, ob das Geschlecht der Erziehungs- oder Lehrperson für die differentielle akademische Entwicklung von Mädchen und Jungen (mit)verantwortlich ist. Das geplante Symposium beleuchtet diese Fragestellung aus soziologischer (Helbig), entwicklungs- und sozialpsychologischer (Wolter, Hannover & Oakhill), erziehungswissenschaftlicher (Denn, Kastens, Theurer, Lotz & Lipowsky) und fachdidaktischer (Jurik, Häusler, Stubben & Seidel) Perspektive. Wie Helbig zeigt, ergeben Makroanalysen von Large Scale-Datensätzen (IGLU 2001, Mikrozensus 2008 und 2009) zu Kompetenzen und formalen Indikatoren von Bildungserfolg (Noten, Übergang auf das Gymnasium) keine Hinweise auf einen differentiellen Einfluss des Lehrkraftgeschlechts oder einer Passung zwischen Schüler/in- und Lehrer/in-Geschlecht auf Mädchen und Jungen. Entsprechend untersuchen die weiteren Beitragenden in unserem Symposium möglicherweise relevante Prozesse der Interaktion zwischen Erziehungspersonen/ Lehrkräften und Lernenden auf der Meso- und Mikroebene. Wolter et al. zeigen, dass Erzieherinnen im Kindergarten häufiger weiblich als männlich konnotierte Lernangebote machen und zu Mädchen engere Bindungen eingehen als zu Jungen, wobei die Interaktion dieser Variablen die Rechtschreibkompetenzentwicklung von Mädchen und Jungen differentiell beeinflusst. Denn et al. untersuchen in der zweiten Grundschulklasse mögliche Unterschiede im Zusammenhänge zwischen Meldeverhalten der Kinder und dem Aufrufverhalten der Lehrkraft in Abhängigkeit des Geschlechts, der Leistung und des Selbstkonzepts des Kindes. Während die Lehrkräfte Mädchen und Jungen nicht unterschiedlich häufig aufrufen, haben Jungen dennoch mehr Redebeiträge, bedingt durch häufigeres unaufgefordertes Zwischenrufen. Weiter wirkt sich das Selbstkonzept des Kindes in Abhängigkeit seines Geschlechts unterschiedlich auf das Aufrufverhalten der Lehrperson aus. Jurik et al. identifizieren mit latenten Klassenanalysen bei Gymnasiasten und Gymnasiastinnen der 8. Klassenstufe unterschiedliche Profile des Zusammenspiels kognitiver und motivational-affektiver Lernvoraussetzungen und zeigen, dass die Häufigkeit der Profile interaktiv von Schüler/in-Geschlecht und Geschlechtskonnotation des Schulfaches (Deutsch vs. Mathematik) bestimmt ist. Videoanalysen des Unterrichts können weiterführend auf Interaktionen zwischen Lehrkraftgeschlecht und Umgang mit verschiedenen Profilen bei Lernenden verweisen. Zusammengefasst sprechen die im geplanten Symposium dargestellten Befunde gegen die Annahme, dass es das biologische Geschlecht der Erziehungs- oder Lehrperson als solche ist, das die akademische Entwicklung von Mädchen und Jungen differentiell beeinflusst. Vielmehr scheinen es (durch Geschlechtsstereotype bedingt) geschlechtstypisiert unterschiedlich ausgeprägte Merkmale der Beziehungs- und Unterrichtsgestaltung der Lehrkraft zu sein, die interaktiv mit (ebenfalls durch Geschlechtsstereotype bedingt) geschlechtstypisiert ausgeprägten Selbstkonzepten und Präferenzen der Lernenden die Kompetenzentwicklung von Mädchen und Jungen differentiell beeinflussen.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
IG-Farben-Haus  
Raum 311

**Unterschiede in der schulischen Entwicklung von Mädchen und Jungen**

**2**



## ***Die empirische Tragfähigkeit der Same-Sex Hypothese in der Bildungsforschung***

Spätestens seit Freud wird angenommen, dass sich Kinder in ihrer geschlechtlichen Entwicklung vor allem an gleichgeschlechtlichen Sozialisationsagenten orientieren. Auch wenn Freuds Theorie mittlerweile verworfen bzw. weiterentwickelt wurde, blieb die zentrale Hypothese bestehen: Jungen und Mädchen brauchen zur Entwicklung ihrer geschlechtlichen Identität gleichgeschlechtliche Vorbilder. Diese Annahme wurde im Zuge der Krise der Jungen im Bildungssystem auf den Bildungserfolg von Jungen und Mädchen übertragen. Vor allem das Fehlen männlicher Vorbilder in der Lehrerschaft (Feminisierung der Schule) und im Haushalt (Anstieg der Zahl weiblicher Alleinerzieherhaushalte) wurde als ursächlich für das Scheitern der Jungen im Bildungssystem gesehen.

In dem Beitrag will ich darlegen, warum gleichgeschlechtliche „Sozialisationsagenten“ überhaupt einen positiven Einfluss auf den Bildungserfolg von Jungen und Mädchen haben sollten. Dabei konzentriere ich mich auf Lehrkräfte, Eltern und Geschwister. Im Anschluss daran zeige ich anhand von drei Studien, dass weder gleichgeschlechtliche Lehrkräfte, noch das Vorhandensein des gleichgeschlechtlichen Elternteils oder gleichgeschlechtlicher älterer Geschwister im Elternhaus den Bildungserfolg von Jungen und Mädchen beeinflussen.

Für die Prüfung des Einflusses des Lehrkraftgeschlechts wurden die Daten der IGLU-Studie 2001 genutzt. Hierbei wurde der Einfluss des Lehrkraftgeschlechts auf Kompetenzen und Noten von Jungen und Mädchen untersucht. Für die Prüfung des Einflusses des gleichgeschlechtlichen Elternteils und einer möglichen Vorbildwirkung gleichgeschlechtlicher älterer Geschwister wurden Daten der Mikrozensus 2008 und 2009 zum Übergang von der Grundschule auf das Gymnasium genutzt.

## ***Geschlechtstypisierte Lernangebote und Erzieherin-Kind-Beziehungsqualität im Kindergarten: Einfluss auf die Rechtschreibkompetenzentwicklung von Mädchen und Jungen***

Im Bereich der vorschulischen Bildung sind Erziehungspersonen fast ausnahmslos weiblich. Während die Politik in Reaktion auf die "Bildungskrise" der Jungen und auf Defizite von Mädchen in der Mathematik reflexhaft "mehr Männer für die Kitas" fordert, ist bisher ungeklärt, auf welche Weise das Geschlecht der Erziehungsperson im Kindergarten überhaupt einen Einfluss auf die Kompetenzentwicklung der Mädchen und Jungen haben sollte. Unsere Studie prüft die Annahme, dass es die Geschlechtstypizität von Lernaktivitäten ist, die die Erziehungspersonen den Kindern "vorlebt" oder anbietet, und die Qualität der Beziehung, die die Erziehungsperson zu Mädchen bzw. Jungen eingeht. Empirisch ist gut belegt, dass Personen ihre Interessen und Aktivitäten geschlechtstypisiert ausrichten, z.B. verbringen Frauen mehr Zeit mit haushaltsbezogenen und zwischenmenschlichen Aktivitäten, Männer hingegen vergleichsweise mehr mit sportbezogenen und außerhäuslichen Aktivitäten (Athenstaedt, Mikula, & Bredt, 2009; Hyde, 2013; Sue, Rounds, & Armstrong, 2009). Vor diesem Hintergrund haben wir angenommen, dass auch Erziehungspersonen im Kindergarten vorzugsweise Aktivitäten "vorleben" und anbieten, die zu ihrer eigenen Geschlechtszugehörigkeit kongruent sind (z.B. eine Erzieherin wahrscheinlicher mit den Kindern bastelt als Fußball spielt). Weiter haben wir vor dem Hintergrund der Befunde der Bindungsforschung (Ahnert, Pinquart & Lamb, 2006; Pianta & Stuhlman, 2004) erwartet, dass Erziehungspersonen in der Kita zu Mädchen eine engere Bindung herstellen als zu Jungen.

In unserer Studie (Wolter, Glüer & Hannover, submitted) haben wir 135 unabhängige Erzieherin-Kind-Dyaden aus verschiedenen Berliner Kindergärten untersucht. Die Erzieherinnen wurden zu den Lernangeboten befragt, die sie typischerweise in einer Woche in ihrer Kindergartengruppe anbieten. 12 unabhängige Expert/inn/en kategorisierten die angebotenen Aktivitäten danach, ob sie von Mädchen stärker präferiert werden, von Jungen oder von beiden Geschlechtern gleichermaßen gemocht werden. Die endgültige Skala umfasste 6 männlich-typisierte (z.B. Umgang mit audio-visuellen Medien oder komplexen Konstruktionsspielzeugen), 6 weiblich-typisierte (z.B. Singen und Tanzen, Rollenspiele) und 7 geschlechtsneutrale Aktivitäten (z.B. Basteln, Versteckspielen). Weiter beschrieben die Erzieherinnen die Nähe der Beziehung zum jeweiligen Kind auf einer Subskala der Student-Teacher Relationship Scale (Pianta, 2001). Die Kompetenzentwicklung der Kinder wurde durch eine zweifache Messung (am Ende der Kindergartenzeit und am Ende der 1. Schulklasse) der Rechtschreibkompetenz erfasst.

Erwartungsgemäß machten die Erzieherinnen deutlich mehr weiblich als männlich geschlechtstypisierte Aktivitätsangebote und gingen zu Mädchen qualitativ hochwertigere Bindungen ein als zu Jungen. In dem Ausmaß, in dem die Erzieherin einem Kind, zu dem es in einer nahen Beziehung stand, Lernangebote machte, die in ihrer Konnotation zum biologischen Geschlecht des Kindes "passten", profitierte dieses Kind in seiner Kompetenzentwicklung: Jungen zeigten höhere Leistungen in Rechtschreibung am Ende der ersten Klasse, wenn sie im Kindergarten eine nahe Beziehung zur Erzieherin gehabt hatten, die im Mittel stärker maskulin konnotierte als feminin konnotierte Lernangebote gemacht hatte, während Mädchen von einer nahen Beziehung zur Erzieherin in ihrer späteren schulischen Rechtschreibkompetenz in dem Ausmaß profitierten, wie die Erzieherin im Kindergarten im Mittel stärker feminin als maskulin konnotierte Lernangebote gemacht hatte.

In einer Folgestudie (Wolter & Oakhill, in Vorbereitung) haben wir Lehrerinnen in Grundschulen in England zu ihren Lernangeboten im Lesen befragt und die von den Lehrkräften angegebenen Aktivitätsangebote und Materialien nach ihrer Geschlechtskonnotation eingeschätzt. Erste Ergebnisse deuten ebenfalls auf eine weibliche Konnotation der Lernangebote hin.

In Hinblick auf die Debatte zur Überrepräsentation von Frauen in der frühen Bildung sprechen unsere Ergebnisse dafür, dass es nicht das biologische Geschlecht der Lehrperson ist, welches einen differenziellen Effekt auf die Kompetenzentwicklung von Jungen und Mädchen hat. Vielmehr können weibliche Erziehungs- und Lehrpersonen Mädchen und Jungen gleichermaßen fördern, wenn sie darauf achten, nicht vorrangig Angebote zu machen, die ausschließlich eine Geschlechtsgruppe ansprechen und Mädchen und Jungen gleichermaßen in alle Lernaktivitäten, auch solche, die in den Augen der Kinder nicht zu ihrem Geschlecht "passen, einzubeziehen. In der Folge könnten auch qualitativ hochwertige Beziehungsqualitäten zu männlichen Kindern entstehen.

Ahnert, L., Pinquart, M., & Lamb, M. (2006). Security of children's relationships with nonparental care providers: a meta-analysis. *Child Development*, 74, 664-679.

Athenstaedt, U., Mikula, G., & Bredt, C. (2009). Gender role self-concept and leisure activities of adolescents. *Sex Roles*, 60, 399-409.

Hyde, J. S. (2013). Gender similarities and differences. *Annual Review of Psychology*, Online first, retrieved from <http://www.annualreviews.org/doi/pdf/10.1146/annurev-psych-010213-115057> (expected final online publication date is 2014, 65).

Pianta, R. (2001). STRS: Student-Teacher Relationship Scale: Professional Manual. Florida: Psychological Assessment Resources.

Pianta, R. & Stuhlman, M. (2004). Teacher-child relationships and children's success in the first years of school. *School Psychology Review*, 33, 444-458.

Sue, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, 135, 859-884.

Wolter, I., Glüer, M., & Hannover, B. (submitted). Gender-typicality of activity offerings and child-teacher relationship closeness in German "Kindergarten". Influences on the development of spelling competence as an indicator of early basic literacy in boys and girls.

Wolter, I. & Oakhill, J. (in prep.). Gender-typed learning offerings in primary school: Teacher's impact on gender differences in reading competence.



Ann-Katrin Denn (Universität Kassel), Claudia Kastens (Universität Wuppertal), Caroline Theurer (Universität Kassel), Miriam Lotz (Universität Bamberg), Frank Lipowsky (Universität Kassel)

## **Wer hat im Unterricht was zu melden? Geschlechtsspezifische Lehrer-Schüler-Kommunikation und Zusammenhänge mit Leistung und Selbstkonzept**

Für die Domäne Mathematik gilt als stabiler Befund, dass auch unter Kontrolle der Mathematikleistungen Jungen ein höheres Selbstkonzept aufweisen als Mädchen (Bossong, 2006; Dickhäuser & Stiensmeier-Pelster, 2003; Eccles, Wigfield, Harold & Blumenfeld, 1993; Gabriel, Mösko & Lipowsky, 2011; Tiedemann & Faber, 1995). Lehmann (2006) sieht als eine mögliche Ursache für dieses Befundmuster die Benachteiligung von Mädchen in der Unterrichtskommunikation im Mathematikunterricht an. In Studien zur Lehrer-Schüler-Interaktion wurden Jungen mehrfach als die präferierten Interaktionspartner der Lehrer nachgewiesen (Brophy & Good, 1970; Jones & Wheatley, 1990; Myhill, 2002), die nicht nur mehr Möglichkeiten erhielten, am Unterricht teilzunehmen, sondern auch mehr Lob und Aufmerksamkeit der Lehrperson bekamen (Frasch & Wagner, 1982) und dadurch mehr Gelegenheiten hatten, positive Selbstkonzepte auszubilden. Neben geschlechtsabhängigen Unterschieden in der Unterrichtskommunikation werden auch Schülermerkmale wie Leistung und Selbstkonzept mit der Beteiligung am Unterricht in Zusammenhang gebracht. Schüler mit hoher Leistungsstärke und hohem Selbstkonzept scheinen mehr Interaktionsmöglichkeiten im Unterricht zu haben als Schüler mit ungünstigeren Lernvoraussetzungen (Gage & Berliner, 1986; Jurik, Gröschner, & Seidel, 2013; Lipowsky, Pauli, & Rakoczy, 2008).

Ziel der vorliegenden Studie ist es, das Aufrufverhalten von Lehrpersonen im Grundschul-Mathematikunterricht und das Meldeverhalten der Schüler unter besonderer Berücksichtigung etwaiger geschlechtsspezifischer Unterschiede zu analysieren. Im Fokus steht dabei die Frage, mit welchen Schülern die Lehrperson im Unterrichtsgespräch hauptsächlich interagiert und ob sich in ihrem Aufrufverhalten eine Benachteiligung der Mädchen nachweisen lässt. Zusätzlich soll analysiert werden, ob Selbstkonzept und Leistung der Schüler in Zusammenhang mit dem Aufrufverhalten der Lehrperson oder dem Meldeverhalten der Schüler stehen und ob sich differenzielle Effekte des Aufrufverhaltens der Lehrperson auf die Selbstkonzeptentwicklung der Schüler ergeben.

Die Daten stammen aus dem Projekt PERLE, das die Persönlichkeits- und Lernentwicklung von Grundschulkindern über einen Zeitraum von vier Jahren untersucht (Lipowsky, Faust & Kastens, 2013) und unter anderem eine Videostudie im Fach Mathematik während des zweiten Schuljahres beinhaltet. Die Stichprobe besteht aus 19 Klassen mit N=382 Schülern (N weiblich = 208; N männlich = 176). Das Meldeverhalten der Schüler und das Aufrufverhalten der Lehrperson wurden anhand von niedrig inferenten Kodierungen während des öffentlichen Unterrichtsgesprächs erhoben. Anhand der Kodierungen lässt sich feststellen, wie oft sich ein Schüler im Unterricht gemeldet hat, wie viele Male er von der Lehrperson aufgerufen wurde und ob diese Aufrufe mit vorheriger oder ohne vorherige Meldung des Schülers erfolgten. Die Kodierungen wurden von geschulten Beobachtern durchgeführt und durch Übereinstimmungsberechnungen begleitet, die immer über den geforderten Mindestwerten lagen (prozentuale Übereinstimmung > 85%). Das Selbstkonzept der Schüler wurde durch einen eigens entwickelten Fragebogen erfasst (Poloczek, Karst, Praetorius & Lipowsky, 2011), zur Erfassung der Leistungen wurde ein standardisierter Test verwendet, mit dem die mathematischen Fähigkeiten der Schüler erfasst wurden (Karst, Mösko, Lipowsky & Faust, 2011).

Erste Analysen zum Aufrufverhalten der Lehrkräfte zeigen keine Unterschiede im Aufrufen von Mädchen und Jungen. Interaktionsanalysen unter Berücksichtigung von Leistung und Selbstkonzept deuten aber an, dass Mädchen mit einem hohen Selbstkonzept eher aufgerufen werden, wenn sie sich vorher gemeldet haben, während Jungen mit einem hohen Selbstkonzept häufiger ohne vorherige Meldung aufgerufen werden. Die individuellen Leistungen der Schüler scheinen für das Aufrufverhalten der Lehrpersonen weniger relevant zu sein. Insgesamt fallen Jungen durch mehr Redebeiträge im Unterricht auf, da sie sich Beitragmöglichkeiten durch eine höhere Zahl von Zwischenrufe verschaffen.

In einem ergänzenden Analyseschritt wird im Vortrag neben dem Aufrufverhalten der Lehrkräfte das Meldeverhalten der Schüler betrachtet. Es ist anzunehmen, dass neben dem Geschlecht auch das Selbstkonzept und die Leistung als Prädiktoren für das Meldeverhalten der Schüler fungieren. Zur Prüfung dieser Annahme werden Aufrufquoten berechnet, welche die Meldehäufigkeit der Schüler in Bezug zur Anzahl der Aufrufe setzen.

Bossong, B. (2006). Wie sympathisch bin ich meiner Lehrerin und wie schätzt sie meine Fähigkeiten ein? Zum Einfluss wahrgenommener Lehrereinstellungen auf Selbstkonzepte der Fähigkeit bei Grundschulern. In I. Hosenfeld & F.W. Schrader (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven* (S. 263- 274). Münster: Waxmann.

Brophy, J. & Good, T. (1970). Teachers' communication expectations for children's classroom performance: some behavioral data. *Journal of Educational Psychology*, 61, 365-374.

Dickhäuser, O. & Stiensmeier-Pelster, J. (2003). Wahrgenommene Lehrereinschätzungen und das Fähigkeitsselbstkonzept von Jungen und Mädchen in der Grundschule. *Psychologie in Erziehung und Unterricht*, 50, 182-190.

Eccles, J., Wigfield, A., Harold, R.D. & Blumenfeld, P. (1993). Age and gender difference in children's self- and task perceptions during elementary school. *Child Development*, 64(3), 830-847.

Frasch, H. & Wagner, A. C. (1982). "Auf Jungen achtet man einfach mehr..." In I. Brehmer (Hrsg.), *Sexismus in der Schule. Der heimliche Lehrplan der Frauendiskriminierung* (S. 260-278). Weinheim: Beltz.

Gabriel, K., Mösko, E. & Lipowsky, F. (2011). Selbstkonzeptentwicklung von Jungen und Mädchen im Anfangsunterricht - Ergebnisse aus der PERLE-Studie. In F. Hellmich (Hrsg.), *Selbstkonzepte im Grundschulalter. Modelle - empirische Ergebnisse - pädagogische Konsequenzen* (S. 133-158). Stuttgart: Kohlhammer.

Gage, N.L. & Berliner, D.C. (1986). *Pädagogische Psychologie*. Weinheim: Beltz.

Karst, K., Mösko, E., Lipowsky, F. & Faust, G. (Hrsg.) (2011). *Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulern“ (PERLE)*. 2. PERLE-Instrumente: Schüler, Eltern (Messzeitpunkte 2 & 3). Materialien zur Bildungsforschung, 23/3. Frankfurt am Main: GFPP.

Jones, M.G. & Wheatley, J. (1990). Gender differences in teacher-student interactions in science classrooms. *Journal of Research in Science Teaching* 27(9), 861-874.

Jurik, V., Gröschner, A. & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. *Learning and Instruction*, 23, 33-42.

Lehmann, R. (2006). Mädchen und Mathematik in der gymnasialen Sekundarstufe I - Ergebnisse einer Längsschnittstudie. In I. Hosenfeld (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven* (S. 107-120). Münster: Waxmann.

Lipowsky, F., Pauli, C. & Rakoczy, K. (2008). Schülerbeteiligung und Unterrichtsqualität. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise. Analyse und Bedeutung unterrichtlichen Handelns* (S. 67-90). Münster: Waxmann.

Lipowsky, F., Faust, G. & Kastens, C. (Hrsg.) (2013). *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren*. Münster: Waxmann.

Myhill, D. (2002). Bad boys and good girls? Patterns of interaction and response in whole class teaching. *Cambridge Journal of Education* 34(1), 35-49.

Poloczek, S., Karst, K., Praetorius, A.-K. & Lipowsky, F. (2011). Generalisten oder Spezialisten? Bereichsspezifität und leistungsbezogene Zusammenhänge des schulischen Selbstkonzepts von Schulanfängern. *Zeitschrift für Pädagogische Psychologie*, 25(3), 173-183.

Tiedemann, J. & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27(1), 61-71.



Verena Jurik (TUM School of Education), Janina Häusler (TUM School of Education), Sina Stubben (TUM School of Education), Tina Seidel (TUM School of Education)

## „Interaction“ - Die Rolle kognitiver und motivational-affektiver Schülervoraussetzungen für Lehrer-Schüler-Interaktionen im Deutsch- und Mathematikunterricht

### 1. Theoretischer Hintergrund

Individuelle Schülervoraussetzungen wie Vorwissen, Interesse und Selbstkonzept spielen eine wichtige Rolle für das erfolgreiche Lernen (Shuell, 1996). Bisherige Studien zum Zusammenspiel kognitiver und motivational-affektiver Schülervoraussetzungen konzentrieren sich üblicherweise auf ein Schulfach (Seidel, 2011). Fächervergleichende Studien im motivational-affektiven Bereich zeigen, dass Schüler mit einem hohen Selbstkonzept der Begabung in Mathematik mit einer hohen Wahrscheinlichkeit ein niedriges Selbstkonzept für das Fach Deutsch aufweisen und umgekehrt (Marsh, Walker, & Debus, 1991). Darüber hinaus liegt ein systematischer Zusammenhang mit genderspezifischen Attributen von Schulfächern vor. Die Fächer Mathematik und Naturwissenschaften gelten als männlich konnotiert, während Deutsch als weiblich konnotiert wahrgenommen wird (Hannover & Kessels, 2004). Im Physikunterricht wurden bereits fünf Schülerprofile entlang der Dimensionen kognitiver Grundfähigkeit, Vorwissen, Interesse und Selbstkonzept der Begabung identifiziert (Seidel, 2006). Die Profile umfassen Gruppen von (a) über alle Dimensionen starke, (b) kognitiv fähige aber Universitätsinteressierte, (c) kognitive fähige aber motivational-affektiv unterschätzende, (d) kognitiv eingeschränkt fähig aber motivational-affektiv überschätzende und (e) allgemein schwache Schüler. Bezüglich dieser Schülerprofile sind Mädchen mit einer höheren Wahrscheinlichkeit in den Gruppen (b), (c) und (e) vertreten. Die vorgestellte Studie „Interaction“ überträgt die Profilanalysen auf den Deutsch- und Mathematikunterricht, um die Robustheit des gewählten methodischen Zugangs zu prüfen. Durch das differenzierte Zusammenspiel der verschiedenen Aspekte (kognitive und motivational-affektive Voraussetzungen, Geschlecht und Fächer) werden die Befunde bisheriger Studien erweitert.

### 2. Fragestellungen

Im vorliegenden Beitrag wurden individuelle Schülerprofile im Deutsch- und Mathematikunterricht untersucht. Die Fragestellungen lauten:

(1) Lassen sich die für den Physikunterricht beschriebenen Profile für die Fächer Deutsch und Mathematik in der achten Jahrgangsstufe in Gymnasien replizieren?

Es wurde angenommen, dass die für den Physikunterricht beschriebenen Schülerprofile repliziert werden können.

(2) Zeigen sich für Mädchen und Jungen Geschlechterunterschiede in den Profilverteilungen für den Deutsch- und Mathematikunterricht?

In Bezug auf genderspezifische Unterschiede wurde erwartet, dass Mädchen im Fach Deutsch häufiger über alle Dimensionen starke Profile verfügen, während sie in Mathematik häufiger schwache Profile aufweisen. Eine umgekehrte Verteilung wurde für Jungen angenommen.

### 3. Methode

Die Stichprobe umfasste N = 16 Gymnasialklassen der 8. Jahrgangsstufe aus Bayern. Die Schülerinnen und Schüler wurden bezüglich ihrer Voraussetzungen zu Beginn des Schuljahres 2013/2014 in einer Unterrichtsstunde (45 Minuten) getestet und befragt. Erhebungsverfahren für individuelle Schülervoraussetzungen: Die Verfahren zur Bestimmung der Schülerprofile umfassten die Bereiche kognitive Fähigkeiten, Vorwissen in Mathematik und Lesen, Interesse am Fach, Selbstkonzept der Begabung. Kognitive Fähigkeiten: Die kognitiven Fähigkeiten der Schülerinnen und Schüler wurden mit der Subskala Figurenanalogien des Kognitiven Fähigkeitstest (KFT) von Heller und Perleth (2000) erfasst. Die Subskala umfasste 25 Items. Vorwissen in Mathematik und Lesen: Das Vorwissen der Schülerinnen und Schüler wurde über die letzten Jahreszeugnisnoten in den Fächern Mathematik und Deutsch erfasst. Interesse an Mathematik und Deutsch: Das Interesse wurde über PISA-Skalen mit vier Items für den Mathematikunterricht und drei Items für den Deutschunterricht erfasst. Selbstkonzept: Das Selbstkonzept wurde für Mathematik mit fünf PISA-Items und für Deutsch mit sechs PISA-Items erfasst. Zur Bestimmung der Schülerprofile wurden latente Klassenanalysen (LCA) mit Hilfe der Software Mplus durchgeführt.

### 4. Ergebnisse

Erste Analysen replizierten die für den Physikunterricht beschriebenen Profile. Darüber hinaus zeigten sich geschlechterspezifische Unterschiede in den Profilverteilungen für den Mathematik- und Deutschunterricht.

### 5. Ausblick

Im Verlauf des Schuljahres 2013/2014 wird es im Projekt „Interaction“ zwei weitere Messzeitpunkte geben: In der Mitte des Schuljahres wird je eine Unterrichtsstunde im Mathematik- und Deutschunterricht videographiert. Zum Schuljahresende werden die Schülerinnen und Schüler erneut zu ihren Voraussetzungen befragt. So können Effekte der Schülerprofile und deren Entwicklung untersucht werden. Darüber hinaus wird auf die Rolle der Lehrkräfte fokussiert: Diese schätzen zum einen zu Schuljahresbeginn und -ende die Schülerinnen und Schüler hinsichtlich ihrer Voraussetzungen ein. Zum anderen wird eine Analyse der videographierten Lehrer-Schüler-Interaktionen stattfinden.

Hannover, B., & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction*, 14(1), 51–68.

Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision. Göttingen: Beltz Test GmbH.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology*, 16, 331–345.

Seidel, T. (2006). The role of student characteristics in studying micro teaching-learning environments. *Learning Environments Research*, 9(3), 253–271.

Seidel, T. (2011). Lehrerhandeln im Unterricht. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 605–629). Münster: Waxmann.

Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. Calfee (Hrsg.), *Handbook of Educational Psychology* (S. 726–764). New York.

Chair(s): Christian Weinzierl (Leibniz Universität Hannover), Jörg Jost (Universität Paderborn)

Diskutant/in: Markus Schmitt (Pädagogische Hochschule Heidelberg)

Im Zentrum des Symposiums steht die Frage nach verschiedenen Perspektiven auf den Forschungsgegenstand Textproduktion und -rezeption in Lehr- und Lernkontexten. Fragen nach der Kompetenz von Lernern, produktiv wie rezeptiv mit Texten umzugehen (i. e. Textkompetenz), verlangen eine differenzierte Betrachtung beider Perspektiven jeweils für sich genommen wie auch wechselseitig zueinander.

Aus den Blickwinkeln der Psychologie sowie der Sprachdidaktik wird über Forschungszugänge berichtet, die sich durch eine unterschiedliche Gewichtung des Produktions- bzw. Rezeptionsaspekts sowie durch verschiedene Blickwinkel auf deren wechselseitiges Zusammenspiel auszeichnen. Neben Unterschieden in methodischen Zugangsweisen und Korngrößen der Betrachtung werden Gemeinsamkeiten und (potenzielle) Synergien fokussiert, die durch eine verstärkte interdisziplinäre Zusammenarbeit befördert werden können:

- Der Beitrag von Jörg Jost, Matthias Knopp, Michael Becker-Mrotzek und Joachim Grabowski fragt aus schreibdidaktischer und psychologischer Perspektive im Kontext der Erforschung von Teilkomponenten von Schreibkompetenz nach dem wechselseitigen Verhältnis von Produktion und Rezeption bei der Herstellung von Kohärenz sowie nach dem Zusammenhang zwischen Kohärenz und Perspektivenübernahme. Dabei wird eine integrierte Betrachtungsweise aus Sicht beider Fächer eingenommen, was sich daran zeigt, dass bei der Konzeption, der Operationalisierung der Variablen sowie der Datenanalyse beide Perspektiven gleichberechtigt Eingang finden.
- Christian Weinzierl fragt in seinem Beitrag aus Perspektive der Psychologie im Rahmen eines experimentellen Zugangs nach den kognitiven Grundlagen, die für die Reproduktion eines vorhandenen Textes durch handschriftliches Abschreiben nötig sind. Dabei wird gezeigt, wie eng rezeptive (Enkodierung des Vorlagentextes) und produktive (graphomotorische Transkription) Aspekte beim Abschreiben zusammenspielen. Zusätzlich werden Anknüpfungspunkte an kognitive Grundlagen der Textrezeption bzw. des Lesens diskutiert.
- In dem Beitrag von Wolfgang Lenhard, Herbert Baier, Alexandra Lenhard, Wolfgang Schneider und Joachim Hoffmann wird aus psychologischer Perspektive der Frage nachgegangen, inwieweit sich das Training von Textproduktionsfertigkeiten gezielt zur Förderung rezeptiver Fertigkeiten des Leseverständnisses nutzen lässt. Durch den Einbezug schreibdidaktischer Kriterien bei der Konzeption der Trainings bietet der Vortrag auch Verbindungspunkte zur Perspektive der Schreibdidaktik.
- Vor dem Hintergrund einer Interventionsstudie gehen Steffen Otteberg und Astrid Neumann in ihrem sprachdidaktischen Beitrag der Frage nach, ob und wie fachliches Lernen durch das Bearbeiten von „Schreibaufgaben mit Profil“ in Storytelling-Arrangements im Fachunterricht unterstützt werden kann. Sie fragen weiterhin danach, welchen Einfluss das gemeinsame, durch mündliches Aushandeln unterstützte Schreiben auf das fachliche Lernen hat.
- Markus Schmitt fasst in seinem Diskussionsbeitrag die im Symposium eingenommenen disziplinären Perspektiven zusammen und systematisiert diese im Hinblick auf theoretische und methodische Anschlussmöglichkeiten. Dabei werden auch Synergien thematisiert, die sich aus der interdisziplinären Beschäftigung für die Erforschung des komplexen Forschungsgegenstands ergeben.

- Bachmann, T. & Becker-Mrotzek, M. (2010). Schreibaufgaben situieren und profilieren. In T. Pohl & T. Steinhoff (Hrsg.), Textformen als Lernformen (S. 191–220). Köln: Gilles & Francke.
- Carey, S.; Evans, R.; Honda, M.; Jay, E. & Unger C. (1989). 'An experiment is when you try it and see if it works': a study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education* 11, 514–529.
- Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach. In M. J. Snowling & C. Hulme (Hrsg.), *The science of reading* (S. 6–23). Malden, MA: Blackwell Pub.
- Galbraith, D. (2009). Writing as discovery. *British Journal of Educational Psychology* 6, 1–23.
- Grabowski, J., Weinzierl, C., & Schmitt, M. (2010). Second and fourth graders' copying ability: from graphical to linguistic processing. *Journal of Research in Reading*, 33(1), 39–53.
- Hattie, J. (2009). *Visible Learning. A Synthesis of over 800 Meta-analyses relating to achievement*. London/New York: Routledge.
- Hayes, J. R. (2012). Modeling and Remodeling Writing. *Written Communication*, 29(3), 369–388.
- Knopp, M., Jost, J., Nachtwei, N., Becker-Mrotzek, M., & Grabowski, J. (2012). Teilkomponenten von Schreibkompetenz untersuchen. Bericht aus einem interdisziplinären empirischen Projekt. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, M. Rothgangel, L.-H. Schön, H. J. Vollmer & H.-G. Weigand (Eds.), *Formate Fachdidaktischer Forschung. Empirische Projekte – historische Analysen – theoretische Grundlegungen Fachdidaktische Forschungen* (Vol. 2, S. 47-66). Münster, New York, München, Berlin: Waxmann.
- Lenhard, W., Baier, H., Lenhard, A., Hoffmann, J. & Schneider, W. (2013). *conText - Förderung des Leseverständnisses durch das Arbeiten mit Texten*. Göttingen: Hogrefe.
- Neumann, A. (2013). Schreiben in der Sekundarstufe II. In S. Gailberger & F. Wietze (Hrsg.), *Handbuch Kompetenzorientierter Deutschunterricht* (255–275). Baltmannsweiler: Beltz.
- Otteberg, Steffen (2012): *Lesen, Schreiben, Experimentieren – Wissenschaftsverständnis und Textkompetenz integrativ fördern*. Lüneburg (unveröffentlicht).
- Rickheit, G. (Ed.). (1991). *Kohärenzprozesse. Modellierung von Sprachverarbeitung in Texten und Diskursen*. Opladen: Westdeutscher Verlag.
- Schekatz-Schopmeyer, Sonja (2010): *Storytelling – Eine narrative Methode zur Vermittlung naturwissenschaftlicher Inhalte im Sachunterricht der Grundschule*. Göttingen: Cuvillier.
- Schmitt, M. (2011). *Perspektivisches Denken als Voraussetzung für adressatenorientiertes Schreiben*. Dissertation. Pädagogische Hochschule, Heidelberg.
- Tanskanen, S.-K. (2006). *Collaborating towards Coherence. Lexical cohesion in English Discourse* (Vol. 146). Amsterdam, Philadelphia: John Benjamins.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G168





## ***Etablierung und Rekonstruktion von Kohärenz als Indikator für Schreibkompetenz***

Die Fähigkeit, als Schreiber einen kohärenzstiftenden Pfad in Texten anzulegen, wird als eine Teilkomponente von Schreibkompetenz verstanden (vgl. Knopp, Jost, Nachtwei, Becker-Mrotzek, & Grabowski, 2012). Dabei verweist die Kohärenz – auch aus Sicht produktionsseitiger Prozesse – auf die Rezeption (hier: das Lesen). Dahinter steht die Annahme von Kohärenz als mentaler Konstruktionsleistung von Rezipienten (Rickheit, 1991). Im Beitrag wird auf der Grundlage empirischer psychologischer und sprachdidaktischer Forschung der Frage nachgegangen, welchen Zusammenhang es zwischen der Fähigkeit zur Etablierung von Kohärenz in Texten beim Schreiben (Produktebene repräsentiert durch messbare Indikatoren von Kohäsion/Kohärenz) und der Fähigkeit zur Herstellung von Kohärenz als mentaler Konstruktionsleistung (z. B. Schwarz-Friesel, 2001) beim Lesen und Verstehen gibt. Neben der produktiven und rezeptiven Perspektive auf Kohärenz als Teil von Schreibkompetenz wird nach dem Zusammenhang der Etablierung von Kohärenz im Text im Ergebnis von Kooperationsbemühungen (Signalisieren und Erkennen von Kohärenz, z. B. Tanskanen, 2006) einerseits und den Fähigkeiten zur Perspektivenübernahme von Schreibern (Schmitt, 2011) andererseits gefragt. Im Rahmen der Untersuchung wurden eine Reihe verschiedenster Maße erhoben, die von Reaktionszeiten bei Perspektivenübernahme über die Bearbeitung von isolierten Aufgaben zur produktiven und rezeptiven Herstellung von Kohärenz bis hin zur Etablierung von Kohärenz in Texten reichen. Ergebnisse aus Regressionsanalysen zeigen, dass die Textqualität (in den Textsorten Instruktion, Bericht, Argumentation) anhand der Komponenten „Kohärenzstiftung“ und „Adressatenorientierung“ vorhergesagt werden können. Anhand der Analyse linearer Strukturgleichungsmodelle werden die angenommenen Zusammenhangsstrukturen der Variablen zu Kohärenz und Perspektivenübernahme untermauert.

## **Wie erwerben Kinder die Fertigkeit effizient abzuschreiben? Indizien durch Schreibprozessanalysen**

Welche mentalen Vorgänge für effizientes Abschreiben von tragender Bedeutung sind, ist bislang kaum erforscht, obwohl diese Anforderung eine wichtige Grundfertigkeit im Rahmen alltäglichen Unterrichts darstellt (vgl. Grabowski, Weinzierl, & Schmitt, 2010). Im Rahmen eines psychologisch-experimentellen Zugangs wird über die Untersuchung kognitiver Prozesse beim Durchführen reproduktiver Abschreibeaufgaben berichtet, für welche Sprachproduktions- und -rezeptionsaspekte gleichermaßen eine wichtige Rolle spielen. Im Vordergrund steht dabei die Frage nach der Ausbildung kognitiver Strategien für eine verbesserte Effizienz dieser basalen Sprachproduktionsaufgabe in der Primarstufe. Im Rahmen zweier Experimente wurden von Grundschulern der zweiten und vierten Klassenstufe Vorlagen abgeschrieben, die sich im Grad ihrer linguistischen Kodierbarkeit unterscheiden (Text, Ziffernfolgen, Konsonantenreihen, abstrakte geometrische Objekte); neben einer Kontrollbedingung wurden zusätzlich parallel Zweitaufgaben durchgeführt, die die verschiedenen Komponenten des Arbeitsgedächtnisses (phonologische Schleife, visuell-räumlicher Notizblock, zentrale Exekutive) selektiv belasten. Zur Abbildung von Abschreibeffizienz dienen Indikatoren auf Produkt- und Prozessebene (u. a. Häufigkeit von Schreibpausen, Anzahl abgeschriebener Zeichen zwischen zwei Schreibpausen). Die in der varianzanalytischen Untersuchung innerhalb dieses zweifaktoriellen Designs durchgängig signifikanten Haupt- und Interaktionseffekte ( $p < .001$ ) deuten auf die Ausbildung einer arbeitsgedächtnisabhängigen Strategie hin, die bei hoher linguistischer Kodierbarkeit der Vorlage auf die Verarbeitung größerer linguistischer Einheiten im Arbeitsgedächtnis abzielt. Neben dem Verweis auf Chunking-Prozesse im Arbeitsgedächtnis werden die Ergebnisse im Licht des Dual-route-Lesemodells (Coltheart, 2005) diskutiert.

Wolfgang Lenhard (Universität Würzburg), Herbert Baier (IT Baier, Wintershausen), Alexandra Lenhard (Psychometrica, Dettelbach), Wolfgang Schneider (Universität Würzburg), Joachim Hoffmann (Universität Würzburg)

### ***Die Effekte schreibdidaktischer Ansätze auf das Leseverstehen***

Lesen und Schreiben sind wechselseitige Prozesse, die einerseits spezifische Prozesse erfordern, zum anderen aber auch modalitätsunabhängige Kompetenzen teilen. Hieraus ergibt sich die Möglichkeit, schreibdidaktische Ansätze gezielt zur Förderung des Leseverständnisses zu nutzen und beide Aspekte zu fördern, z. B. indem Texte zusammengefasst werden. Zur Überprüfung der Effekte dieser Vorgehensweise trainierten 148 Schülerinnen und Schüler der 6. Klasse der Hauptschule ein Jahr mit einem computerbasierten Programm, das inhaltliches Feedback zu Textzusammenfassungen gibt (Lenhard et al., 2013), erhielten ein Strategietraining oder fungierten als Kontrollgruppe. Zur Analyse des Inhalts der Zusammenfassungen kamen semantische Technologien (sog. Latente Semantische Analyse) zum Einsatz. Sowohl das Strategietraining als auch das computerbasierte Training waren der Kontrollgruppe in einer Reihe lesespezifischer Fähigkeiten (Flüssigkeit, Strategiewissen und Leseverständnis) überlegen, wobei das Computerprogramm effektiver war als das Strategietraining. Die Ergebnisse zeigen, dass die Lesekompetenz von einer Schreibförderung profitieren kann.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G168

## ***Storytelling als Schreibarrangement zur Unterstützung des naturwissenschaftlichen Lernens in der Grundschule***

Schriftsprachkompetenzen sind seit der Einführung der Bildungsstandards ab 2004 in allen fachlichen Domänen gefordert. Schreibkompetenzen werden zur Lernunterstützung, aber auch als Prüfungsform genutzt (Neumann, 2013). „Aufgaben mit Profil“ (Bachmann & Becker-Mrotzek, 2010) in kollaborativen Schreibsituationen erweisen sich als besonders lernförderlich. Diese lernförderlichen Aufgaben können im Fachunterricht in Storytelling-Arrangements (Schekatz-Schopmeyer, 2010) den Wissensaufbau durch Schreiben (Galbraith, 2009) und das Wissensmanagement forcieren. Dies soll am Beispiel der Fallstudie zum „Lesen, Schreiben, Experimentieren“ (Ottoberg, 2012) im naturwissenschaftlichen Unterricht der Klasse 4 gezeigt werden. Zielsetzung der Untersuchung ist zu prüfen, ob fachliches Lernen durch Schreiben forciert werden kann. Dabei soll die Frage beantwortet werden, wie inhaltliches Lernen im Fach durch Schreiben unter a) spezifischen Schreibarrangements und b) in gemeinsamen Schreibsituationen durch mündliche Aushandlungsprozesse unterstützt werden kann.

Die Intervention erfolgte mittels Storytelling in mehreren naturwissenschaftlichen Problemlöseaufgaben. Darüber hinaus wurde das „Nature of Science“-Interview (Carey et al., 1989) an den Kontext der Untersuchung angepasst. Methodisch wurden die Effekte zwischen Pre- und Posttests an den geschriebenen Texten und den Antworten im Interview gemessen. Die Informationen aus den Prozessbeobachtungen während der Experimentier- und Schreibphasen werden zur Erklärung genutzt.

Die Schreibenden verbessern sowohl ihr naturwissenschaftliches Verständnis als auch ihre Schreibfähigkeiten. Die Führung der Lernenden durch das Storytelling ermöglichte dabei vor allem die motivationale Stützung des Schreibprojektes (Hayes, 2012).

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G168

Steffen Schindler (Otto-Friedrich-Universität Bamberg) und Markus Lörz (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

## ***Expansion und institutionelle Differenzierung des Bildungssystems: Soziale Inklusion oder neue Ungleichheiten?***

Chair(s): Steffen Schindler (Universität Bamberg), Markus Lörz (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

Diskutant/in: Rolf Becker (Universität Bern)

Die Bildungsexpansion war begleitet von zahlreichen institutionellen Differenzierungsprozessen. Sowohl im allgemeinen als auch im tertiären Bildungssystem wurden verschiedene alternative Bildungswege eingeführt und sukzessive ausgebaut. In der Sekundarstufe traten beispielsweise neben das klassische allgemeinbildende Gymnasium zahlreiche alternative (berufliche) Bildungswege zur Hochschulreife. Im tertiären Bildungssystem traten die Fachhochschulen, und in jüngster Zeit einige neuartige Institutionen, wie duale Hochschulbildungsgänge, neben die Universitätsversitäten. Es gibt folglich mittlerweile sehr viele verschiedene Wege durchs Bildungssystem und der Anteil der Hochqualifizierten ist in den letzten fünf Jahrzehnten deutlich gestiegen.

Vor dem Hintergrund dieser veränderten Rahmenbedingungen beschäftigt sich das Symposium mit der Frage, inwieweit sich die verschiedenen Differenzierungsprozesse auf die sozialen Ungleichheiten beim Erwerb von höherer Bildung ausgewirkt haben. Führt institutionelle Differenzierung zur Inklusion sozialer Gruppen, die beim Zugang zu höherer Bildung bislang unterrepräsentiert waren? Oder führt die Diversifizierung von Bildungsgängen zur Ablenkung bestimmter sozialer Gruppen in weniger prestigeträchtige Bildungs- und Berufswege und zur Entstehung neuer Ungleichheiten? Im Symposium sollen diese Fragen aus unterschiedlichen disziplinären Blickwinkeln (Erziehungswissenschaft, Ökonomie, Soziologie) für verschiedene Etappen des Bildungssystems untersucht werden. Die Unterscheidung zwischen vertikalen und horizontalen Ungleichheiten und deren Entwicklung in den vergangenen Jahrzehnten steht dabei ebenso im Vordergrund wie die Frage nach den Mechanismen, die diesen Entwicklungen zugrundeliegen.

Der Beitrag von Steffen Schindler zeichnet die Differenzierungsprozesse im Sekundarschulsystem seit den 1970er-Jahren nach und veranschaulicht, wie sich dies auf die sozialen Ungleichheiten beim Erreichen der Hochschulreife und im Übergang an die Hochschulen ausgewirkt hat. Der Beitrag von David Reimer, Kristian Karlson, Mads Meier Jæger und Anders Holm stellt eine Analyse zu ähnlichen Entwicklungen im dänischen Schulsystem vor, die den kausalen Einfluss der Differenzierungsprozesse auf die Bildungsungleichheit anhand ökonomischer Verfahren bestimmt. Der Beitrag von Kai Maaz und Michael Becker richtet den Blick auf soziale Selektivitäten beim Übergang ins Studium und beschäftigt sich mit der Frage, welche Mechanismen den beobachtbaren Unterschieden zugrunde liegen. Auf Basis der TOSCA-Studie wird die Bedeutung primärer und sekundärer Herkunftseffekte beim Übergang aus den allgemein- und berufsbildenden Gymnasien an die Hochschulen bestimmt. Der Beitrag von Markus Lörz widmet sich der Frage, inwieweit der Ausbau der beruflichen Bildungswege zu veränderten Ungleichheitsmechanismen beim Übergang ins Studium geführt hat. Hierbei liegt der Fokus auf der Entscheidung an einer Universität oder einer Fachhochschule zu studieren und den sich im Zeitverlauf ändernden Mechanismen sozialer Disparitäten. In einem Diskussionsbeitrag von Rolf Becker werden die Erkenntnisse der Einzelbeiträge in einen Zusammenhang gestellt, vor dem Hintergrund der Ausgangsfragestellung des Symposiums diskutiert und Perspektiven für Forschung und Bildungspolitik aufgezeigt.



## ***Differenzierungsprozesse im Sekundarschulbereich – Auswirkungen auf die soziale Ungleichheit beim Erreichen der Hochschulreife und in der Studierneigung***

Seit den 1970er-Jahren vollzogen sich im deutschen Sekundarschulsystem einige weitreichende Veränderungen. Insbesondere ist eine Diversifizierung der Wege zur Hochschulreife zu beobachten, da zunehmend mehr Institutionen, insbesondere auch im berufsbildenden Bereich, eine Studienberechtigung vergeben können. So wird heute ein Großteil der Studienberechtigungen eines Abschlussjahrgangs nicht mehr klassisch als Abitur am allgemeinbildenden Gymnasium erworben. Der Beitrag zeichnet diese Entwicklungen nach und beschäftigt sich mit der Frage, inwieweit diese Öffnungsprozesse der Sekundarstufe II zu Veränderungen der sozialen Ungleichheit an den beiden wichtigen Bildungsetappen auf dem Weg an die Hochschulen geführt haben: Dem Erreichen der Hochschulreife und dem Übergang an die Hochschulen. Es wird argumentiert, dass die Diversifizierungsprozesse im Sekundarschulbereich zur Reduzierung der Ungleichheit beim Erreichen eines Abschlusses der Sekundarstufe II beigetragen haben, dass jedoch innerhalb der Sekundarstufe II neue (qualitative) Ungleichheiten entstanden sind, die sich anhand der Unterscheidungen zwischen beruflichen/allgemeinbildenden Institutionen oder Abitur/Fachhochschulreife festmachen lassen. Durch diese qualitativen Unterschiede innerhalb der Studienberechtigung werden zudem sozialgruppenspezifische Veränderungen in der Studienberechtigtenpopulation vermittelt, die zu einem Anstieg sozialer Disparitäten in den Studierneigungen geführt haben. Trotz dieser Prozesse scheint die Diversifizierung im Sekundarschulbereich jedoch insgesamt zu einem leichten Abbau sozialer Ungleichheiten beim Hochschulzugang beizutragen.

David Reimer (Universität Aarhus), Mads Meier Jaeger (Universität Kopenhagen), Kristian Karlson (Universität Aarhus), Anders Holm (Universität Kopenhagen)

## ***Die Rolle von beruflichen Gymnasien bei der Generierung von Bildungsungleichheit in Dänemark***

Ziel des vorliegenden Beitrags ist es, die Rolle von beruflichen Gymnasien bei der Generierung von Bildungsungleichheit in Dänemark zu untersuchen. In Dänemark wurden, ähnlich wie in Deutschland, neben den etablierten allgemeinbildenden Gymnasien in den 1970er und 1980er Jahren berufliche Gymnasien geschaffen, mit dem Ziel, breiteren Bevölkerungsschichten den Erwerb des Abiturs und den Zugang zur Tertiärbildung zu ermöglichen. In unserer Untersuchung schätzen wir auf Basis von informationsreichen dänischen Registerdaten für zwei Kohorten, die im Jahr 2001 und 2002 die neunte Klasse (Folkeskole) abgeschlossen haben, den Effekt von beruflichen vs. allgemeinbildenden Gymnasien auf die Wahl einer Hochschulausbildung. Die von uns geschätzten Übergangsmo­delle stellen explizit unbeobachtete Unterschiede zwischen Schülern der verschiedenen Gymnasialarten in Rechnung. Wir kommen zum Ergebnis (1) dass die beruflich orientierten Gymnasien gegenüber den allgemeinbildenden Gymnasien sozial weniger selektiv sind, d.h. dieser Schultyp wird vermehrt von Schülerinnen und Schülern aus weniger privilegierten Verhältnissen gewählt (2) dass der Besuch eines beruflichen Gymnasiums einen deutlichen negativen Effekt auf die Wahrscheinlichkeit, ein Hochschulstudium aufzunehmen, hat und (3) insgesamt das Vorhandensein des beruflichen Gymnasiums die Chancen von Schülerinnen und Schülern aus unteren Herkunftsschichten, ein Studium zu wählen, erhöht. Allerdings ist dies mehrheitlich auf die Wahl eines Hochschultyps unterhalb des Universitätversitätsniveaus zurückzuführen. Zusammenge­nommen deuten diese Resultate auf ein interessantes Paradox hin. Die Ausdifferenzierung von Gymnasialtypen hat in Hinblick auf den Eintritt in ein Hochschulstudium negative Konsequenzen auf der Individualebene, da berufliche Gymnasien in gewisser Weise eine „Ablenkungsfunktion“ ausüben. Auf der Makroebene hingegen führen berufliche Gymnasialtypen zu einer Verringerung von Bildungsungleichheit. Wir diskutieren die zugrundeliegenden Mechanismen und Implikationen unserer Befunde – insbesondere auch in Hinblick auf Resultate aus anderen Ländern.

## ***Soziale Ungleichheit beim Hochschulzugang: Eine Zerlegung in primäre und sekundäre Effekte der sozialen Herkunft***

Seit den 1960er Jahren stehen eine Expansion und eine Öffnung weiterführender Bildungsgänge, eine höhere Partizipation an Hochschulbildung und eine größere Chancengerechtigkeit beim Hochschulzugang im Zentrum der Diskussion um das Bildungswesen. Eine Öffnung der Wege zur Hochschulreife und zur Hochschule sollte einerseits dem Bedarf an Hochschulabsolventen bzw. qualifizierten Arbeitskräften nachkommen und andererseits die sozialen Disparitäten in der Bildungsbeteiligung entschärfen. Dass auch an dieser späten Übergangsschwelle, beim Übergang in ein Hochschulstudium, Effekte der sozialen Herkunft zum Tragen kommen ist mittlerweile in einigen empirischen Studien belegt worden. Vor diesem Hintergrund stellt sich die Frage, auf welche Weise Merkmale sozialer Herkunft beim Übergang von der Schule in die Hochschule wirksam werden, wie groß deren Effekte sind und wie man den sozialen Disparitäten beim Hochschulzugang begegnen kann. Hierbei gilt in Anlehnung von Boudon (1974) zu unterscheiden, inwieweit die Wirkungen sozialer Herkunft auf die Bildungsbeteiligung im tertiären Bereich (a) kumulative Effekte der mehr oder weniger gelungenen Unterfütterung der Schullaufbahn darstellen – dies bezieht sich vor allem auf den Erwerb der für den Hochschulzugang vorausgesetzten Merkmale, also maßgeblich Noten und Kompetenzen (primäre Effekte) – oder aber (b) auf bildungs- und sozialschichtabhängige Bildungsaspirationen und -entscheidungen (sekundäre Effekte) zurückgeführt werden können. Für die wissenschaftliche Auseinandersetzung mit sozialen Ungleichheiten der Bildungsbeteiligung hat sich die von Boudon vorgenommene Unterscheidung sozialer Herkunftseffekte in primäre und sekundäre Effekte bzw. Disparitäten als sehr nützlich erwiesen. In vielen Untersuchungen wurde diese Unterscheidung berücksichtigt und ihre Bedeutung für unterschiedliche Bildungsübergänge nachgewiesen. Einigkeit besteht darin, dass sowohl primäre als auch sekundäre Effekte auch beim Übergang in ein Hochschulstudium eine Rolle spielen. Analysen zur relativen Bedeutung von Herkunftseffekten ergaben allerdings auch, dass Ungleichheiten am Übergang zur Hochschule überwiegend auf sekundäre Effekte, also Unterschiede im Entscheidungsverhalten, zurückzuführen scheinen (Schindler & Reimer, 2010). Primäre Herkunftseffekte konnten dagegen nur in begrenztem Maße zur Erklärung der Unterschiede im Übergangsverhalten beitragen (Schindler & Reimer, 2010; Schindler & Lörz, 2011).

Der geplante Beitrag setzt an der aktuellen Forschung zu primären und sekundären Herkunftseffekten beim Übergang in ein Hochschulstudium an. Im Mittelpunkt steht dabei die Zerlegung des sozialen Herkunftseffektes des Hochschulzugangs in primäre und sekundäre Effekte. Es wird (1) überprüft, ob sich sowohl primäre als auch sekundäre Teileffekte identifizieren lassen und wie sich diese Effekte auf die Wahrscheinlichkeit, ein Hochschulstudium aufzunehmen, auswirken. Dabei soll unterschieden werden zwischen Effekten, die direkt auf den Hochschulzugang wirken (ausschließlich sekundäre Effekte) und jene, die sowohl in direkter als auch indirekter Weise (sekundäre und primäre Effekte) den Hochschulzugang beeinflussen. Diese Befunde sollen (2) mit Befunden früherer Bildungsübergänge (insbesondere der Übergang in ein Gymnasium) verglichen werden.

Datengrundlage der vorgestellten Analysen ist die TOSCA-Studie. Mit dieser Studie liegen längsschnittliche Informationen von Abiturienten an allgemeinbildenden und beruflichen Oberstufen in Baden-Württemberg vor. In TOSCA wurden neben den für den Hochschulzugang relevanten Zertifikaten (z.B. Abiturzeugnis bzw. Abiturnote) auch Kompetenzen mit standardisierten Leistungstests erfasst. Damit ist es möglich primäre und sekundäre Effekte korrekt zu spezifizieren. Auf Grundlage binär logistischer Regressionsanalysen mit der Entscheidung Hochschulstudium Ja oder Nein als Kriterium werden die Anteile aufgeklärter Varianz geschätzt und durch die Identifikation von direkten und indirekten Effekten die relativen Anteile von primären und sekundären Herkunftseffekten quantifiziert.

***Differenzierungsprozesse im Hochschulbereich: Haben sich mit dem Ausbau der beruflichen Bildungswege die Ungleichheitsmechanismen verändert?***

In den vergangenen Jahrzehnten haben sich die Zugangswege zum Studium verändert und die Bildungsbeteiligung hat insgesamt deutlich zugenommen. Es führen mittlerweile sehr verschiedene Wege zu einem Studium und zu einem Studienabschluss. Die Einführung und der Ausbau beruflicher Bildungswege im Sekundarschulbereich waren begleitet von zahlreichen Veränderungen im Hochschulbereich. Insbesondere die Einführung der Fachhochschulen Anfang der 1970er Jahre hat das Studium in Deutschland verändert. Trotz dieser Öffnungs- und Veränderungsprozesse der letzten Jahrzehnte bestehen beim Übergang ins Studium und im Studienverlauf weiterhin deutliche herkunftsspezifische Unterschiede.

Auf Basis der für Deutschland repräsentativen DZHW-Studienberechtigendaten (1976-2010) beschäftigt sich der Vortrag mit der Frage, wie sich die Ungleichheiten beim Zugang zu höherer Bildung und der Art der Bildungsbeteiligung in den vergangenen Jahrzehnten entwickelt haben und in welcher Weise es zu einer Veränderung der Ungleichheitsmechanismen gekommen ist.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G191

Valentina Piwowar (Freie Universität Berlin), Theresa Dicke (Universität Duisburg-Essen), Felicitas Thiel (Freie Universität Berlin), Detlev Leutner (Universität Duisburg-Essen)

## **Wissen, Können und Überzeugungen im Klassenmanagement. Empirische Zugänge zu einem zentralen Aspekt der professionellen Kompetenz von Lehrkräften**

Chair(s): Valentina Piwowar (Freie Universität Berlin), Theresa Dicke (Universität Duisburg-Essen)

Diskutant/in: Johannes Mayr (Alpen-Adria-Universität Klagenfurt)

Klassenmanagement gilt als zentraler Prädiktor für schulisches und sozio-emotionales Lernen von Schülern (Wang, Haertel, & Walberg, 1993) und wird als Kernbereich der Unterrichtsqualität konzipiert (Ophardt & Thiel, 2008). Im Symposium wird hingegen untersucht, wie verschiedene Aspekte der professionellen Kompetenz im Klassenmanagement zusammenhängen. Dabei werden unterschiedliche berufliche Ausbildungsstadien in den Blick genommen (Studierende, Referendare, berufstätige Lehrkräfte). Studien haben wiederholt gezeigt, dass Klassenmanagement mit Selbstwirksamkeitserwartungen (Brouwers & Tomic, 2000) und dem beruflichen Wohlbefinden von Lehrkräften (Friedman, 2006) positiv assoziiert ist. Darüber hinaus identifizieren Lehrkräfte im Klassenmanagement den größten Entwicklungsbedarf, insbesondere am Beginn ihrer beruflichen Karriere (Pigge & Marso, 1997).

Klassenmanagement als Teil der professionellen Kompetenz von Lehrkräften zu konzeptualisieren impliziert die theoretische Annahme, dass Lehrerhandeln multipel determiniert, also als ein Zusammenspiel aus Wissen und Können von Lehrkräften, aber auch berufsbezogenen Überzeugungen ist (in Bezug auf die Rolle als Lehrkraft und das Vertrauen in das eigene Können; Baumert & Kunter, 2011; Blömeke, Felbrich, & Müller, 2008).

Im ersten Beitrag stellen Piwowar, Ophardt und Thiel Instrumente zur Erfassung von Kompetenzen im Klassenmanagement vor: einen Fragebogen für Schüler und einen Fragebogen für Beobachter, die das Können von Lehrkräften erfassen. In Validierungsanalysen werden Zusammenhänge zur Lehrerselbstwirksamkeit und der beruflichen Belastung von Lehrkräften und Referendaren untersucht.

Im zweiten Beitrag analysieren Gold, Hellermann, Holodynski und Steffensky den Zusammenhang von Wissen über und professioneller Wahrnehmung von Klassenmanagement bei Studierenden mit Hilfe von video- und textvignettenbasierten Testinstrumenten. Darüber hinaus wird überprüft, ob die professionelle Wahrnehmung domänenspezifisch ist und sich in zwei distinkte Konstrukte unterscheiden lässt (Lernunterstützung vs. Klassenmanagement) oder eindimensional zu konzipieren ist.

Im dritten Beitrag stellen Seiz und Kunter ein Instrument zur Erfassung von zwei unterrichtsbezogenen Rollenüberzeugungen dar (instructor vs. socialiser) und untersuchen differenzielle Zusammenhänge mit zwei Verhaltensdimensionen von Klassenmanagement (kontrollierendes vs. fürsorgliches Verhalten), die durch Befragungen von Grundschullehrkräften und ihren Schülern erfasst wurden.

Im vierten Beitrag überprüfen Dicke, Parker, Marsh, Kunter, Schmeck und Leutner, welchen Einfluss Selbstwirksamkeitsüberzeugungen im Klassenmanagement zu Beginn des Referendariats vermittelt über das Ausmaß an Unterrichtsstörungen auf die emotionale Erschöpfung ein Jahr nach Beginn des Referendariats haben. Darüber hinaus wird untersucht, ob die Stärke dieser Mediation wiederum von der individuellen Höhe der Selbstwirksamkeit im Klassenmanagement abhängt. Der sich daraus ergebende konditionale indirekte Effekt wird mit Hilfe eines moderierten Mediationsmodells überprüft.

In der abschließenden Diskussion werden die Beiträge durch einen in diesem Forschungsfeld in besonderem Maße ausgewiesenen Forscher integrierend interpretiert, der die Konzeptualisierung und Erfassung von Klassenmanagement im deutschsprachigen Raum maßgeblich geprägt hat.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.

Blömeke, S., Felbrich, A. & Müller, C. (2008b). Theoretischer Rahmen und Untersuchungsdesign. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare: Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* (S. 15–48). Münster: Waxmann.

Brouwers, A. & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, 16, 239-253.

Friedman, I. A. (2006). Classroom management and teacher stress and burnout. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management* (pp. 925–945). Mahwah, NJ: Lawrence Erlbaum Associates.

Ophardt, D. & Thiel, F. (2008). Klassenmanagement als Basisdimension der Unterrichtsqualität. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion. Inhaltsfelder, Forschungsperspektiven und methodische Zugänge* (2. Aufl., S. 259–284). Wiesbaden: Verlag für Sozialwissenschaften.

Pigge, F. L. & Marso, R. N. (1997). A seven year longitudinal multi-factor assessment of teaching concerns development through preparation and early years of teaching. *Teaching and Teacher Education*, 13, 225–235.

Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of educational research*, 63, 249–294.



## **Multidimensionale und multiperspektivistische Erfassung von Klassenmanagement. Entwicklung und Validierung eines Schüler- und eines Beobachterfragebogens für die Sekundarstufe.**

Klassenmanagement gilt als ein Kernbereich professioneller Kompetenz von Lehrkräften (Brophy & Good, 1986; Kunter et al., 2005). Zentraler Bestandteil der Entwicklung und empirischen Erforschung von Lehrerkompetenzen ist deren systematische und theoriegeleitete Erfassung auf Basis objektiver, reliabler und valider Instrumente. Psychometrisch zufriedenstellende Instrumente, die eine differenzierte Erfassung von Kompetenzen im Klassenmanagement ermöglichen, sind im deutschen Sprachraum bislang jedoch kaum etabliert: Die Operationalisierung erfolgt in der Regel eindimensional im Sinne von Zeitnutzung und Störungsfreiheit oder beschränkt sich auf die von KoUniversitatn (1970) beschriebenen Strategien erfolgreichen Klassenmanagements. In Abgrenzung zu dieser eindimensionalen Operationalisierung wird Klassenmanagement international (vgl. Evertson & Weinstein, 2006) sowie national (z. B. Mayr, 2006; Ophardt & Thiel, 2013 oder Schönbachler, 2008) zunehmend breiter konzipiert.

### Ziele

Ziel der vorliegenden Arbeit war daher die Entwicklung und Überprüfung eines deutschsprachigen Instrumentes zur multidimensionalen und multiperspektivistischen Beschreibung von Kompetenzen des Klassenmanagements (KODEK) in Form von standardisierten Fragebögen für Beobachter (KODEK-B) und Schüler (KODEK-S). Im Beitrag werden Reliabilitätsmaße sowie Validierungsanalysen der entwickelten Instrumente dargestellt.

### Methode

Die Instrumente wurden unter Rückgriff auf bereits existierende Instrumente sowie auf Basis empirischer Studien theoriegeleitet entwickelt und in mehreren Pilotierungsphasen optimiert. Die Items wurden nicht gespiegelt, sondern perspektivenspezifisch formuliert. Die Endformen der Instrumente wurden an einer Gelegenheitsstichprobe von 88 Sekundarschullehrkräften implementiert. Die Berufserfahrung reicht von Berufsanfängern nach dem ersten Jahr des Referendariats bis zu 39 Jahren ( $M = 8.0$ ;  $SD = 10.56$ ) bei einem Alter von 22 bis 62 Jahren ( $M = 36.73$ ;  $SD = 10.14$ ), 67% der Lehrkräfte sind weiblich, 41 Lehrkräfte unterrichten an Gymnasien. Die Unterrichtsevaluation fand in allen Fächern mit Ausnahme von Musik, Sport und Kunst statt. An den Schülerbefragungen nahmen 1811 Schüler teil. Zur Validierung wurden konvergente (Störverhalten und Schülerbeteiligung, erfasst durch das Münchener Aufmerksamkeitsinventar, Helmke & Renkl, 1992) sowie diskriminante Kriterien (Erfassung von Lehrerselbstwirksamkeit, Schmitz & Schwarzer, 2000 und beruflicher Belastung, Enzmann & Kleiber, 1989 im Selbstbericht) erhoben; es wurden latente Korrelationen berechnet. Darüber hinaus wurde die mehrdimensionale Struktur des Schülerfragebogens mit Hilfe einer explorativen Mehrebenen-Faktorenanalyse überprüft.

### Ergebnisse

Die Instrumente weisen zufriedenstellende bis sehr gute Reliabilitätsmaße auf, die sowohl hinreichende Übereinstimmung innerhalb von Klassen, als auch hinreichend Variabilität zwischen den Klassen indizieren ( $MD_{\alpha}$ , KODEK-S = .76,  $MD_{ICC(1)}$ , KODEK-S = .40,  $MD_{ICC(2)}$ , KODEK-S = .88;  $MD_{\alpha}$ , KODEK-B = .81,  $MD_{ICC(1)}$ , KODEK-B = .88,  $MD_{ICC(2)}$ , KODEK-B = .92). Es liegen mittlere bis hohe Übereinstimmungen in der Beurteilung von Klassenmanagement zwischen Schülern und Beobachtern vor ( $MD = .55$ ,  $.31 < r < .78$ ) sowie zu den konvergenten Kriterien Schülerbeteiligung ( $MD_{KODEK-S} = .46$ ,  $MD_{KODEK-B} = .28$ ) und -aufmerksamkeit ( $MD_{KODEK-S} = .39$ ,  $MD_{KODEK-B} = .54$ ). Erwartungsgemäß zeigen sich maximal mittlere positive Zusammenhänge mit der Lehrerselbstwirksamkeit ( $MD_{KODEK-S} = .28$ ,  $MD_{KODEK-B} = .20$ ). Erwartungswidrige Zusammenhänge zeigen sich in der Subgruppe der Referendare: Während die berufliche Belastung bei Lehrkräften im Beruf negativ mit Klassenmanagement assoziiert sind ( $MD_{KODEK-S} = -.17$ ,  $MD_{KODEK-B} = -.10$ ), ist dieser Zusammenhang bei Referendaren positiv: Mehr Kompetenzen hängen mit mehr Belastungserleben zusammen ( $MD_{KODEK-S} = .16$ ,  $MD_{KODEK-B} = .18$ ), auch wenn es sich hierbei um kleine Effekte handelt. In den explorativen Faktorenanalysen konnten fünf von neun theoretisch abgeleiteten Klassenmanagementstrategien im individuellen Schülerurteil extrahiert werden; die übrigen vier Strategien fielen zu zwei Faktoren zusammen. In der geteilten Wahrnehmung wurden drei übergeordnete Dimensionen differenziert.

### Ausblick

Die gefundenen Ergebnisse werden integrierend diskutiert. In weiterführenden Studien sollten die Sensitivität und prädiktive Validität der Instrumente überprüft sowie die zum Teil erwartungswidrigen Zusammenhänge von Klassenmanagement und beruflicher Belastung genauer untersucht werden.

Bernadette Gold (Westfälische Wilhelms-Universität Münster), Christina Hellermann (Westfälische Wilhelms-Universität Münster), Manfred Holodynski (Westfälische Wilhelms-Universität Münster), Mirjam Steffensky (IPN Kiel)

## ***Professionelle Wahrnehmung von Klassenführung: Zusammenhänge mit Wissen über Klassenführung und der professionellen Wahrnehmung von Lernunterstützung***

### Theoretischer Hintergrund

Klassenführung ist ein Qualitätsmerkmal von Unterricht, für das konsistent durch Metaanalysen (Hattie, 2009; Wang, Haertel & Walberg, 1993), große Vergleichsstudien (Klieme, Schümer & Knoll, 2001; Kunter et al., 2011) sowie zahlreiche Einzelstudien positive Effekte hinsichtlich von Schülermerkmalen (Helmke & Weinert, 1997; Kunter, Baumert & Köller, 2007; Rakoczy et al., 2007), aber auch Lehrergesundheit und Arbeitszufriedenheit (Klassen & Chiu, 2010; Klusmann, Kunter, Voss, & Baumert, 2012) belegt worden sind. Dabei wird trotz vieler – teilweise uneinheitlicher – Definitionen von Klassenführung die Herstellung und Aufrechterhaltung einer positiven und lernförderlichen Lernumgebung als ihre Funktion formuliert (Brophy, 1999; Emmer & Stough, 2001). Eine wesentliche Voraussetzung für lernwirksamen Unterricht und damit auch für effektive Klassenführung ist die Fähigkeit, lernrelevante Situationen im Unterricht erkennen und theoriebasiert analysieren zu können (Seidel, Blomberg & Stürmer, 2010; Sherin, 2007). Solch eine professionelle Wahrnehmung (PW) ist domänenspezifisch und kann sich auf fachliche, fachdidaktische oder pädagogisch-psychologische Aspekte von Unterricht beziehen. Es ist unklar, wie die PW eines pädagogischen Bereichs (Klassenführung) mit der PW eines fachdidaktischen Bereichs (Lernunterstützung) zusammenhängt und ob hier eher von einem eindimensionalen Konstrukt der „professionellen Wahrnehmung von Unterricht“ im Gegensatz zu einer domänenspezifischen Fähigkeit, die von dem domänenspezifischen Professionswissen abhängt, auszugehen ist.

Erste Studien konnten einen positiven Zusammenhang zwischen PW und Schülerleistung aufzeigen (Kersting, Givvin, Thompson, Santagata, & Stigler, 2012; Roth et al., 2011). Auch Novizen und Experten unterscheiden sich deutlich in ihrer PW (Hogan, Rabinowitz, & Craven, 2003). Sie gilt als Vermittler zwischen Wissen und Handeln (Schwindt, 2008); dabei wird professionelles Wissen als die Grundlage von PW von Unterricht angenommen. Manche Arbeitsgruppen ordnen diese Fähigkeit eindeutig dem professionellen Wissen zu und gehen nicht von einer Unterscheidung der beiden Konstrukte aus (Kersting, 2008). Es stellt sich die Frage, wie die PW von Klassenführung mit Klassenführungswissen zusammenhängt bzw. ob es zwei empirisch unterscheidbare Konstrukte darstellen.

### Ziel der vorliegenden Studie

Das Ziel der Studie ist es, erste Ergebnisse über Zusammenhänge zwischen unterschiedlichen Dimensionen der PW sowie zwischen PW von Klassenführung und Wissen über Klassenführung zu erlangen. Dabei wird die Hypothese aufgestellt, dass PW inhaltsbezogen ist und die beiden Dimensionen der PW (Klassenführung und Lernunterstützung) empirisch trennbar sind. Des Weiteren wird - unter der Annahme der Inhaltsbezogenheit von PW und der Notwendigkeit von Wissen für die PW - davon ausgegangen, dass das Wissen über Klassenführung und die PW von Klassenführung hoch zusammenhängen, aber dennoch diskriminant sind.

### Methode

Zur Untersuchung der (latenten) Zusammenhänge bearbeiteten 193 Bachelorstudierende jeweils zwei videobasierte Instrumente zur Erfassung der PW von Klassenführung bzw. der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht, die im Projekt „ViU: Early Science“ entwickelt und validiert wurden, sowie einen Situational Judgment Test zum Wissen über Klassenführung (SJT-KF). Die videobasierten Testinstrumente bestehen aus kurzen Videoclips aus dem naturwissenschaftlichen Grundschulunterricht und Aussagen zum gezeigten Unterricht, die auf einer 4-stufigen Skala auf ihre Korrektheit eingeschätzt werden sollen. Der SJT-KF beschreibt klassenführungsbezogene Problemsituationen aus dem Grundschulunterricht und schlägt unterschiedliche Handlungsalternativen vor, die mit Hinblick auf ihre Effektivität bewertet werden sollen.

### Ergebnisse

Der Zusammenhang zwischen der PW von Klassenführung und der PW von Lernunterstützung fiel moderat aus ( $r = .37$ ). Dies weist damit auf zwei distinkte Konstrukte und auf die Domänenspezifität von PW hin. Sie liegt in einem ähnlichen Bereich wie der Zusammenhang zwischen pädagogischem und fachdidaktischem Wissen (Voss, Kunter & Baumert, 2011). Die latente Korrelation zwischen PW von Klassenführung und Wissen über Klassenführung fiel hoch aus ( $r = .62$ ) und unterstützt damit die Annahme, dass die PW von Klassenführung nicht mit dem Wissen darüber gleichzusetzen ist. Diese Studie gibt erste Hinweise auf den Zusammenhang beider Fähigkeiten. Der Frage, ob das professionelle Wissen eine Grundlage für eine theoriebasierte Unterrichtswahrnehmung ist, sollte weiterhin in Langzeitstudien nachgegangen werden.

## **Klassenmanagement – wie sehen Lehrkräfte ihre Rolle?**

Klassenmanagement stellt eine zentrale Dimension der Unterrichtsqualität dar (Kunter & Voss, 2011; Lipowsky et al., 2009) und ist für sowohl den Lernerfolg von Schülerinnen und Schülern (z.B. Wang, Haertel, & Walberg, 1993) als auch für ihre motivationale Entwicklung (Kunter, Baumert, & Köller, 2007) bedeutsam. Studien weisen darauf hin, dass es Unterschiede im Klassenmanagement zwischen Lehrkräften gibt (z.B. Fauth, Decristan, Rieser, Klieme, & Büttner, 2014). Eine mögliche Ursache für diese Unterschiede könnte in variierenden Rollenüberzeugungen von Lehrkräften in Bezug auf ihr Klassenmanagement liegen (Brophy, 1996). Es wird angenommen, dass diese Überzeugungen handlungsleitend sind und sich je nach Ausprägung förderlich oder hinderlich auf das Klassenmanagement auswirken können. Es existieren verschiedene Konzeptualisierungen von Überzeugungen über Klassenmanagement (siehe Woolfolk Hoy & Weinstein, 2006). Ein Vorschlag stammt von Brophy (1996), er unterscheidet in einer qualitativen Studie dichotome Rollenüberzeugungen und benennt diese als instructor vs. socialiser. Instructor sind Lehrkräfte, die ihre Rolle in Bezug auf das Klassenmanagement als die einer Autoritätsperson verstehen und verstärkt Wert auf eine funktionierende Lernumgebung legen. Socialiser hingegen betonen die Relevanz der Lehrer-Schüler-Beziehung und sehen Fürsorge und Geduld als wichtige Eigenschaften der Lehrerrolle (Brophy, 1996).

### Fragestellung

Ausgangspunkt dieser explorativen Untersuchung war das Anliegen, die Rollenüberzeugungen instructor vs. socialiser quantitativ zu erfassen. Für die Überprüfung dieses Anliegens wurden Skalen entwickelt. Um der Frage nachzugehen, ob diese Rollenüberzeugungen differenziell mit dem Klassenmanagement der Lehrkräfte zusammen hängen, wurden zusätzlich Skalen entwickelt, die verschiedene Verhaltensdimensionen der Lehrkräfte hinsichtlich ihres Klassenmanagements abfragen. Diese Skalen wurden in Anlehnung an Nie und Lau (2009) entwickelt, die beschreiben, dass einem umfassenderen Verständnis von Klassenmanagement nach Lehrkräfte neben strukturierendem, kontrollierendem Verhalten (control) ebenso fürsorgliches Verhalten (care) zeigen. Es wird ein differenzieller Zusammenhang zwischen den Überzeugungsskalen und den Verhaltensdimensionen erwartet: Die Überzeugungsskala socialiser hängt enger mit der Verhaltensdimension care zusammen während die Überzeugungsskala instructor eher mit control- Verhalten zusammen hängt. Diese differenziellen Zusammenhänge sollen als erster Validierungshinweis der Überzeugungsskalen dienen. Zusätzlich soll analysiert werden, ob sich diese differenziellen Zusammenhänge ebenfalls aus der Perspektive von Schülerinnen und Schülern zeigen.

### Methode

In einer Online-Befragung wurden 244 Grundschullehrkräften die beschriebenen Skalen vorgelegt. (Beispielitems: Skala instructor: „Ich sehe es als meine Aufgabe an, schnell auf Unterrichtsstörungen zu reagieren.“; Skala socialiser: „Mir ist es wichtig, eine gute Beziehung zu meinen Schüler(inne)n zu haben.“) Zum jetzigen Zeitpunkt liegen Daten von 224 Lehrkräften vor. Da die Erhebung noch nicht abgeschlossen ist, wird bis zur Konferenz mit einem Stichprobenumfang von ca. 300 Lehrkräften gerechnet. Für die Skalenanalyse wurden Reliabilitätsanalysen sowie Faktorenanalysen durchgeführt. Die Zusammenhänge mit den Verhaltensdimensionen wurden anhand manifester Pfadmodelle in Mplus analysiert. Bis zur GEBF-Tagung werden die Analysen mit Strukturgleichungsmodellen analysiert. Für die Analyse der Schülerperspektive ist eine kleine Zusatzbefragung im Sekundarbereich 1 geplant, die Daten liegen zum jetzigen Zeitpunkt jedoch noch nicht vor.

### Ergebnisse

Die Überzeugungsskalen wiesen ausreichende Reliabilitäten auf (socialiser  $\alpha = .70$ ; 5 Items; instructor  $\alpha = .67$ ; 8 Items). Eine gemeinsame faktoranalytische Untersuchung wies auf eine Zwei-Faktorenstruktur hin.

Die Pfadanalyse zeigte hypothesenkonform, dass die Skala socialiser signifikant höher mit der Verhaltensdimension care zusammenhängt als mit der Verhaltensdimension control ( $\beta_{care} = .37$ ;  $\beta_{control} = .16$ ). Der Zusammenhang der Skala instructor ist signifikant höher mit control ( $\beta_{control} = .41$ ) als mit der Verhaltensdimension care ( $\beta_{care} = .15$ ).

Die Überzeugungsskalen korrelieren miteinander mit  $r = .29$ , was darauf hinweist, dass sich die Rollenüberzeugungen nicht grundsätzlich ausschließen. Die Verhaltensdimensionen korrelieren ebenfalls ( $r = .23$ ). Dieser Zusammenhang unterstützt die Annahme, dass beide Verhaltensweisen komplementär sind und gemeinsam zu einem effizienten Klassenmanagement beitragen (Nie & Lau, 2009).

### Ausblick

Die Analysen bieten einen explorativen Ansatzpunkt, um Unterschiede im Klassenmanagement von Lehrkräften durch eine differenzierte Analyse ihrer Überzeugungen interpretieren zu können. Jedoch sind weitere Studien und Validierungshinweise für die Überzeugungsskalen notwendig.

Theresa Dicke (Universität Duisburg-Essen), Philip D. Parker (University of Western Sydney), Herbert W. Marsh (University of Western Sydney), Mareike Kunter (Goethe-Universität Frankfurt), Detlev Leutner (Universität Duisburg-Essen)

## ***Selbstwirksamkeitsüberzeugungen in Bezug auf Klassenmanagement, Unterrichtsstörungen und Emotionale Erschöpfung von Lehramtsanwärter(inne)n***

### Theoretischer Hintergrund

In der Forschung zu Lehrerstress zeigt sich, dass Unterrichtsstörungen die stärkste Einflussgröße beruflicher Beanspruchung unerfahrener Lehrkräfte sind. Zudem werden sowohl Klassenführung, um mit Unterrichtsstörungen adäquat umzugehen, als auch der Umgang mit berufsspezifischer Belastung (die aus der Beanspruchung resultieren kann) als wichtige Aspekte der professionellen Kompetenz von Lehrkräften gesehen (Klusmann et al., 2008). Insbesondere die Überzeugung der eigenen Wirksamkeit in Bezug auf Klassenmanagement repräsentiert eine zentrale Ressource, welche sowohl die Wahrnehmung von Belastung als auch die eingesetzten Strategien zum Umgang mit dieser Belastung beeinflusst (Parker et al., 2012). Die Forschung zum Zusammenspiel von Selbstwirksamkeitsüberzeugungen in Bezug auf Klassenführung (SK) und der erlebten Beanspruchung im Sinne von emotionaler Erschöpfung ist inkonsistent (z.B. Brouwers & Tomic, 1999). Es zeigen sich Hinweise für die Rolle von SK als Prädiktor von emotionaler Erschöpfung, wobei dieser Zusammenhang durch Stressoren mediiert wird (Bandura, 1997; Schwarzer & Hallum, 2008), sowie als Moderator des Zusammenhangs von erlebten Stressoren und der empfundenen Beanspruchung (Bakker & Demerouti, 2007; Jex & Bliese, 1999). Ziel dieser Studie ist es, diese beiden Rollen von SK zusammenzubringen und in einem empirischen Modell zu überprüfen. SK fungiert in diesem Modell somit als Prädiktor und Moderator gleichzeitig. Es ergibt sich eine moderierte Mediation, in der SK emotionale Erschöpfung via Unterrichtsstörungen vorhersagt (Prädiktor). Dabei hängt die Stärke dieser Mediation wiederum von der individuellen Höhe der SK ab (Moderator); somit ergibt sich ein konditionaler indirekter Effekt.

### Methode

#### Stichprobe

Die Stichprobe (N = 1,227) besteht aus Lehramtsanwärter(inne)n in NRW (29.7% Männer) mit einem Durchschnittsalter von M = 29.55 Jahren (SD = 4.47). Die durchschnittliche Anzahl der bisherigen Halbjahre im Referendariat liegt bei M = 1.18 (SD = 0.577). Die Lehramtsanwärter(innen) verteilen sich zu 23.7% auf die Grundschule, zu 32.3% auf Haupt-, Real- und Gesamtschule und zu 44.0% auf das Gymnasium.

#### Instrumente

- Emotionale Erschöpfung (Enzmann & Kleiber, 1989),
- Selbstwirksamkeitsüberzeugung in Bezug auf Klassenmanagement (Teachers' Sense of Efficacy Scale; Tschannen-Moran & Woolfolk Hoy, 2001),
- Wahrgenommenes Ausmaß an Unterrichtsstörungen (Baumert et al., 2008).

#### Analyseverfahren

Die moderierte Mediationsanalyse zur Berechnung des konditionalen indirekten Effekts wurde mit dem MODMED Verfahren von Preacher et al. (2007) durchgeführt, wobei dies in der vorliegenden Studie mit latenten Variablen gelungen ist. Um die Validität der gewonnenen Ergebnisse zu sichern, wurde das Modell unter Einbeziehung von Kovariaten in einem Mehrgruppen-Design und mit Hilfe von längsschnittlichen Daten überprüft (Messzeitpunkt 1 zu Beginn des Referendariats, Messzeitpunkt 2 ca. ein Jahr später).

#### Ergebnisse

Zunächst wurde die angenommene Mediation überprüft, in der SK die emotionale Erschöpfung via Unterrichtsstörungen vorhersagt. Das Modell zeigte eine gute Passung, wobei mehr SK zu weniger Unterrichtsstörungen führt, welche dann wiederum weniger emotionale Erschöpfung bedeuten.

Im Anschluss wurde die moderierte Mediation überprüft, in welcher die Stärke der Mediation vom Niveau der SK abhängt. Auch dieses Modell ergibt eine gute Passung. Es zeigte sich, dass die Mediation nur für jene Lehramtsanwärter(innen) signifikant ist, welche ein niedriges Niveau an SK angegeben hatten. Das Modell passte auch nach Aufnahme von Kovariaten (z.B. Erfahrung). In einer Mehrgruppenanalyse zeigten sich keine Unterschiede für Lehramtsanwärter(innen) verschiedener Schultypen.

Zuletzt wurde ein längsschnittlich-sequentielles Modell spezifiziert. In diesem Modell sagt SK zum Messzeitpunkt 1 die spätere emotionale Erschöpfung via Unterrichtsstörungen (beide Messzeitpunkt 2) vorher. Diese Mediation aber ist vom Ausmaß der SK zum Messzeitpunkt 1 abhängig. Es zeigten sich ähnliche Ergebnisse wie im querschnittlichen Modell mit einem signifikanten konditionalen indirekten Effekt.

#### Fazit

Die Ergebnisse der vorliegenden Studie weisen auf eine noch bedeutsamere Rolle von SK bei der Entstehung von beruflicher Beanspruchung von Lehramtsanwärter(innen) hin als bisher angenommen. Zudem zeigt sich, dass es wichtig ist, individuelle Unterschiede bei der Entstehung von Beanspruchung zu berücksichtigen, was für die Ableitung von Interventionsmaßnahmen von entscheidender Bedeutung ist.

## Entwicklung und Validierung wissenschaftspropädeutischer Messinstrumente

Chair(s): Rainer Bromme (Westfälische Wilhelms-Universität Münster)

Diskutant/in: Maria Opfermann (Universität Duisburg-Essen)

Entscheidungen in vielen Fragen des privaten und öffentlichen Lebens (z.B. Meinungsbildung zu Fragen des Klimaschutzes) erfordern zunehmend ein Verständnis von und die Auseinandersetzung mit wissenschaftlichen Erkenntnissen. Die Schule soll auf einen angemessenen Umgang mit wissenschaftlichen Inhalten und Aussagen vorbereiten und Schüler über den schulischen Kontext hinausgehend befähigen, kompetent wissenschaftliche Informationen auszuwählen und zu beurteilen. Insbesondere die Oberstufe soll zudem die Studierfähigkeit befördern und auf ein Studium vorbereiten. Dies setzt die Entwicklung eines adäquaten Verständnisses von Wissenschaft sowie von der Natur und Genese wissenschaftlichen Wissens voraus. Im Symposium werden Instrumente vorgestellt, die Aussagen über das Wissenschaftsverständnis von jungen Erwachsenen (Oberstufenschüler/innen, Studierenden) – bezogen auf naturwissenschaftliche wie auch gesellschaftswissenschaftliche Disziplinen – ermöglichen. Hierzu werden sowohl Ansätze zur Erfassung wissenschaftspropädeutischer Kompetenzen als auch zur Erfassung epistemischer Überzeugungen präsentiert und diskutiert.

Kerstin Oschatz, Jochen Kramer u.a. gehen in ihrem Beitrag zur Entwicklung eines Testinstrumentes zur Wissenschaftspropädeutik von einem mehrdimensionalen Konstrukt aus, das metawissenschaftliches Wissen, Methodenbewusstsein und metawissenschaftliche Reflexion umfasst (Müsche, 2009). Mit ihrem Instrument erfassen sie Aspekte metawissenschaftlicher Reflexion, die ein Verständnis und kritisches Hinterfragen wissenschaftlicher Forschungswege und Erkenntnisse erfordern. Christiane Bertram, Wolfgang Wagner und Ulrich Trautwein stellen ein Testinstrument zur Messung der epistemologischen Überzeugungen im Fach Geschichte vor, welches auf dem historischen Kompetenz-Strukturmodell der FUER-Gruppe (Körper et al., 2007) basiert. Im Anschluss an Hofer und Pintrich (1997) werden die epistemologischen Überzeugungen zur Natur des Wissens und des Lernens auf die für die historische Erkenntnis konstitutiven Prinzipien der Geschichtswissenschaft bezogen: Partialität, (Retro-) Perspektivität, Selektivität, Konstruktivität und Narrativität (Baumgartner, 1997; Schreiber, 2007). Monika Waldis und Martin Nitsche rekurrieren in ihrem Beitrag zu den epistemologischen Beliefs angehender Lehrpersonen im Fach Geschichte ebenfalls auf die wissenschaftstheoretischen Grundlagen des Fachs Geschichte (Jeismann, 1980; Rösen, 1997; Körper, et al., 2007) und gehen der Frage nach, ob epistemologische Überzeugungen von Lehrerstudierenden fachwissenschaftlich gesättigt sind bzw. weitere Dimensionen enthalten. Die Bedeutung des Wissenschaftsverständnisses von Laien für die Beantwortung alltäglicher Fragen steht auch im Zentrum des Beitrags von Eva Thomm und Rainer Bromme, die einen Fragebogen vorstellen, der von kontextabhängigen Annahmen über die Ursachen von Widersprüchen zwischen Wissenschaftler/innen ausgeht und die quantitative Erfassung subjektiver Annahmen von Laien über die Ursachen wissenschaftlicher Konflikte fokussiert.

Im Zentrum des Symposiums steht die Fragestellung, wie reliabel und valide mit quantitativen Testinstrumenten das Wissenschaftsverständnis, die wissenschaftspropädeutischen Kompetenzen wie auch die epistemologischen Überzeugungen von Lernenden und Laien in den verschiedenen Disziplinen mit quantitativen Testinstrumenten gemessen werden können.

Baumgartner, H. M. (1997). Narrativität. In K. Bergmann (Hrsg.) Handbuch der Geschichtsdidaktik (S.157–160). Seelze-Velber: Kallmeyer.

Becker, A. & Heuer, C. (2012). Erkenntnistheoretische Grundlage historischen Lehrens und Lernens. In M. Barricelli & M. Lücke (Hrsg.): Handbuch Praxis des Geschichtsunterrichts, (Bd.1; S. 77–88). Schwalbach/Ts.: Wochenschau-Verlag.

Hofer, B. K. & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88–140.

Jeismann, K.-H. (1980). "Geschichtsbewusstsein". Überlegungen zur zentralen Kategorie eines neuen Ansatzes der Geschichtsdidaktik. In H. Süßmuth, *Geschichtsdidaktische Positionen* (S. 179–222). Paderborn: Schöningh-Verlag.

Körper, A., Schreiber, W., & Schöner, A. (Hrsg.) (2007). Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Neuried: ars una Verlagsgesellschaft.

Müsche, H. (2009). Wissenschaftspropädeutik aus psychologischer Perspektive – Zur Dimensionierung und Konkretisierung eines bildungstheoretischen Konzeptes. *TriOS*, 4, 61–109.

Rösen, J. (1997). Historisches Erzählen. In K. Bergmann (Hrsg.), *Handbuch der Geschichtsdidaktik* (S. 51–63). Seelze-Velber: Kallmeyer.

Schreiber, W., Körper, A. von Borries, B., Krammer, R., Leutner-Ramme, S., Mebus, S, Ziegler, B. (2007). Historisches Denken. Ein Kompetenz-Strukturmodell (Basisbeitrag). In A. Körper, W. Schreiber, & A. Schöner (Hrsg.), *Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik* (S. 17–53). Neuried: ars una Verlagsgesellschaft.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G165





Kerstin Oschatz (Eberhard Karls Universität Tübingen), Jochen Kramer (Eberhard Karls Universität Tübingen),  
Eva Thomm (Westfälische Wilhelms-Universität Münster), Rainer Bromme (Westfälische Wilhelms-Universität Münster)

## **Entwicklung eines Testinstrumentes zur Erfassung von Wissenschaftspropädeutik in der gymnasialen Oberstufe (TEWI)**

Ein Ziel der gymnasialen Oberstufe ist es, wissenschaftliches Verständnis, kritisches Denken und angemessenes Umgehen mit wissenschaftlichen Erkenntnissen zu vermitteln. Diese Kompetenz Wissenschaftspropädeutik wird als wichtig angesehen für Studienleistung und den beruflichen Erfolg von Abiturienten und Abiturientinnen (KMK, 1995). Was genau unter Wissenschaftspropädeutik verstanden wird ist Gegenstand umfassender Diskussionen (vgl. Hahn, 2009; Huber, 1994; Trautwein & Lüdtke, 2004). Einigkeit besteht weitgehend darin, dass es sich um ein mehrdimensionales Konstrukt handelt, das beispielsweise metawissenschaftliches Wissen, Methodenbewusstsein und metawissenschaftliche Reflexion umfasst (Müsche, 2009) und sich nicht nur auf kognitive Kompetenzen sondern auch Einstellungen, Motivationen und Fertigkeiten erstrecken kann.

Vorgestellt wird ein an den Universitäten Tübingen und Münster neu entwickelter Kompetenztest zur Erfassung von Wissenschaftspropädeutik in der gymnasialen Oberstufe (TEWI). Die dabei fokussierte Kompetenz zielt auf die Ebene der metawissenschaftlichen Reflexion nach Müsche ab und beschränkt sich auf die kognitiven Aspekte von Wissenschaftspropädeutik. Sie erfordert ein Verständnis und kritisches Hinterfragen wissenschaftlicher Forschungswege und Erkenntnisse.

Im TEWI werden Kontroversen zu aktuellen naturwissenschaftlichen Themen vorgelegt. Dabei werden zwei kontroverse wissenschaftliche Positionen beschrieben (z.B. zur Wirkung eines Konditionstrainings im Leistungssport), die von jeweils zwei unterschiedlichen Wissenschaftlerinnen und Wissenschaftlern vertreten werden. Die Darstellungen der Positionen werden durch Informationen gestützt, die unterschiedliche Facetten wissenschaftlichen Wissens und Arbeitens darlegen (z.B. unterschiedliche disziplinäre Perspektiven, Methodenvielfalt). Zu jeder Kontroverse werden anschließend verschiedene Aussagen präsentiert, die daraufhin eingeschätzt werden müssen, ob sie zutreffen oder nicht (forced choice Format). Dazu ist ein Verständnis der geschilderten wissenschaftlichen Forschungswege und Erkenntnisse sowie eine angemessene Bewertung ihrer Relevanz für die kontroversen Positionen erforderlich. Der Test besteht aus fünf Kontroversen mit insgesamt 32 Items.

Im Vortrag wird zunächst die Bestimmung des Konstruktes und der Entwicklungsprozess vorgestellt. Letzterer umfasste verschiedene kognitive Vortests und eine Pilotierungsstudie an 93 Schülerinnen und Schülern der Jahrgangsstufe 11. Anhand dieser Pilotierung konnte ein Kompetenzscore gebildet werden (WLE-Schätzer), der eine Reliabilität von .71 und eine Varianz von 0.83 aufweist. Anschließend werden Ergebnisse einer Validierungsstudie berichtet, zu der 700 Studierende eingeladen werden. In der Validierungsstudie werden die Beziehungen des im TEWI erfassten Teilaspektes von Wissenschaftspropädeutik zu verwandten Konstrukten geprüft (Konstruktvalidierung): a) zur Persönlichkeitseigenschaft need for cognition (Keller, Bohner, & Erb, 2000), b) zu subjektiven Annahmen über die Ursachen konfligierender wissenschaftlicher Ergebnisse (ECSC; Thomm, Hentschke, & Bromme, 2013), c) zum Entwicklungsniveau (FREE; Krettenauer, 2005) und d) zu konnotativen Aspekten (CAEB; Stahl & Bromme, 2007) epistemologischer Überzeugungen sowie e) zur De- und Rekonstruktionsfähigkeit historischer Berichte (Bertram et al., 2013). Geprüft wird, inwiefern es sich dabei um unterscheidbare Teilbereiche von Wissenschaftspropädeutik handelt.

- Bertram, C., Wagner, W. & Trautwein, U. (2013). Chancen und Risiken von Zeitzeugenbefragungen – Entwicklung eines Messinstrumentes für eine Interventionsstudie. In J. Hodel & B. Ziegler (Hrsg.), Forschungswerkstatt Geschichtsdidaktik 12, Beiträge zur Tagung "geschichtsdidaktik empirisch 12" (S. 108–119). Bern: Hep-Verlag.
- Hahn, S. (2009). Wissenschaftspropädeutik in der Sekundarstufe II: Bildungsgeschichtlicher Rückblick und aktuelle Entwicklungen. *TriOS*, 4, 5–37.
- Huber, L. (1994). Nur allgemeine Studierfähigkeit oder doch allgemeine Bildung? Zur Wiederaufnahme der Diskussion über Hochschulreife und die Ziele der Oberstufe. *Die deutsche Schule*, 86, 12–26.
- Keller, J., Bohner, G. & Erb, H.P. (2000). Intuitive und heuristische Urteilsbildung - verschiedene Prozesse? *Zeitschrift für Sozialpsychologie*, 31, 87–101.
- KMK - Kultusministerkonferenz (1995). Weiterentwicklung der Prinzipien der gymnasialen Oberstufe und des Abiturs. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Expertenkommission. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- Krettenauer, T. (2005). Die Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen und das Problem der Übertragbarkeit von Interviewverfahren in standardisierte Fragebogenmethoden. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 69–79.
- Müsche, H. (2009). Wissenschaftspropädeutik aus psychologischer Perspektive – Zur Dimensionierung und Konkretisierung eines bildungstheoretischen Konzeptes. *TriOS*, 4, 61–109.
- Stahl, E. & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, 17, 773–785.
- Thomm, E., Hentschke, J. & Bromme, R. (2013). The ExplainingConflictingScientificClaims (ECSC) Questionnaire: Measuring laypersons' explanations for science conflicts. Zur Veröffentlichung eingereicht.
- Trautwein, U. & Lüdtke, O. (2004). Aspekte von Wissenschaftspropädeutik und Studierfähigkeit. In O. Köller, R. Watermann, U. Trautwein, & O. Lüdtke (Hrsg.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA-Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. (S. 327–366). Opladen: Leske + Budrich.

## **Entwicklung eines Testinstruments zur Messung epistemologischer Überzeugungen im Fach Geschichte**

Seit der narrativen Wende in den 80er-Jahren bestimmt die Kategorie des „Geschichtsbewusstseins“, verstanden als eine Verbindung der Gegenwart mit der Vergangenheit, die sich als historisches Erzählen beschreiben lässt (Rüsen, 1983), die Geschichtsdidaktik. Alle derzeit diskutierten Kompetenzmodelle basieren auf der geschichtstheoretischen Basis von Geschichtsbewusstsein und Narrativität. „Geschichte“ entsteht aufgrund von historischen Fragestellungen als Rekonstruktion von Vergangenheit und wird in historischen Narrationen greifbar, die Bedeutung auch für die historische Orientierung in der Gegenwart und Zukunft haben (Pandel, 2013).

Im Rahmen einer Interventionsstudie zur Wirksamkeit von Zeitzeugenberichten im Geschichtsunterricht wurde ein Kurzinstrument entwickelt, in dem die Grundlagen historischen Denkens in vierzig geschlossen formulierten Items thematisiert werden. Als theoretischer Rahmen bei der Itemgenerierung und Validierung des Tests wurde das in der Fachdidaktik Geschichte weithin anerkannte Kompetenz-Strukturmodell historischen Denkens der sogenannten FUER-Gruppe zugrunde gelegt. Von den Frage-, Methoden- und Orientierungskompetenzen werden die sogenannten „Sachkompetenzen“, d.h. Begriffs- und Strukturierungskompetenzen, abgegrenzt, die unter anderem auch das Verfügen über die epistemologischen Prinzipien im Fach Geschichte beinhalten. Nach Hofer und Pintrich (1997) versteht man unter den epistemologischen bzw. epistemischen Prinzipien oder Überzeugungen („epistemological beliefs“) die Überzeugungen einer Person über die Natur des Wissens und des Lernens. Epistemologische Prinzipien in der Geschichtswissenschaft sind in diesem Sinn „konstitutiv für historische Erkenntnis und schränken die Möglichkeit a priori ein, durch Geschichte eine vergangene ‚Wirklichkeit‘ abzubilden.“ (Schreiber, 2007, S. 32).

In dem vorzustellenden Instrument wurden solche geschichtstheoretischen Voraussetzungen, die auf die Grundlagen spezifisch historischer Erkenntnisprozesse abzielen, in geschlossenen Items formuliert. Testinstrumente, mit denen die epistemologischen Überzeugungen zu verschiedenen Wissenschaftsdisziplinen gemessen werden sollen, wie auch Aufgabenstellungen aus geschichtsdidaktischen empirischen Studien dienen als Anregung für die Formulierung der Items. Da die Interventionsstudie die Wirksamkeit von Zeitzeugenbefragungen untersucht, wurden auch Items formuliert, die explizit die Aussagekraft von Zeitzeugen in den Blick nehmen.

In dem Vortrag werden die Befunde aus zwei Erhebungen vorgestellt, die im Sommer 2011 mit 360 und im Sommer 2012 mit 554 Studierenden durchgeführt wurden. Das Kurzinstrument besteht aus drei Itemgruppen, die zusammengefasst werden unter: (a) „Woher wissen wir etwas über die Vergangenheit?“, (b) „Aussagekraft von Zeitzeugen“ und (c) „Wie gut wissen wir über die Vergangenheit Bescheid?“. In den exploratorischen Faktorenanalysen zeigte sich in den beiden Erhebungen eine sehr ähnliche Faktorenstruktur mit drei Faktoren, die folgenden Erkenntnisbereichen zugeordnet werden können: 1.) der Erkenntnis, dass Geschichte aus der Re-Konstruktion der Vergangenheit entsteht, dass also Quellen und Darstellungen analysiert, interpretiert und verglichen werden müssen, um eine Narration von der Vergangenheit zu erstellen; 2.) der Erkenntnis, dass geschichtliche Narrationen auf ihre Tiefenstruktur hin kritisch befragt, also dekonstruiert werden müssen; und 3.) der Erkenntnis, dass Zeitzeugen eine sehr spezifische Geschichte erzählen, da sie Teil der Vergangenheit sind, aber gleichzeitig mit zeitlichem Abstand berichten. Auch die Befunde der Haupterhebung mit 937 Schüler/innen der neunten Klassenstufe sollen in den Vortrag mit einfließen.

Baumgartner, H. M. (1997). Narrativität. In K. Bergmann, Handbuch der Geschichtsdidaktik (S. 157–160). Seelze-Velber: Kallmeyer.

Hofer, B. K. / Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. Review of Educational Research, 67, 88–140.

Pandel, H.-J. (2013). Geschichtsdidaktik. Eine Theorie für die Praxis. Schwalbach/Ts.: Wochenschau-Verlag.

Rüsen, J. (1983). Historische Vernunft: Die Grundlagen der Geschichtswissenschaft. Grundzüge einer Historik I. Göttingen: Vandenhoeck & Ruprecht.

Schreiber, W. et al. (2007). Historisches Denken. Ein Kompetenz-Strukturmodell. In: A. Körber, W. Schreiber & A. Schöner, Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik (S. 17–53). Neuried: ars una Verlagsgesellschaft.

## **Epistemologische Beliefs von angehenden Lehrpersonen im Fach Geschichte – Erste Ergebnisse aus dem Projekt VisuHist**

Epistemologische Überzeugungen sind subjektive Theorien über die Struktur des Wissens und den Wissenserwerb (Hofer & Pintrich, 1997). Durch eine Vorstrukturierung der wahrgenommenen Inhalte beeinflussen sie die Wahrnehmung und Informationsverarbeitung. Wissenschaftlich angemessene epistemologische Überzeugungen über die Struktur des Wissens und den Wissenserwerb in Geschichte sind als Kernbestandteil einer historischen Grundbildung sowohl von Novizen (Schüler/-innen) als auch von Experten zu betrachten. Epistemologische Beliefs sind somit Inhalt der professionellen Wissensbasis von Lehrpersonen und werden dem Teilbereich des Fachwissens (content knowledge) zugeordnet (Baumert & Kunter, 2006). Es wird angenommen, dass die Sicht darüber, was Geschichte ist und wie historisches Wissen entsteht, einen grossen Einfluss auf die Gestaltung des Geschichtsunterrichts hat (van Norden, 2012).

Die Darstellung epistemologischer Überzeugungen im Fach Geschichte wird von der Geschichtstheorie beeinflusst und ist vor allem am Erkenntnissystem der Geschichtswissenschaft orientiert (Jeismann, 1980; Rüsen, 1997; Körber et al., 2007). Damit bleibt der Blick auf historische Lernprozesse vorerst im Hintergrund. Die Struktur und Genese historischen Wissens ist in Fachkreisen vielfach diskutiert (vgl. Kühberger, 2012). Der englische Geschichtstheoretiker Keith Jenkins beschreibt salopp: „The past has gone and history is what historians make of it when they go to work“ (Jenkins, 2003, 8) und greift damit drei fundamentale Überlegungen auf: Erstens, dass Geschichte als kulturelle Sinnbildung von erkennenden Subjekten zu verstehen ist, zweitens, dass Geschichte erst im Tun, das heisst, im sprachlichen Modus des Erzählens entsteht und drittens, dass Historiker und Historikerinnen die zentralen Akteure sind. Während die erste und zweite Überlegung in Fachkreisen kaum bestritten ist, so bestehen im dritten Punkt unterschiedliche Vorstellungen. So machen beispielsweise Becker und Heuer (2012) darauf aufmerksam, dass Geschichte nicht nur die methodische Denkweise einer Wissenschaft sondern ganz generell die alltägliche Denkweise von Menschen umfasst. Unterschiedliche Positionen sind auch in der Frage nach der Objektivität in der Geschichtswissenschaft festzustellen und umspannt Fragen wie diejenige nach dem Verhältnis der Wissenschaft zu ihrem Forschungsgegenstand oder dem Verhältnis von historischer Wirklichkeit und historischer Darstellung (Rüsen, 1997; Goertz, 2009). Hingegen stellen die prinzipielle Offenheit und Pluralität historischer Erkenntnis, der Konstruktcharakter von Geschichte oder Strukturmomente wie Selektivität, Perspektivität, Retrospektivität, Medialität und Narrativität relativ breit anerkannte Aspekte von Geschichte dar (z.B. Barricelli, 2012; Baumgartner, 1997; Borries & Körber, 2008; Becker & Heuer, 2012).

Im vorliegenden Beitrag werden epistemologische Beliefs zu Geschichte aus der Sicht angehender Geschichtslehrpersonen beschrieben. Die empirischen Daten entstammen dem vom Schweizerischen Nationalfonds geförderten Projekt „VisuHist“, welches die Ausprägung und Genese professionellen Wissens von Geschichtslehrpersonen untersucht. Im Rahmen des Projekts werden rund 250 Geschichtslehrerstudierende in vier Pädagogischen Hochschulen der deutschsprachigen Schweiz sowohl zu Beginn als auch gegen Ende ihrer fachdidaktischen Ausbildung mittels Paper & Pencil-Test zum Fachwissen getestet und zu ihren epistemologischen Überzeugungen befragt. Im ersten Erhebungszeitpunkt wurden die Studierenden mittels offenem Frageformat aufgefordert, ihre Sichtweise auf die Struktur und Beschaffenheit historischen Wissens kurz und treffend zu beschreiben. Die vorliegenden Studierendentexte (n = 100) wurden im Projekt mittels der Methode der qualitativen Inhaltsanalyse (Mayring, 2007) analysiert. Analyseeinheit war der einzelne Satz bzw. Teilsatz. Erste Ergebnisse belegen die häufige Nennung von Aspekten wie Zeitdifferenz, die Unterscheidung von Vergangenheit und Geschichte, Geschichte als Konstrukt, Selektivität, Perspektivität, Retrospektivität und Partialität, Quellenbasierung und die Notwendigkeit von deren Rekonstruktion und Interpretation sowie die Vorläufigkeit historischen Wissens. Weit weniger häufig werden historische Methoden im Sinne von Heuristik, Kritik und Interpretation erwähnt oder Überlegungen zur „Objektivität“ historischen Wissens formuliert. Darüber hinaus wird Geschichte als kulturelle Praxis aufgefasst, an der nicht nur Experten (Historiker/-innen) sondern auch Novizen beteiligt sind. Die vorliegenden Daten verweisen auf eine grosse Heterogenität innerhalb der untersuchten Studierendengruppe. Die Befunde werden im Hinblick auf die Entwicklung eines Messinstruments mit geschlossenen Antwortformaten diskutiert. Zugleich werden erste Schlussfolgerungen für die Lehrerbildung gezogen.

- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9, 469–520.
- Baumgartner, H.M. (1997). Narrativität. In K. Bergmann u.a. (Hrsg.), Handbuch der Geschichtsdidaktik (S. 157–160). Seelze-Velber: Kallmeyer.
- Barricelli, M. (2012). Narrativität. In M. Barricelli & M. Lücke (Hrsg.), Handbuch Praxis des Geschichtsunterrichts, Bd. 1 (S. 255–280). Schwalbach/Ts.: Wochenschau Verlag.
- Becker, A. & Heuer, Ch. (2012). Erkenntnistheoretische Grundlage historischen Lehrens und Lernens. In M. Barricelli & M. Lücke (Hrsg.), Handbuch Praxis des Geschichtsunterrichts, Bd. 1, (S. 77–88). Schwalbach/Ts.: Wochenschau Verlag.
- Von Borries, B. & Körber, A. (2009). Historisches Denken – Zur Bestimmung allgemeiner und domänenspezifischer Kompetenzen und Standards. In M.A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), Perspektiven der Didaktik (S. 293 – 311). Wiesbaden: VS, Verlag für Sozialwissenschaften
- Goertz, H.-J. (2009). Was können wir von der Vergangenheit wissen?. Geschichte in Wissenschaft und Unterricht, 12, 692–706.
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. Review of Educational Research, 67, 88–140.
- Jeismann, K.-E. (1980). "Geschichtsbewusstsein". Überlegungen zur zentralen Kategorie eines neuen Ansatzes der Geschichtsdidaktik. In H. Süsmuth (Hrsg.), Geschichtsdidaktische Positionen. Bestandsaufnahmen und Neuorientierung (S. 179–222). Paderborn: Schöningh.
- Jenkins, K. (2003). Re-thinking History. London: Routledge.
- Kühberger, C. (Hrsg.) (2012). Historisches Wissen. Geschichtsdidaktische Erkundungen zu Art, Tiefe und Umfang für das historische Lernen. Schwalbach/Ts.: Wochenschau Verlag.
- Mayring, P. (2008). Qualitative Inhaltsanalyse: Grundlagen und Techniken. Weinheim: Beltz Verlag.
- Van Norden, J. (2012). Geschichte ist Einstellungssache. Zeitschrift für Didaktik der Gesellschaftswissenschaften, 1, 54-75.
- Pandel, H.-J. (1997). Zeit. In K. Bergmann u. a. (Hrsg.), Handbuch der Geschichtsdidaktik (S. 10–15). Seelze-Velber: Kallmeyer.
- Pandel, H.-J. (2010). Historisches Erzählen. Narrativität im Geschichtsunterricht. Schwalbach/Ts.: Wochenschau Verlag.
- Rohlfes, J. (2005). Geschichte und ihre Didaktik. Göttingen: Vandenhoeck und Ruprecht.
- Rüsen, J. (1986). Rekonstruktion der Vergangenheit. Grundzüge einer Historik II: Die Prinzipien der historischen Forschung. Göttingen: Vandenhoeck und Ruprecht.
- Rüsen, J. (1997). Historisches Erzählen. In K. Bergmann u. a. (Hrsg.), Handbuch der Geschichtsdidaktik (S. 51–63). Seelze-Velber: Kallmeyer.
- Wiersing, E. (2007). Geschichte des Historischen Denkens. Zugleich eine Einführung in die Theorie der Geschichte. Paderborn: Schöningh.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G165



## **Messung subjektiver Annahmen über die Ursachen konfligierender wissenschaftlicher Aussagen verschiedener Experten**

Wissenschaftliches Wissen entsteht in einem Diskurs, es ist stets Gegenstand von Widerspruch, und Veränderung, insoweit ist es fragil (Kajanne & Pirttilä-Backman, 1999; cf. Upshur, 2000). Andererseits aber soll es zum Verständnis der Welt (bzw. des Weltausschnittes, für den eine Disziplin zuständig ist) beitragen und vielfach auch Handlungsmöglichkeiten begründen. Wenn Laien sich nun zu wissenschaftlichen Themen informieren, weil sie solche Erklärungen oder Handlungsmöglichkeiten suchen (Feinstein, 2012) wird diese Fragilität zum Problem. Ein Beispiel: Wenn Laien bei der Recherche zu Gesundheitsfragen im Internet auf widersprüchliche Geltungsbehauptungen unterschiedlicher Wissenschaftler stoßen müssen sie diese Fragilität kognitiv und möglicherweise auch emotional bewältigen.

Für die Bewältigung von solchen Widersprüchen (was kann man nun eigentlich glauben?) ist es wichtig, sich zu erklären, wie diese überhaupt zustande kommen können. Solche subjektiven Erklärungen wurden in der Entwicklungs- wie auch in der Pädagogischen Psychologie bislang im Zusammenhang mit den Konstrukten des kritischen Denkens und der epistemologischen Überzeugungen betrachtet (King & Kitchener, 1994; Kuhn, 1991). Hierbei wird zumeist ein generelles, abstraktes Verständnis von wissenschaftlichem Wissen und Wissenschaft untersucht, weniger dagegen die Ursachenannahmen, die sich auf den konkreten Prozeß der wissenschaftlichen Wissenschaftsproduktion und auf die Wissenschaftler beziehen.

Qualitative Studien zeigen, dass Individuen differenzierte, kontextabhängige Annahmen über die Ursachen von Widersprüchen zwischen Wissenschaftlern haben (Bromme, Thomm, & Wolf, 2013; Kajanne & Pirttilä-Backman, 1999). Auf Grundlage dieser Befunde entwickelten wir ein Instrument zur quantitativen Erfassung subjektiver Annahmen von Laien über die Ursachen wissenschaftlicher Konflikte, den ExplainingConflictingScientificClaims Fragebogen (ECSC; Thomm, Hentschke, & Bromme, 2013). Ziel war es, die jeweiligen Erklärungsmuster zu beschreiben, die Laien zur Auflösung konfligierender wissenschaftlicher Aussagen heranziehen. In diesem Beitrag stellen wir die Entwicklung und Überprüfung des ECSC vor.

Der ECSC kombiniert die Präsentation eines Szenarios konkreter widersprüchliche Aussagen zweier Wissenschaftler (ähnlich wie bei King & Kitchener, 1994) mit der Beantwortung eines standardisierten Fragebogens. Der Fragebogen umfasst 23 Items, die je eine Erklärung (z.B. die Wissenschaftler wenden unterschiedliche Methoden an) für den Konflikt nennen. Die Probanden sollen dann angeben, in welchem Maße sie die einzelnen Erklärungen für das konkrete Szenario für angemessen halten. Die Szenarios behandelten dabei Themen der Medizin sowie der Klimaforschung.

Die Konstruktion des ECSC erfolgte in einer ersten Studie (N =285 Studierenden und Berufstätige). Explorative Faktorenanalysen (Hauptkomponentenanalyse, Varimax-Rotation) ergaben eine Vier-Faktoren-Lösung. Die endgültige Lösung erhielt 23 der ursprünglich 35 Items, die sich gleichmäßig auf die vier inhaltlich konsistenten Dimensionen (Cronbach's  $\alpha = .72-.92$ ) verteilen und 51 % der Gesamtvarianz aufklären. Widersprüchliche wissenschaftliche Aussagen werden demnach durch Unterschiede im Forschungsprozess, Komplexität des Themas, Verfolgen von Interessen, Kompetenzunterschiede erklärt.

Eine Testung der gefundenen Faktorenstruktur erfolgte in einer zweiten Studie (N =262 Studierende). Konfirmatorische Faktorenanalysen bestätigten die Vier-Faktoren-Struktur (CFI =.91; RMSEA =.06, range .05-.07; SRMR =.07). In den vorgelegten Szenarios wurden zudem Merkmale der sich widersprechenden Fachleute variiert; z.B. wurde zu dem konkreten Widerspruch (Wissenschaftler A sagt das Gegenteil von Wissenschaftler B) die Information hinzugefügt, ob diese in einer Universität oder in einem Privatunternehmen arbeiten. Die experimentelle Manipulation diente der Testung, inwieweit die Dimensionen des Fragebogens für den Einfluss von Quellenmerkmalen auf subjektive Konflikterklärungen sensitiv sind.

Es zeigte sich ein signifikanter Einfluss der Quellenmanipulation auf die bevorzugten Erklärungen: Die Probanden attribuierten den Widerspruch insbesondere dann auf das Verfolgen von Interessen und auf Kompetenzunterschiede, wenn einem der Fachleute ein kommerzielles Interesse zugeschrieben werden konnte. Die Ergebnisse der Studien zeigen, dass der ECSC es ermöglicht, differenzierte Annahmen von Studierenden über die Ursachen wissenschaftlicher Widersprüche reliabel zu erfassen. Abschließend diskutieren wir die Rolle subjektiver Konflikterklärungen im Kontext der Entwicklung von propädeutischem Wissenschaftsverständnis und Scientific Literacy.

Bromme, R., Thomm, E., & Wolf, V. (accepted). From understanding to deference: Laypersons' and medical students' views on conflicts within medicine. *International Journal of Science Education*, Part B: Community Education and Public Engagement.

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, 95, 168–185.

Kajanne, A., & Pirttilä-Backman, A. (1999). Laypeople's viewpoints about the reasons for expert controversy regarding food additives. *Public Understanding of Science*, 8, 303–315.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

Thomm, E., Hentschke, J. & Bromme, R. (2013). The ExplainingConflictingScientificClaims (ECSC) Questionnaire: Measuring laypersons' explanations for science conflicts. Zur Veröffentlichung eingereicht.

Upshur, R. E. G. (2000). Seven characteristics of medical evidence. *Journal of Evaluation in Clinical Practice*, 6, 93–97.

## **Computerbasiertes Assessment (CBA) im Klassenraum**

Chair(s): Frank Goldhammer (DIPF), Birgit Harks (DIPF)

Diskutant/in: Susanne Narciss (TU Dresden)

Ein wichtiger Anwendungszweck von Assessment ist lernbegleitende Diagnostik inklusive Erfassung individueller Lernverläufe und Rückmeldung der diagnostischen Informationen an Schüler(innen) und Lehrkräfte. Letzteres bildet ein Kernelement von individueller Förderung und adaptivem Unterricht. Vorteile einer computerbasierten Realisierung von Assessment bestehen unter anderem in der automatischen zeitnahen Rückmeldung ohne manuelle Dateneingabe, der Nutzung von Prozessdaten für Rückmeldung und der genaueren Individualdiagnostik. Die Beiträge des Symposiums „Computerbasiertes Assessment (CBA) im Klassenraum“ befassen sich aus psychologischer und informationstechnologischer Perspektiven mit der Frage, wie computerbasierte Diagnostik und Rückmeldung zur Unterstützung von Lernprozessen im schulischen Bereich wirksam eingesetzt werden können und welche Herausforderungen für den erfolgreichen Transfer in die alltägliche Unterrichtspraxis bewältigt werden müssen. Die Beiträge diskutieren dazu diagnostische Anforderungen, untersuchen die Wirksamkeit von Assessmentmaßnahmen und stellen sich darüber hinaus Infrastruktur- und Implementierungsfragen.

Der Beitrag von Harks et al. stellt Ergebnisse zur leistungs-, metakognitions- und motivationsförderlichen Wirkung von lösungsprozessbezogener Rückmeldung vor und leitet darauf aufbauend Gestaltungshinweise für ein computerbasiertes Rückmeldeformat ab, das um die spezifischen Möglichkeiten von CBA erweitert ist. Über ein bereits umgesetztes webbasiertes Feedbackinstrument berichtet der Beitrag von Daume und Nachtigall. Das „Schüler als Experten für Unterricht“ (SEfU)-Instrument erlaubt Lehrkräften, sich Rückmeldung über die Qualität ihres Unterrichts aus Sicht der Schüler(innen) einzuholen und stellt somit Wissen zur Unterrichtsentwicklung zur Verfügung. Es werden Ergebnisse einer Nutzerbefragung präsentiert. Der Beitrag von Förster et al. gibt einen Überblick über das quop-System, das Lehrkräften und Schüler(innen) eine internetbasierte Lernverlaufsdagnostik mit Messverfahren für Lesen und Mathematik anbietet. Vorgestellt werden Ergebnisse zur Implementierung, psychometrischen Qualität und leistungsförderlichen Wirkung des Verfahrens. Der abschließende Beitrag von Breiter widmet sich aus der Informatik-Perspektive wichtigen technisch-organisatorischen Fragen, die bei der Realisierung von CBA in Schulen zu beachten sind. Zentrale Herausforderungen beziehen sich auf Bereitstellung und Betrieb von Informationstechnologie, Informationssicherheit und Datenschutz sowie die Benutzbarkeit und Verständlichkeit der Datenvisualisierungen.

Die vier Beiträge fokussieren unterschiedliche Aspekte von CBA im Klassenraum und geben somit einen Überblick über die vielfältigen Faktoren, die für eine erfolgreiche computerbasierte Diagnostik und Rückmeldung im schulischen Kontext zusammenwirken müssen.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G131



**Formatives Assessment: Die (computerbasierte) Gestaltung der Rückmeldekomponente**

Rückmeldung ist eine zentrale Komponente formativen Assessments (Wiliam & Thompson, 2008) und einer der bedeutendsten Einflussfaktoren schulischen Lernens (Hattie, 2009). Dennoch mangelt es an empirisch fundierten Informationen zur Gestaltung von Rückmeldung für formatives Assessment (Kingston & Nash, 2011). Ziel des vorliegenden Beitrags ist es zur Gestaltung von formativer Rückmeldung im Mathematikunterricht (Sekundarstufe I) beizutragen. Hierzu werden zunächst Ergebnisse des Projekts Conditions and Consequences of Classroom Assessment (Co2CA) zur Wirkung von Papier-Bleistift-basierter Testrückmeldung vorgestellt. Hierauf aufbauend werden Gestaltungsprinzipien für computerbasierte Rückmeldung abgeleitet.

Obwohl Rückmeldung Lernen im Durchschnitt positiv beeinflusst, wirkt sie nicht immer gleichermaßen stark und nicht automatisch positiv (Bangert-Drowns et al., 1991). Bisherige Rückmeldeforschung zeigt, dass Rückmeldung insbesondere dann lernförderlich ist, wenn sie sich konkret auf Aufgabenleistung, Verstehens- bzw. Regulationsprozesse bezieht, über den individuellen Lernstand relativ zum Lernziel informiert, Strategien liefert (Hattie & Timperley, 2007), eine individuumsbezogene bzw. kriteriale Bezugsnorm verwendet (Krampen, 1987) und nach einem Misserfolg instabile Attributionen nahelegt (Dresel & Haugwitz, 2008).

**Papier-Bleistift-basierte Rückmeldung**

Theoretischer Hintergrund: Papier-Bleistift-basierte Testrückmeldung, die die oben genannten lernförderlichen Rückmeldemerkmale in Form von prozessbezogenen Stärken-Schwächen-Strategie-Profilen kombiniert, wurde im Rahmen des Co2CA-Projekts entwickelt (lösungsprozessbezogene Rückmeldung). Die Wirkung von lösungsprozessbezogener Rückmeldung auf Leistungs-, Interessens- und Selbstevaluationsänderung wurde mit der Wirkung von Rückmeldung in Form von Noten (der im Unterricht üblichen Form schriftlicher Rückmeldung) verglichen. Für ein genaueres Verständnis der Rückmeldeeffekte wurde zusätzlich die vermittelnde Rolle der wahrgenommenen Nützlichkeit von Rückmeldung untersucht.

Hypothesen: Es wurde erwartet, dass (1) lösungsprozessbezogene Rückmeldung im Vergleich zu Noten vermittelt über die wahrgenommene Nützlichkeit der Rückmeldung einen indirekten Effekt auf Leistungs-, Interessens- und Selbstevaluationsänderung hat; (2) lösungsprozessbezogene Rückmeldung im Vergleich zu Noten einen totalen Effekt auf Leistungs-, Interessens- und Selbstevaluationsänderung hat.

Methode: 146 Schüler(innen) der neunten Jahrgangsstufe (Realschule) wurden einer von zwei experimentellen Bedingungen (lösungsprozessbezogene Rückmeldung oder Noten) zugewiesen. Wahrgenommene Nützlichkeit (Cronbachs  $\alpha = .81$ ), Interesse (Cronbachs  $\alpha = .90$ ) und Selbstevaluation wurden mit Fragebögen, Mathematikleistung durch Tests (EAP/PV Reliabilität = .73) erfasst. Die statistische Auswertung erfolgte mittels Pfadanalysen in MPLus.

Ergebnisse: (1) Lösungsprozessbezogene Rückmeldung wird als nützlicher wahrgenommen als Noten,  $\beta = .60$ , 95%-KI [.28, .90]. Die wahrgenommene Nützlichkeit der Rückmeldung ist positiv mit der Änderung von Leistung,  $\beta = .20$ , 95%-KI [.04, .35], und Interesse,  $\beta = .21$ , 95%-KI [.03, .40], nicht aber mit der Änderung von Selbstevaluation,  $\beta = -.06$ , 95%-KI [-.22, .10], assoziiert. In Übereinstimmung hiermit hat lösungsprozessbezogene Rückmeldung im Vergleich zu Noten, vermittelt über ihre wahrgenommene Nützlichkeit, einen positiven indirekten Effekt auf Leistungs-,  $\beta = .12$ , 95%-KI [.04, .25], und Interessensänderung,  $\beta = .13$ , 95%-KI [.02, .31], aber nicht auf die Änderung der Selbstevaluation,  $\beta = -.04$ , 95%-KI [-.15, .05]. (2) Ohne Berücksichtigung von wahrgenommener Nützlichkeit zeigen sich keine totalen Effekte auf Leistungs-,  $\beta = .24$ , 95%-KI [-.08, .56], Interessens-,  $\beta = .15$ , 95%-KI [-.18, .49], oder Selbstevaluationsänderung,  $\beta = .04$ , 95%-KI [-.28, .36]. Leistungs- und Interessenseffekte deuten sich zwar auf deskriptiver Ebene an, werden aber nicht signifikant. Möglicherweise können entsprechende Effekte durch eine intensiviertere Rückmeldeintervention, wie sie durch den Einsatz von Computern ermöglicht wird, verstärkt werden.

Computerbasierte Rückmeldung: Aufbauend auf diesen Erkenntnissen sollen lösungsprozessbezogene Rückmeldung und entsprechende Diagnostiktools computerbasiert in den Unterricht übertragen und erweitert werden. Eine computerbasierte Umsetzung hat gegenüber Papierinstrumenten den Vorteil, dass Rückmeldung automatisiert, zeitnah und intensiviert erfolgen kann. Zudem liefern computerbasierte prozessbezogene Verhaltensdaten (wie Interaktionssequenzen während der Aufgabenlösung, z.B. Mausclicks) genauere Hinweise zum Lösungsprozess, die für die Generierung einer noch differenzierteren, lösungsprozessbezogenen Rückmeldung genutzt werden können. Neben lösungsprozessbezogener Testrückmeldung können auch strategische Hinweise nach der Bearbeitung einzelner Aufgaben gegeben werden (Koedinger et al., 2010; Narciss & Huth, 2006). Außerdem ist eine ergänzende Rückmeldung in Form von Kompetenzstufenmodellen möglich (Harks et al., 2013). Entsprechende Umsetzungsbeispiele, Implementationsansätze und Transferfragestellungen werden präsentiert und diskutiert.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A. & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61 (2), 213-238. doi:10.3102/00346543061002213

Dresel, M. & Haugwitz, M. (2008). A computer-based approach to fostering motivation and self-regulated learning. *The Journal of Experimental Education*, 77 (1), 3-18. doi:10.3200/JEXE.77.1.3-20

Harks, B., Rakoczy, K., Klieme, E., Hattie, J. & Besser, M. (im Druck). Indirekte und moderierte Effekte von Rückmeldung auf Leistung und Motivation. In A. Müller & H. Ditton (Hrsg.), *Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Waxmann: Münster.

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81-112. doi:10.3102/003465430298487

Kingston, N. & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30, 28-37. doi:10.1111/j.1745-3992.2011.00220.x

Koedinger, K. R., McLaughlin, E. & Heffernan, N. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, 43 (1), 489-510. doi:org/10.2190/EC.43.4.d

Krampen, G. (1987). Differential effects of teacher comments. *Journal of Educational Psychology*, 79, 137-146. doi:10.1037//0022-0663.79.2.137

Narciss, S. & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction*, 16 (4), 310-322.

doi:10.1016/j.learninstruc.2006.07.003

William, D. & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). New York, NY: Lawrence Erlbaum.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G131



## ***CBA im Klassenraum – Schüler als Experten für Unterricht (SEfU)***

Das webbasierte Feedbackinstrument SEfU gibt Lehrerinnen und Lehrern die Möglichkeit, sich Rückmeldung über die Qualität ihres Unterrichts von den Schülerinnen und Schülern zeitökonomisch und arbeitssparend einzuholen. SEfU wird auf freiwilliger Basis für Lehrerinnen und Lehrer in Thüringen, Nordrhein-Westfalen und Sachsen kostenlos angeboten. Neben allgemeinen Fragebögen zum Unterricht können auch eigene erstellt werden. Dazu stehen Module zur Verfügung, welche zentrale Aspekte des Unterrichts (wie z.B. „diagnostische Kompetenz“ oder „Zeitnutzung“) in einem Fragenblock erfassen. Außerdem besteht die Möglichkeit, eigene Fragen zu formulieren. Die Auswertung erfolgt automatisch und kann in Form eines Ergebnisberichts heruntergeladen werden. Die Durchführung in SEfU läuft über einen passwortgeschützten Bereich ab, auf den nur die angemeldete Lehrkraft Zugriff hat. Dadurch wird gewährleistet, dass die Datenhoheit bei der Lehrerin oder bei dem Lehrer bleibt.

In der Hattie-Studie (vgl. Hattie, 2012) wurde deutlich hervorgehoben, dass Schülerfeedback einen positiven Einfluss auf Schülerleistungen hat. SEfU (Schüler als Experten für Unterricht) unterstützt Lehrerinnen und Lehrer in diesem Feedbackprozess (vgl. Börner & Nachtigall, 2013). Entscheidend für die Unterrichtsentwicklung ist allerdings nicht nur das Instrument an sich, sondern welche Konsequenzen die Lehrerin oder der Lehrer aus den gewonnenen Informationen ziehen.

Insofern sind zum einen die Verfahrensweisen nach Abschluss der SEfU Befragung (also nach dem Erzeugen des Ergebnisberichts) wichtig: Wie kommunizieren Lehrerinnen und Lehrer das Ergebnis mit ihrer Klasse? Werden Änderungen mit Schülerinnen und Schülern vereinbart oder weitere spezifischere Befragungen durchgeführt (siehe Module)? Werden Ergebnisse im Kollegium diskutiert? Dazu wird eine empirische Untersuchung in Form einer Nutzerbefragung der SEfU-Teilnehmerinnen und Teilnehmer durchgeführt. Zum anderen werden in dieser Befragung Stärken und Schwächen des SEfU-Instruments erhoben, um die Nutzbarkeit des Verfahrens für Feedbackprozesse zu überprüfen.

Durch Auswertung der erhobenen Daten und bereits bestehender Nutzungsstatistiken können Konsequenzen und perspektivische Entwicklungsmöglichkeiten für computerbasiertes Schülerfeedback abgeleitet werden.

Hattie, J. (2012): Visible learning for teachers: maximizing impact on learning. London: Routledge.

Börner, H./Nachtigall, C. (2013): Unterricht reflektieren – Schüler als Experten nutzen. Praxis Schule. 4/2013, S. 46–49.

## **quop: Ein Ansatz internetbasierter Lernverlaufsdagnostik mit Testkonzepten für Lesen und Mathematik**

### Theoretischer Hintergrund

Der Ansatz der Lernverlaufsdagnostik ist eine attraktive Möglichkeit, Lehrkräften und Schülern eine zeitnahe Rückmeldung zu Leistungsniveau und Lernzuwächsen zu geben (Deno, 2003; Fuchs, 2004). Lernverlaufsdagnostik kann damit eine Grundlage für die Optimierung von Lerngelegenheiten darstellen. Die Anforderungen an eine aussagekräftige und gleichzeitig für die schulische Praxis geeignete Lernverlaufsdagnostik sind allerdings hoch (Förster & Souvignier, 2011): Die Testverfahren müssen kurz und inhaltlich so angelegt sein, dass sie unmittelbare Hinweise auf individuellen Förderbedarf geben. Eine automatisierte Auswertung und Dokumentation der Ergebnisse trägt zur Nutzerfreundlichkeit bei. Die Testverfahren selbst müssen sich jenseits klassischer Anforderungen an die psychometrische Güte als parallel und als sensitiv für Veränderungen erweisen. Computerbasierte Testungen ermöglichen dabei zeitlich flexible, eigenständige Testdurchführungen der Schülerinnen und Schüler sowie eine automatisierte Auswertung und Rückmeldung an die Lehrkräfte. Sie stellen somit eine vielversprechende Möglichkeit dar, den skizzierten Herausforderungen gerecht zu werden. Ein internetbasierter Ansatz bietet darüber hinaus den Vorteil, dass kontinuierlich aktuelle Normwerte zur Verfügung gestellt und Programmaktualisierungen zentral gesteuert werden können. Vor diesem Hintergrund wurde das quop-System der Lernverlaufsdagnostik mit bislang fünf Testreihen für Lesen (Klassenstufe drei und vier) und Mathematik (Klassenstufe eins, zwei und drei) in der Grundschule entwickelt und erprobt.

### Fragestellung

Zunächst wurde geprüft, ob sich das internetbasierte System zur Lernverlaufsdagnostik überhaupt in die schulische Praxis implementieren lässt. Darüber hinaus wurden die psychometrische Qualität der fünf Testreihen sowie Effekte des Einsatzes auf Leistungsentwicklungen von Schülerinnen und Schülern untersucht.

### Methode

Die Untersuchungen zur Implementation, zur psychometrischen Güte der Testkonzepte und zu Effekten der Lernverlaufsdagnostik wurden in mehreren Einzelstudien vorgenommen, die hier zusammenfassend dargestellt werden. An einer Befragung zur Implementation des Ansatzes internetbasierter Lernverlaufsdagnostik nahmen 40 Lehrkräfte teil. Die Untersuchung der psychometrischen Güte der Testkonzepte wurde in einer Stichprobe von insgesamt 2.291 Kindern vorgenommen, Effekte der Lernverlaufsdagnostik wurden bislang in drei Evaluationsstudien mit insgesamt 2.002 Schülerinnen und Schülern überprüft. Die Testreihen umfassten jeweils acht Messzeitpunkte und wurden in einem zeitlichen Rahmen von sechs Monaten absolviert (3 Wochen Abstand von Test zu Test). Die Testkonzepte sind so angelegt, dass differenzierte Hinweise auf individuellen Förderbedarf abgeleitet werden können: Die Lesetests für die 3. und 4. Klassenstufe liefern mit der Lesegeschwindigkeit, der Lesegenauigkeit sowie dem Verständnis bei textimmanenten und schlussfolgernden Anforderungen mehrere Indikatoren für die Lesefähigkeit. Die Mathematiktests für die 1.-3. Klassenstufe erfassen insbesondere Vorläuferkompetenzen und curriculare Kompetenzen in den Bereichen Rechnen und Geometrie. Die Konstruktion der einzelnen Testverfahren folgt der Überlegung, konzeptuell vollständig vergleichbare Paralleltests vorzulegen. Daher wurden Aufgabenmerkmale auf Itemebene so genau spezifiziert, dass formal exakt vergleichbare Einzeltests eine Testreihe bilden. Die Validierung der Testverfahren wurde anhand standardisierter Testverfahren (z.B. DEMAT; SLS; HAMLET) vorgenommen. Zur Evaluation der Effekte der Lernverlaufsdagnostik wurden jeweils Prätest-Posttest-Kontrollgruppendesigns verwendet, anhand derer geprüft wurde, ob sich die Lernzuwächse in standardisierten Testverfahren zwischen den verschiedenen Untersuchungsbedingungen unterscheiden.

### Ergebnisse

In der schulischen Praxis zeigen Rückmeldungen von Lehrkräften und Schülern, dass der Ansatz einer internetbasierten Lernverlaufsdagnostik auf hohe Akzeptanz stößt und damit eine geeignete Basis für ein wirksames formatives Assessment darstellt. Für alle Testreihen gilt, dass hohe interne Konsistenzen und gute Validitätskennwerte vorliegen. Die Schwierigkeiten der Tests weisen auf eine unmittelbare Vergleichbarkeit hin, ebenso sind die Paralleltestkorrelationen hoch. Die Sensitivität ist bei einigen Skalen – mitunter aufgrund von Deckeneffekten – eingeschränkt. Mit dem überwiegenden Teil der Skalen lassen sich aber signifikante Leistungszuwächse von Messzeitpunkt zu Messzeitpunkt abbilden. Alle drei Evaluationsstudien zeigen höhere Lernzuwächse von Schülerinnen und Schülern, deren Lehrkräfte die Lernverlaufsdagnostik in ihrem Unterricht einsetzten und deuten somit darauf hin, dass Informationen über Lernverläufe eine wertvolle Grundlage für effektives Unterrichten darstellen.

Deno, S. L. (2003). Curriculum-based Measures: Development and perspectives. *Assessment for Effective Intervention*, 28, 3–11.

Förster, N. & Souvignier, E. (2011). Curriculum-based measurement: Developing a computer-based assessment instrument for monitoring student reading progress on multiple indicators. *Learning Disabilities: A Contemporary Journal*, 9, 21-44.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33, 188-192.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G131

## Technisch-organisatorische Anforderungen an CBA in Schulen

Ansätze computerbasierter Verfahren für Diagnostik und Rückmeldung in der Schule werden in Deutschland eher punktuell eingesetzt und stehen am Anfang einer breiten Verwendung. Gründe hierfür lassen sich bei den Verfahren selbst ausmachen, sind aber v.a. in den technisch-organisatorischen Rahmenbedingungen für das IT-Management in deutschen Schulen zu sehen.

Aus Informatik-Perspektive sind funktionale und nicht-funktionale Anforderungen zu unterscheiden (Sommerville, 2011). Die funktionale Eignung der Systeme für den Nutzungszweck wird hier vorausgesetzt. Die nicht-funktionalen Anforderungen reichen von rechtlichen Vorgaben über Informationssicherheit bis zur softwareergonomischen Gestaltung. In einer empirisch fundierten Machbarkeitsstudie wurden verschiedene Umsetzungsszenarien analysiert. Ausgangspunkt war dabei der Einsatz computerbasierter Verfahren in Large-scale Assessments (Breiter, Groß, & Stauke, 2013), was für diesen Beitrag durch zentrale Fragen für die Nutzung im Unterricht erweitert wurde. Dabei liegt der Schwerpunkt auf drei Kernaspekten: (a) Bereitstellung und Betrieb; (b) Informationssicherheit und Datenschutz; (c) Benutzbarkeit.

(a) Das deutsche Schulsystem hat durch die Trennung in innere und äußere Schulangelegenheiten eine Governance-Struktur, die prinzipiell den Einsatz von IT-Systemen erschwert. Zuständig für die Beschaffung von Geräten und Software, die Bereitstellung und den Betrieb der IT-Infrastruktur sind die Kommunen als Schulträger. Das Land setzt die inhaltlichen Vorgaben. Eine Ausstattung über das erforderliche Niveau hinaus erfolgt nur dann durch die Schulträger, wenn das Land eine entsprechende Verordnung erlässt und Kosten übernimmt. Dies betrifft insbesondere die Personalkosten für den technischen Support. Dieser Abstimmungsprozess ist konfliktbehaftet und in allen Bundesländern politisch heikel (Breiter, Fischer, & Stolpmann, 2006). Die alternative Nutzung eigener Endgeräte der Schülerinnen und Schüler („Bring-your-own-device“) löst das Beschaffungsproblem, schafft aber neue Schwierigkeiten wegen Kompatibilitätsproblemen, Haftungsfragen und Fragen der Gleichbehandlung (Welling & Stolpmann, 2013).

(b) Der Einsatz computerbasierter Assessments in Schulen setzt zwingend die Verarbeitung personenbezogener Daten voraus, um schülerindividuelle Rückmeldungen zu ermöglichen. Die Erhebung, Speicherung und Verarbeitung ist nur dann zulässig, wenn die Betroffenen (bzw. deren Erziehungsberechtigte) zugestimmt haben oder es eine gesetzliche Grundlage gibt. Aufgrund der Sensibilität von Eltern und der kritischen Öffentlichkeit in Bezug auf die maschinelle Verarbeitung von personenbezogenen Daten in der Schule ist es zweifelhaft, ob hierzu allgemein gültige Regelungen getroffen werden können. Dies hat zur Konsequenz, dass jede Schule dies mit dem örtlichen Datenschutzbeauftragten zu regeln hat. Der zweite Aspekt betrifft die verlässliche Speicherung der Daten, ihre Unveränderbarkeit und die Sicherstellung der Verfügbarkeit. Für eine sinnvolle, ggf. langfristige Nutzung der Daten für die Diagnostik (bspw. über mehrere Schuljahre) müssen sich alle Beteiligten darauf verlassen können, dass die Systeme funktionsfähig und verfügbar sind. Im Falle eigener Endgeräte ist dies gar nicht, bei schuleigenen Geräten nur in den Grenzen der Leistungsfähigkeit des Schulträgers zu gewährleisten, mit dem Risiko des Ausfalls (bei Prüfungen) und des Datenverlustes.

(c) Der dritte Bereich betrifft die Benutzbarkeit und Verständlichkeit der Datenvisualisierungen. Die Gestaltung der Mensch-Maschine-Schnittstelle gilt als zentraler Faktor für die Nutzerakzeptanz (Preece, Rogers, & Sharp, 2007). Hierauf weisen nur wenige Forschungsergebnisse zu CBA ausdrücklich hin (Georgiadou, Triantafillou, & Economides, 2005; Valenti, Cucchiarelli, & Panti, 2002). Es kann daher nicht davon ausgegangen werden, dass Lehrkräfte in der Unterrichtssituation mit komplexen Interaktionsformen zurechtkommen – gleiches gilt für Schülerinnen und Schüler. Stauke (2012) und Koch (2011) haben zeigen können, dass die Visualisierung der Rückmeldungen aus den Test entscheidenden Anteil am Verständnis der Ergebnisse und damit für die wirksame Unterstützung von Lernprozessen hat. Erschwerend kommt hinzu, dass aufgrund der Inklusion die Anforderungen an die Barrierefreiheit in den Vordergrund rücken, was bisher in Deutschland noch nicht Gegenstand der Forschung ist (Reich & Petter, 2009). Im Kern lässt sich schließen, dass unter den derzeitigen technisch-organisatorischen Rahmenbedingungen schulischen Handelns ein breiter Einsatz von computerbasierter Assessments und damit ein Transfer in die alltägliche Unterrichtspraxis eher unwahrscheinlich ist, solange nicht zentrale Probleme der technischen Verfügbarkeit, der Benutzbarkeit und des Datenschutzes gelöst sind.

Breiter, A., Fischer, A., & Stolpmann, B. E. (2006). IT-Service-Management – neue Herausforderungen für kommunale Schulträger. In M. Wind & D. Kröger (Hrsg.), *Handbuch IT in der Verwaltung* (S. 254-274). Berlin: Springer.

Breiter, A., Groß, L. M., & Stauke, E. (2013). Computer-Based Large-Scale Assessments in Germany. In D. Passey, A. Breiter & A. J. Visscher (Hrsg.), *Next Generation of IT in Educational Management* (S. 41-54). Berlin: Springer.

Georgiadou, E., Triantafillou, E., & Economides, A. A. (2005). Evaluation parameters for computer-adaptive testing. *British Journal of Educational Technology*, 37(2), 261-278.

Koch, U. (2011). Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten. Münster: Waxmann.

Preece, J., Rogers, Y., & Sharp, H. (2007). *Interaction Design: Beyond Human-Computer Interaction* (2nd edition ed.). New York, NY: John Wiley & Sons.

Reich, K., & Petter, C. (2009). eInclusion, eAccessibility and Design-for-All Issues in the Context of European Computer-based Assessment. In F. Scheuermann & J. Björnsson (Hrsg.), *The Transition to Computer-Based Assessment. New Approaches to Skills Assessment and Implications for Large-scale Testing* (S. 68-74). Luxembourg: Office for Official Publications of the European CommUniversitätsitäties.

Sommerville, I. (2011). *Software engineering* (9th edition ed.). Harlow: Addison-Wesley.

Stauke, E. (2012). Konzeption und Anforderungsdefinition eines elektronischen Rückmeldesystems für schulische Leistungsdaten. Dr.-Ing., Universität Bremen, Bremen.

Valenti, S., Cucchiarelli, A., & Panti, M. (2002). Computer Based Assessment Systems Evaluation via the ISO 9126 Quality Model. *Journal of Information Technology Education*, 1(3), 157-175.

Welling, S., & Stolpmann, B. E. (2013). Mobile Computing in der Schule – Zentrale Herausforderungen am Beispiel eines Schulversuchs zur Einführung von Tablet-PCs. In R. Schulz-Zander, B. Eickelmann, H. Moser, H. Niesyto & P. Grell (Hrsg.), *Jahrbuch Medienpädagogik* 9 (S. 193-217). Wiesbaden: Springer VS.



## Zur Vergleichbarkeit von Kompetenzstufenklassifikationen auf Grundlage des Gemeinsamen Europäischen Referenzrahmens für Sprachen

Chair(s): Johanna Fleckenstein (IPN Kiel)

Die Beherrschung des Englischen als globale Verkehrssprache (Lingua Franca) zählt zu den zentralen Basisqualifikationen für die Bewältigung beruflicher und gesellschaftlicher Anforderungen (Tenorth, 2001). In einer globalisierten Welt gelten sichere Englischkenntnisse häufig als notwendige Einstellungsvoraussetzungen und auch im Universitätsstudium sind englische Lehr- und Forschungstexte kaum zu umgehen (Köller & Trautwein, 2004). Dies gilt insbesondere für die steigende Anzahl international ausgerichteter Studiengänge sowie für ein Studium im englischsprachigen Ausland.

Der Gemeinsame Europäische Referenzrahmen (GER) wurde entwickelt, um die Vergleichbarkeit und gegenseitige Anerkennung von Sprachqualifikationen zu erleichtern (Europarat, 2001). Er klassifiziert fremdsprachliche Kompetenzen auf sechs Stufen und bündelt diese in übergeordnete Kategorien: Basic User (A1/A2), Independent User (B1/B2) und Proficient User (C1/C2). Diese Stufen werden häufig für die Klassifikation von Leistungen in standardisierten Sprachtests eingesetzt (Little, 2007). Die Validität und Äquivalenz solcher GER-basierter Kompetenzstufenklassifikationen ist Gegenstand kritischer Diskussionen (Weir, 2005), jedoch fehlt es bislang noch weitgehend an empirischen Linking-Studien.

Es existieren verschiedene, auf dem GER verortete Tests für Englisch als Fremdsprache (EFL). Diese Studie untersucht zum einen, inwiefern sich deren Kompetenzstufenklassifikationen empirisch voneinander unterscheiden: Stellen die deutschen Bildungsstandards den gleichen Anspruch an einen Proficient User wie der international anerkannte Test of English as a Foreign Language (TOEFL)? Zum anderen stellt sich die Frage nach der praktischen Relevanz der Bildungsstandards: Decken sich die Erwartungen, denen hierzulande die Englischleistungen von Abiturientinnen und Abiturienten genügen müssen mit denjenigen, die in ihrer weiteren akademischen und beruflichen Laufbahn an sie gestellt werden?

Der parallele Einsatz von Aufgaben aus den Bildungsstandards und dem TOEFL ermöglicht eine Verlinkung der beiden Skalen. Dieses Vorgehen dient der Verortung und Validierung nationaler Bildungsstandards sowie deren Überprüfung auf praktische Relevanz in einer globalisierten, englischsprachigen Welt.

Die hier untersuchte Stichprobe umfasst N=1134 Abiturientinnen und Abiturienten an allgemeinbildenden und berufsbildenden Gymnasien in Schleswig-Holstein. Wie bereits in den Studien TOSCA und TOSCA-Repeat (Köller et al., 2004; Trautwein et al., 2010) wurden hier die Englischleistungen anhand einer Kurzform des TOEFL gemessen. Zusätzlich wurden Aufgaben aus dem Ländervergleich 2009 (s. Köller, Knigge & Tesch, 2010) eingesetzt, die der Überprüfung der Bildungsstandards für das Leseverstehen in Englisch am Ende der Sekundarstufe I dienen. Letztere wurden als eindimensionales 1-PL Rasch-Modell in ConQuest skaliert. Personenfähigkeiten (PVs) wurden mit Hintergrundmodell geschätzt, auf die 500/100-Metrik des Ländervergleichs transformiert und in fünf Kompetenzstufen (A1-C1) kategorisiert. Die Daten aus den drei TOEFL Subtests (Listening Comprehension/Structure and Written Expression/Vocabulary and Reading) wurden anhand der Transformationsvorschriften des Educational Testing Service (ETS, 1997) analog zu TOSCA auf die entsprechende Metrik transformiert und zu einem globalen Testwert aggregiert.

Der eingesetzte Test zu den Bildungsstandards zeigt eine akzeptable PV-Reliabilität von .82, der TOEFL ein Cronbachs  $\alpha$  von .92. Der Vergleich der Kompetenzstufenklassifikation beider Tests auf dem GER ergab einen Rangkorrelationskoeffizienten von  $\rho=.42$ . Trotz weitgehender Übereinstimmung existiert eine Verschiebung in der Verortung der Stufen: So liegen 60.7% der Schülerinnen und Schüler, die laut Bildungsstandards auf GER-Niveau B verortet sind, bei dem TOEFL auf der gleichen Stufe. Ein Drittel von ihnen liegt laut TOEFL auf Niveau A. Dies spricht dafür, dass der TOEFL im Durchschnitt ein höheres Maß an Sprachkompetenz zur Einordnung auf einer GER-Stufe verlangt als die Bildungsstandards. Zur Validierung und internationalen Einordnung der Bildungsstandards wurden die TOEFL-Werte für die Zulassung zu international ausgerichteter Studiengänge und amerikanischen Universitätsveritäten als Kriterium herangezogen. Beispielsweise erreichen 19% der Abiturientinnen und Abiturienten die sprachbezogenen Aufnahmekriterien einer prestigeträchtigen, öffentlichen Universität in den USA (TOEFL-Wert = 550).

Die Ergebnisse zeigen, dass die Entwicklungsprozesse von EFL Tests durchaus zu divergierenden Kompetenzstufenklassifikationen anhand des GER gelangen können. Die Studie hat zudem bildungspolitische Implikationen, indem sie die deutschen Bildungsstandards anhand von externen Validitätskriterien und praktischer Relevanz überprüft.

Educational Testing Service (ETS) (1997). TOEFL 1997 edition. Test & score manual. Princeton, NJ: ETS.

Europarat (2001). Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen. Berlin: Langenscheidt.

Köller, O., Knigge, M. & Tesch, B. (2010). Sprachliche Kompetenzen im Ländervergleich. Münster: Waxmann.

Köller, O., Trautwein, U. (2004). Englischleistungen von Schülerinnen und Schülern an allgemeinbildenden und beruflichen Gymnasien. In: Köller, O., Watermann, R., Trautwein, U., Lüdtke, O. (Hrsg.). Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. S. 285-326. Opladen: Leske + Budrich.

Köller, O., Watermann, R., Trautwein, U. & Lüdtke, O. (Hrsg.). (2004). Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. S. 69-111. Opladen: Leske + Budrich.

Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. The Modern Language Journal, 91, 645-55.

Tenorth, H.-E. (2001). Englisch: Ein Kerncurriculum, seine Notwendigkeit und seine Gestalt - Zusammenfassung. In H.-E. Tenorth (Hrsg.). Kerncurriculum Oberstufe. Mathematik - Deutsch - Englisch. Expertisen im Auftrag der Ständigen Konferenz der Kultusminister. S. 156-161. Weinheim: Beltz.

Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (Hrsg.). (2010). Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Weir, C. J. (2005). Limitations of the Common European Framework for developing comparable examinations and tests. Language Testing, 22, 281-300.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G135



## ***Leistungsentwicklung in der ersten Fremdsprache Englisch im Verlauf der Sekundarstufe II***

Über die längsschnittliche Entwicklung schulischer Leistungen in der ersten Fremdsprache Englisch in der Sekundarstufe II ist empirisch so gut wie nichts bekannt. Die vorliegende Studie setzt hier an, indem sie Lernstände allgemeinbildender Gymnasiasten der Jahrgänge 11 und 13 (N = 1190) in Schleswig-Holstein zueinander in Beziehung setzt. Die Lernstände werden abgetragen auf der Metrik der Bildungsstandards für das Fach Englisch (vgl. Köller, Knigge, Tesch, 2010). Dies ermöglicht erstmals direkte kriteriale Aussagen darüber, welche Kompetenzniveaus im Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER; s. Europarat, 2001) mit dem Bildungszertifikat des Abiturs korrespondieren. Unlängst wurde das GER-Niveau B2 (bzw. für rezeptive Kompetenzen im Fach Englisch: C1) als Regelstandard für das Erreichen der Allgemeinen Hochschulreife als politische Zielvorgabe festgesetzt (vgl. KMK, 2012). Zudem prüfen wir eine Korrespondenz der erreichten Lernstände am Ende der Klasse 13 mit schriftlichen Abiturnoten im Fach Englisch.

Zur Überprüfung der Fragestellungen wurden im Rahmen der LISA-Studie (z.B. Retelsdorf, Becker, Köller & Möller, 2012) Testungen am Ende der Jahrgangsstufe 11 (Testung: April 2011) und 13 (Testung: April 2013) realisiert. Die längsschnittlich untersuchte Stichprobe stammt aus der Kohorte, die 2009 für Schleswig-Holstein in den Ländervergleich zur Überprüfung der Bildungsstandards im Fach Englisch am Ende der Sekundarstufe I einging. Die Schülerinnen und Schüler (SuS) besuchen fünf verschiedene Profilerbestufen (naturwissenschaftlich, 30.9%; sprachlich, 22.5%; gesellschaftswissenschaftlich, 30.3%; ästhetisch, 9.0%; sportlich, 7.3%). Die eingesetzten Aufgabenmengen (Klasse 11: 133 bzw. 118 Items Lese- und Hörverstehen, PV-Reliabilitäten: .86 bzw. .85; Klasse 13: 42 bzw. 32 Items, PV-Reliabilität jeweils .82) weisen substantielle Überschneidungsbereiche zu den Aufgabenmengen des Ländervergleichs von 2009 auf (Jg. 11: 82 bzw. 80, Jg. 13: 24 bzw. 28 Items). Für beide Messzeitpunkte erfolgte ein Equating auf die Metrik des Ländervergleichs im Rahmen der Item Response Theorie. Die resultierenden Plausible Values wurden auf die 500/100 Metrik des Ländervergleichs transformiert. In einer Nacherhebung wurden schriftliche Abiturnoten im Fach Englisch erhoben (Punkte 0 bis 15; transformiert in ganze Noten von 1 bis 5); N = 922 der untersuchten SuS hatten Englisch als schriftliches Prüfungsfach belegt.

Bislang wird für die untersuchten Kompetenzen innerhalb der Sekundarstufe I von einem Leistungszuwachs von 35-40 Punkten pro Schuljahr ausgegangen (s. Köller, Knigge & Tesch, 2010). Unsere Ergebnisse zeigen, dass sich die Entwicklung im Leseverstehen in der Sekundarstufe II deutlich verlangsamt (Jg. 9 zu Jg. 11: 28 Punkte pro Schuljahr, Jg. 11 zu Jg. 13: 9 Punkte pro Schuljahr). Im Hörverstehen zeigen sich unverändert stabile Zuwachsraten (32 bzw. 36 Punkte pro Schuljahr für dieselben Zeiträume). Trotz erheblicher Niveauunterschiede zwischen den Profilerbestufen zeigen sich nur vereinzelte Hinweise auf differenzielle Lernzuwächse. Die Analysen zur Korrespondenz mit den schriftlichen Abiturnoten belegen beispielsweise, dass selbst SuS mit der Note 5 im Leseverstehen noch ein Niveau erreichen, das als Bildungsstandard für den Mittleren Schulabschluss gilt (d.h. B1.2). Im Hören erreichen dieselben SuS im Mittel sogar Leistungen, die als Bildungsstandard für das Abitur in Englisch insgesamt definiert wurden (d.h. B2). Zudem zeigen sich deutliche Hinweise für eine unterschiedliche Bewertung ein und derselben Leistungen an den Profilerbestufen: So beträgt etwa die Spannbreite der erreichten Lernstände für die Abiturnote „3“ 40 Punkte im Lese- und 45 Punkte im Hörverstehen.

Europarat (2001). Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Berlin: Langenscheidt.

KMK (2012). Bildungsstandards für die fortgeführte Fremdsprache (Englisch / Französisch) für die Allgemeine Hochschulreife (Beschluss der Kultusministerkonferenz vom 18.10.2012).

Köller, O., Knigge, M. & Tesch, B. (2010) (Hrsg.). Sprachliche Kompetenzen im Ländervergleich. Überprüfung der Erreichung der Bildungsstandards für den Mittleren Schulabschluss für Deutsch und die erste Fremdsprache in der neunten Jahrgangsstufe. Münster: Waxmann.

Retelsdorf, J., Becker, M., Köller, O. & Möller, J. (2012). Reading development in a tracked school system: A longitudinal study over 3 years using propensity score matching. *British Journal of Educational Psychology*, 82, 647-671.

## ***Mathematische Leistungsentwicklung in englischem Immersionsunterricht***

Immersionunterricht stellt eine Integration von Sprach- und Fachunterricht dar, bei dem Schülerinnen und Schüler in einer Fremdsprache (z. B. Englisch) unterrichtet werden. Internationale Studie belegen, dass immersiv unterrichtete Schülerinnen und Schüler trotz Unterrichts in einer Fremdsprache ähnliche bis bessere erstsprachliche und mathematische Leistungen erbringen als konventionell unterrichtete Schülerinnen und Schüler, sofern das Lesen und Schreiben in der Erstsprache erlernt wurde (z. B. Cheng et al., 2010; Genesee, 2004, 2007; Genesee & Jared, 2008). Insbesondere für Mathematik finden sich konsistente Belege für Vorteile bei immersiv unterrichteten Schülerinnen und Schülern, die auf einen förderlichen Effekt von Immersionunterricht auf mathematische Leistungen hindeuten (z. B. Cheng et al., 2010; Dubé & MacFarlane, 1991; Turnbull, Hart & Lapkin, 2003). Für Deutschland fehlen bisher längsschnittliche Untersuchungen zur Entwicklung von mathematischen Leistungen im Immersionunterricht. In der vorliegenden Studie wurde daher längsschnittlich die mathematische Leistungsentwicklung mithilfe von deutschsprachigen, curriculumvaliden Mathematiktests untersucht. Die Stichprobe umfasste N = 590 teilimmersiv unterrichtete Grundschülerinnen und Grundschüler (Unterrichtssprache in allen Fächern ist die L2 Englisch mit Ausnahme des Faches Deutsch), die von der ersten bis zur vierten Jahrgangsstufe getestet wurden. Die Ergebnisse der Wachstumskurvenanalysen zeigen bereits Ende der ersten Jahrgangsstufe Vorteile zugunsten der teilimmersiv unterrichteten Schülerinnen und Schüler sowie im gesamten Verlauf eine schnellere Leistungsentwicklung als bei den konventionell unterrichteten Schülerinnen und Schülern. Dies ist besonders bemerkenswert vor dem Hintergrund, dass Mathematik auf Englisch unterrichtet wird und die Tests deutschsprachig waren (vgl. auch De Courcy & Burston, 2000). Die Befunde werden in Hinblick auf positive Effekte des Immersionunterrichts auf kognitive Funktionen, insbesondere Exekutivfunktionen, sowie interlinguale Transferprozesse, aber auch plausible alternative Erklärungen wie (fremd-)sprachbedingte spezifische Lehrmethoden und Lernmechanismen diskutiert.

Cheng, L., Li, M., Kirby, J. R., Quiang, H., & Wade-Woolley, L. (2010). English language immersion and students' academic achievement in English, Chinese and mathematics. *Evaluation & Research in Education*, 23, 151–169.

De Courcy, M. C. & Burston, M. (2000). Learning mathematics through French in Australia. *Language and Education*, 14, 75-95.

Dubé, L. & MacFarlane, A. (1991). Middle immersion: Is it a better option than early or late? *Immersion Journal*, 14, 21-27.

Genesee, F. (2004). What do we know about bilingual education for majority-language students? In T. K. Bhatia & W. C. Ritchie (Hrsg.), *The handbook of bilingualism* (S. 547–576). Malden, MA: Blackwell Publishing.

Genesee, F., & Jared, D. (2008). Literacy development in early French immersion programs. *Canadian Psychology/Psychologie canadienne*, 49, 140–147.

Turnbull, M., Hart, D. & Lapkin, S. (2003). Grade six French immersion students' performance on large-scale reading, writing, and mathematics tests: Building explanations. *The Alberta Journal of Educational Research*, XLIX, 6-23.

## **Selektion und Preparation als Ursache für die hohen Sprachkompetenzen bilingualer Lerner? Kritische Evaluation einer Längsschnittstudie**

In der Fremdsprachendidaktik wird bilingualer Sachfachunterricht (bili SFU; engl. CLIL) als eines der vielversprechendsten Fremdsprachenlernprinzipien bzw. -konzepte des 21. Jahrhunderts angesehen (García, 2009). In zahlreichen Studien (z.B. Klieme, 2008; Zydariß, 2007) erwiesen sich die Teilnehmer des bilingualen Sachfachunterrichts (bili SFU, engl. CLIL) gegen Ende der Sekundarstufe I als fremdsprachlich deutlich kompetenter als ihre peers in Regelklassen. Vielfach existieren jedoch nur (explorative) querschnittliche Studien mit geringen Probandenzahlen und qualitativen Designs. Zudem deuten die fallstudienartigen Daten einiger Forscher auf nicht zu unterschätzende Verzerrungseffekte hin (z.B. Fehling, 2008; Bredenbröcker, 2000; Burmeister, 1994), die durch die Eingangsselektion der Schüler und ihrer Vorbereitung auf den bili SFU in den Klassen 5 und 6 verursacht werden und damit die Vergleichbarkeit von CLIL- und Regelschülern beeinträchtigen.

Um sowohl diese Verzerrungseffekte als auch den tatsächlichen Einfluss des CLIL-Unterrichts näher zu untersuchen, wurde in den Jahren 2011-2013 eine quasi-experimentelle Längsschnittstudie (N=1.275) durchgeführt. Dazu wurde am Ende der 6.Klasse, d.h. unmittelbar vor dem Beginn des bili SFU, sowie ein und zwei Jahre danach u.a. allgemeine englische Sprachkompetenz, verbale kognitive Fähigkeiten, Interesse und Selbstkonzept von Schülern bilingualer und nicht-bilingualer Zweige erhoben.

Das Erkenntnisinteresse der Studie besteht demnach in der Untersuchung der Unterschiede, die bereits vor und damit unabhängig vom bilingualen Sachfachunterricht bestehen und in bisherigen Studien nicht berücksichtigt wurden; zudem wird so ermöglicht, die Leistungsfähigkeit des bilingualen Sachfachunterrichts, so wie er momentan durchgeführt wird, im Hinblick auf die fremdsprachliche und affektiv-motivationale Entwicklung der teilnehmenden Schüler deutlich fehlerbereinigter zu evaluieren und wichtige Erkenntnisse im Bereich einer essentiellen fremdsprachendidaktischen Forschungslücke zu erlangen.

Die Stichprobe besteht aus 476 Schülern bilingualer Gymnasialzweige (16 Klassen) und 429 Schülern (15 Klassen) nicht-bilingualer Gymnasialzweige der gleichen Schulen; dies wird ergänzt durch eine Referenzgruppe von 170 Schülern (7 Klassen), die von Schulen ohne bilinguale Zweige stammen, wo keine Selektions- und Preparationseffekte beobachtbar sind. Zur Messung der allgemeinen Sprachkompetenz dienten die C-Tests des Hamburger Schulleistungstests für 6./7. Sowie 8./9. Klassen (Behörde für Schule, Jugend und Berufsbildung. Amt für Schule, Hamburg, 1998 u. 2000); verbale kognitive Fähigkeiten wurden mit Hilfe zweier Skalen des KFT (Heller & Perleth, 2000) erhoben. Die eingesetzten Messinstrumente für die in diesem Vortrag zentralen affektiv-motivationalen Konstrukte basieren auf etablierten Messinstrumenten (Interesse: u.a. Baumert et al., 1986; akademisches Selbstkonzept: u.a. Frey et al., 2009), wurden jedoch zum Teil leicht adaptiert. Für den ersten Messzeitpunkt liegt Cronbach's Alpha bei .77 (5 items) und .88 (9 items) für die beiden Interessenskalen und bei .9 für die Skala zum akademischen Selbstkonzept.

Verbale kognitive Fähigkeiten ( $r=.51$ ), Selbstkonzept ( $r=.41$ ) und Interesse ( $r=.24$ ) wiesen hochsignifikante Pearson-Korrelationen mit der Sprachkompetenz auf. Parametrische (ANOVA und t-Tests) wie auch nicht-parametrische Verfahren (Kruskal-Wallis und Mann-Whitney-Tests) zeigten für alle Variablen bereits vor dem bili SFU die Existenz hochsignifikanter Unterschiede zwischen beiden Schülergruppen bei substantiellen Effektstärken; eine Referenzgruppe aus 170 Schülern, die zwei Schulen ohne bilingualen Sachfachunterricht besuchen, liegt entsprechend der fehlenden Selektion und Preparation auf einem mittleren Niveau zwischen den CLIL und den nicht-CLIL Schülern im Bereich der fremdsprachlichen Kompetenzen (Kruskal-Wallis:  $H(2)=189$ .  $p<.000$ ; Effektstärken im Vergleich aller drei Gruppen:  $.2<r<.52$ , mit dem größten Unterschied zwischen CLIL und nicht-CLIL Schülern). Im darauffolgenden ersten Jahr bilingualen Unterrichts entwickelten sich die Rasch-Fähigkeitswerte aller drei Gruppen parallel, d.h. die a priori existierenden Unterschiede blieben, quantitativ gesehen, gleich groß. [Die vollständigen (Längsschnitt-)Analysen können durch die noch laufende Dateneingabe erst auf der Konferenz präsentiert werden].

Die vorläufigen Ergebnisse bestätigen die Vermutung, dass in bisherigen, querschnittlichen Studien mit nur einer Messung die Effekte des bili SFU überschätzt wurden. Selektions- und Preparationseffekte wurden zwar häufig in der Diskussion im Hinblick auf die eingeschränkte Generalisierbarkeit der Ergebnisse genannt, jedoch scheint die hier zu Tage getretene Bedeutung dieser Einflüsse nicht in ausreichendem Maße berücksichtigt worden zu sein.

- Baumert, J., Roeder, P. M., Sang, F., & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. Zeitschrift für Pädagogik, 32, 639–660. Behörde für Schule, Jugend und Berufsbildung. Amt für Schule, Hamburg (Ed.). (1998). Der Hamburger Schulleistungstest für sechste und siebte Klassen – SL-HAM 6/7. Hamburg. Behörde für Schule, Jugend und Berufsbildung. Amt für Schule, Hamburg (Ed.). (2000). Der Hamburger Schulleistungstest für achte und neunte Klassen – SL-HAM 8/9. Hamburg.
- Bredenbröcker, W. (2000). Förderung der fremdsprachlichen Kompetenz durch bilingualen Unterricht: Empirische Untersuchungen. Frankfurt am Main: Lang.
- Burmeister, P. (1994). Englisch im Bili-Vorlauf: Pilotstudie zur Leistungsfähigkeit des verstärkten Vorlaufs in der 5. Jahrgangsstufe deutsch-englisch bilingualer Zweige in Schleswig-Holstein. Kiel: I&F.
- Fehling, S. (2008). Language awareness und bilingualer Unterricht: Eine komparative Studie (2. Aufl.). Frankfurt am Main: Lang.
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., et al. (2009). PISA-2006-Skalenhandbuch: Dokumentation der Erhebungsinstrumente. Münster: Waxmann.
- Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest (KFT) für 4.-12. Klassen, Revision. Göttingen: Beltz.
- Klieme, E. (Hrsg.). (2008). Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie. Weinheim: Beltz.
- Zydariß, W. (2007). Deutsch-Englische Züge in Berlin (DEZIBEL): Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien. Frankfurt am Main: Lang.



## ***Historische Kompetenzen: Zum Potential inhaltsanalytischer Auswertungen qualitativer Daten zur Komplexitätsreduktion***

Chair(s): Waltraud Schreiber (Katholische Universität Eichstätt-Ingolstadt)

Diskutant/in: Bodo v. Borries (Universität Hamburg)

Historische Kompetenzen zu messen stellt Empirische Bildungsforschung und Geschichtsdidaktik vor Herausforderungen. Deren Zeit- und Kulturabhängigkeit erschwert die Entwicklung kontextreduzierter Testsituationen und die prinzipielle und vielschichtige Perspektivität von Geschichte die Definition, „richtiger“ Antworten. Die Güte von Denkprozessen offenbart sich in schwer erschließbaren, (selbst-)reflexiven Prozessen.

Als Grundbedingung für die empirische Erschließung hat sich der Bezug auf klar strukturierte Kompetenzmodelle erwiesen. Das international anschlussfähige FUER-Modell bietet die Grundlage für die im Symposium vorgestellten Studien. Als Kompetenzstrukturmodell erlaubt es, zu präzisieren, worin historisches Denken besteht; als theoretische Fundierung inhaltsanalytischer Verfahren ermöglicht es die Komplexitätsreduzierung qualitativ erhobener Daten. So entsteht eine Basis für die Verschränkung mit quantitativen Erhebungsmethoden. Dies verdeutlicht in methodischer Hinsicht der Beitrag von Michael Werner (KU Eichstätt), in inhaltlicher Hinsicht der von Stefanie Zabold (KU Eichstätt). Werner zeigt mit seiner Studie „Das Potential geschlossener Aufgabenformate zur Aktivierung historischen Denkens“, wie kognitive Interviews nutzbar sind, um nachzuweisen, dass geschlossene Items historische Denkprozesse hervorrufen können. Er bezieht sich dabei auf Aufgabenformate der HiTCH-Studie, in der Bildungsforscher der Universität Tübingen und deutschsprachige Geschichtsdidaktiker zusammenarbeiten.

Stefanie Zabold erhebt in ihrer Studie „Ausprägungen historischen Denkens von Grundschulern vor dem ersten Geschichtsunterricht“ zum einen historische Themen, zu denen 9-Jährige sich äußern können, zum anderen dabei vorfindliche Denkstrukturen. Sie kann zeigen, dass Dimensionierungen und Ausprägungen breit streuen, sich aber auch gruppieren lassen. Bei der noch nicht ganz abgeschlossenen Auswertung scheinen sich mehrere „Denker-Typen“ herauszukristallisieren.

Florian Sochatzy (KU Eichstätt)/Martin Merkt (IWM Tübingen) und Christoph Kühberger (PH Salzburg) erforschen Kompetenzentwicklungen anhand historischer Filme. Im ersten Fall untersuchen Psychologen und Geschichtsdidaktiker Training und Cueing als „Instruktionale Maßnahmen zur Unterstützung der Analyse von Propagandafilmen“. Während die Wirksamkeit des digitalen Methodentrainings in allen Experimentalgruppen nachgewiesen werden kann, tragen Cues nur in bestimmten Hinsichten zu verbesserter Leistung bei.

Christoph Kühberger befasst sich mit der „Wahrnehmung des Mediums ‚historischer Spielfilm‘ als Rekonstruktion der Vergangenheit“. Schüleressays zu ausgewählten Filmausschnitten spielen eine zentrale Rolle unter den in einem Methodenmix erhobenen und inhaltsanalytisch ausgewerteten Daten. Repliziert konnten die Schwierigkeiten der Schüler werden, belegbare und fiktionale Momente im Umgang mit Geschichte zu unterscheiden. Ansatzweise gelang es, konzeptionelle Denkstrukturen zu identifizieren.

Das Symposium will zeigen, dass Geschichtsdidaktiker mit ihren Forschungspartnern auf dem Weg sind, die als schwierig zu erschließend geltende Domäne des Historischen empirisch in den Griff zu bekommen, wobei den theoriebasierten Inhaltsanalysen eine Schlüsselfunktion bei der Reduzierung der Komplexität zukommt.

## Das Potential geschlossener Aufgabenformate zur Aktivierung historischen Denkens

Theoretischer Hintergrund: Wie in anderen „schwach strukturierten“ Domänen, gilt der Kompetenzmessung im Fach Geschichte große Skepsis. Einerseits wird wegen der prinzipiellen Perspektivität von Geschichte auf die Problematik verwiesen, es könne jenseits reiner Kenntnisabfrage Antwortverhalten nicht ohne Weiteres als objektiv „richtig“ gewertet werden. Auch die besondere Herausforderung, dass wegen der Dynamik historischen Denkens die Güte und nicht das Ergebnis von Denkprozessen gemessen werden muss, ist der Fachdidaktik bewusst (Körber, 2008). Angesichts dieser Schwierigkeiten einer Messung historischer Kompetenzen, steht in Frage, ob geschlossene Aufgabenformate überhaupt eine hinreichende Anregungsstruktur und Narrativität als Proprium des Historischen aufweisen, um historisches Denken als Voraussetzung für die Kompetenzmessung auszulösen.

Fragestellung: Als Voraussetzung für die quantitative Messbarkeit historischer Kompetenzen, soll zunächst untersucht werden, ob und inwieweit für standardisierte Tests typische Aufgabenformate überhaupt historisches Denken bei den Probanden auslösen. Nur so würden Rückschlüsse auf Kompetenzausprägungen ermöglicht. Können Schüler also auch bei der Arbeit mit geschlossenen Aufgabenformaten methodisch reguliert mit Quellen aus der Vergangenheit und historischen Darstellungen aufgrund gegebener oder selbst gestellter Fragestellungen umgehen und Verbindungen zwischen Vergangenheit, Gegenwart und Zukunft herstellen, kategorial ordnen und auf sich beziehen? In Frage stehen hier quantitative Formate, die im Projekt HiTCH der Universitäten Eichstätt, Hamburg und Tübingen auf Grundlage des Kompetenzstrukturmodells nach FUER (Körber et. al., 2010) zur Messung von Kompetenzen historischen Denkens entwickelt wurden.

Methode: Um das Anregungspotential geschlossener Formate für historisches Denken jenseits einer Expertenbewertung zu prüfen, wurden diese Aufgabenformate aus HiTCH in Cognitive Labs eingesetzt. Derartige Kognitive Interviews haben sich in der Bildungsforschung sowohl für Pretesting und Itemoptimierung (Willis, 2005; Frederking, 2011) als auch für Erkenntnisgewinn über zu messende Gegenstände (Alavi, 2005) bewährt. Anhand der artikulierten Bearbeitungs- und Denkprozesse werden die Aufgaben hinsichtlich ihrer Wirkung beschreibbar. Es wurden insgesamt 12 Testpersonen von je zwei Schülerinnen und zwei Schülern der 9. Klasse in der Hauptschule, der Realschule und am Gymnasium mit einem Durchschnittsalter von 14,75 Jahren in einem durchschnittlich 60-minütigen kognitiven Interview befragt. Als Leitfaden für die komplett mit Audioaufnahmen gesicherten Interviews diente die Fragebogenstruktur eines im Kontext von HiTCH entwickelten Testheftes zum Thema Nürnberger Prozesse mit verschiedenen geschlossenen Formaten. Der Methode des kognitiven Interviews entsprechend wurden die Probanden dazu angehalten, die Aufgabenstellungen und ihre jeweiligen Bearbeitungsschritte zu kommentieren. Zur Ergänzung des lauten Denkens oder als Ausgleich mangelnder Bereitschaft oder Fähigkeit zur Verbalisierung wurden sog. Probes eingesetzt. Diese tastenden Fragen erheben u.a. Aufgabenverstehen und Lösungsstrategien oder das vorgängige Begriffsverständnis und die Begründung der Wahl von Antwortalternativen. Den Prinzipien und Methoden der qualitativen Inhaltsanalyse folgend (Mayring, 2010; Kuckartz, 2012) wurde das vollständig transkribierte und nach Test- und Interviewfragen strukturierte Material codiert. Da dieses die theoretische Grundlage der Fragebogenentwicklung in HiTCH bildet, wurde das Kategoriensystem deduktiv aus dem Kompetenzstrukturmodell nach FUER entwickelt. Mit Hilfe der Codierung von Denkprozessen wie z.B. der kategorialen und zeitlichen Analyse von historischen Fragen, wurden die Schüleräußerungen dahingehend auswertbar, inwieweit durch die Aufgabenstellungen eine Aktivierung einer Performanz historischen Denkens gelingt.

Ergebnisse: Auf Basis der Cognitive Labs lässt sich nachweisen, dass sich allein durch geschlossene Aufgabenformate historisches Denken induzieren lässt. Zudem zeigt das Material Möglichkeiten auf, die geschlossenen Aufgabenformate zu optimieren: Da die Cognitive Labs zusätzlich als Pretesting-Methode angesehen wurden, konnten Verständnisschwierigkeiten aufgrund von unbekanntem Begriffen o.ä. erkannt werden, die ohne Verbindung zu den Fähigkeiten, Fertigkeiten und Bereitschaften historischen Denkens die Aufgabenbearbeitung erschwerten. Dies ist ebenso nutzbar für eine Weiterentwicklung von Aufgabenformaten in HiTCH wie die Analyse und Abstraktion der Anregungs- und Anforderungsstrukturen der im Cognitive Lab erfolgreich erprobten Formate, um Analogie zu diesen weiteren Aufgaben zur Messung historischer Kompetenzen zu konstruieren.

Alavi, S. M. (2005). On the adequacy of verbal protocols in examining an underlying construct of a test. *Studies on educational evaluation* 31 (1), 1–26.

Frederking, V., Roick, T. & Steinhäuser, L. (2011). Literarästhetische Urteilskompetenz - Forschungsansatz und Zwischenergebnisse. In H. Bayrhuber, B. Muszynski, B. Ralle, M. Rothgangel, L-H. Schön, H. J. Vollmer & H-G. Weigand (Hrsg.) *Empirische Fundierung in den Fachdidaktiken* (S.75-94). Münster: Waxmann.

Jabin, T. B., Straf, M. L., Tanur, J. M. & Tourangeau, M. (1984). *Cognitive Aspects of Survey Methodology: Building a Bridge between disciplines*. Washington D.C.: National Academy Press.

Körber, A., Schreiber, W. & Schöner, A. (Hrsg.). (2010). *Kompetenzen historischen Denkens: ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik*. Neuried: ars una Verlagsgesellschaft.

Körber, A. (2008). Sind Kompetenzen historischen Denkens messbar? In V. Frederking (Hrsg.) *Schwer messbare Kompetenzen. Herausforderungen für die empirische Fachdidaktik* (S. 65-84). Baltmannsweiler: Schneider Verlag Hohengehren.

Kuckartz, U. (2012). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim: Beltz Juventa.

Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Beltz Juventa.

Willis, G. B. (2005). *Cognitive Interviewing. A tool for improving questionnaire design*. Thousand Oaks: Sage.



## Ausprägungen historischen Denkens von Grundschulern vor dem Geschichtsunterricht

Nachdem die in Anlehnung an Piaget erwachsene Diskussion der Frage, ob Grundschüler überhaupt historisch denken könnten, überwunden ist, gilt es herauszufinden, wie deren historisches Denken konkret aussieht. In der vorzustellenden Studie wird untersucht, welche Ausprägungen historischer Kompetenzen Lernanfänger mit in die erste Geschichtsstunde bringen. Als theoretische Grundlage dient das Kompetenzstrukturmodell der FUER-Gruppe. Im Schwerpunkt wird die Sachkompetenz mit den Kernkompetenzen Begriffs- und Strukturierungskompetenz untersucht. Zu diesem Kompetenzbereich ist insbesondere bei den Lernanfängern mit unterschiedlichen Kompetenzausprägungen zu rechnen, da die Geschichtskultur als einziger Einflussfaktor auf historisches Denken vor dessen gezielter Förderung im Geschichtsunterricht hier die vielfältigsten Angebote macht. Orientierungs-, Methoden- und Fragekompetenz spielen in der Studie eine untergeordnete Rolle, werden aber berücksichtigt, um den Denkprozess als Ganzes betrachten zu können.

Das Sample sind 9-jähriger Kinder, die jeweils in Bayern die Schule besuchen unter-sucht. Die Auswahl ergibt sich daraus, dass nach dem bayerischen Grundschullehrplan Geschichtsunterricht erstmals in der dritten Klasse vorgesehen ist. Die Studie umfasst 461 Probanden aus 25 dritten Klassen.

Die Erhebung war in zwei Phasen untergliedert, um einerseits zu erfahren, zu welchen histori-schen Themen junge Lerner über Vorstellungen verfügen und andererseits zu untersuchen, wie sich an diesen Themen das historische Denken gestaltet: Die erste Phase der Erhebung betraf immer die Klasse als Ganzes. Dabei wurde ein Verfahren angewandt, das sich an die Methode Gruppendiskussion anlehnt und es ermöglichte, Aussagen über die Themenfelder zu eruieren. In einer zweiten Phase wurden teilstandardisierte Einzelinterviews geführt: Nach vorgegebenen Kriterien wurden jeweils aus den beteiligten Schulklassen Probanden ausgewählt, die in einem ersten Schritt Aussagen zu einem selbst gewählten, historischen Thema trafen. Sie wurden aufgefordert, die Einzelaussagen miteinander zu verknüpfen und argumentativ Strukturen zu bilden. Außerdem sollten sie erklären, woher sie ihre Vorstellungen zu früheren Zeiten haben. Nach dem Nutzen, sich mit vergangenen Zeiten zu beschäftigen wurden sie ebenfalls befragt.

Die Transkripte aller Gruppengespräche und der auf Grundlage der Sample-Struktur ausgewählten 19 Einzelinterviews wurden anhand von Inhalts- und Strukturanalysen computergestützt (MAXQDA) ausgewertet. Grundlage für die Entwicklung des Codebaumes bildet auf der deduzierenden Ebene das FUER-Modell. Dabei konnte auf das Codebuch von Schreiber et al. aus dem Schulbuch-Analyse-Projekt zurückgegriffen werden. Durch induktive Arbeiten am Material wurde der Codebaum weiter ausdifferenziert.

Nach Auswertung der Transkripte der Gruppengespräche mittels Frequenzanalysen zeigt sich, dass die Probandengruppe über Vorstellungen vor allem zu den folgenden historischen Themen verfügt: Nationalsozialismus, Steinzeit, Alte Römer, Alte Ägypter und Ritter.

Auf diesen Themen beziehen sich auch die Einzelinterviews. Hier zeigt sich, dass bezogen auf das historische Denken vor Beginn des Geschichtsunterrichts ein überraschend weites Spektrum an Denkstrukturen identifiziert werden kann. Während manchen Interviews kaum sichtbare „Spuren“ historischen Denkens zu Tage fördern, zeigen andere Kompetenzentwicklungen in allen Bereichen. Die erwarteten Unterschiede in der Sachkompetenz lassen sich als Verfügen können über Konzepte versus Kennen isolierter Begriffe beschreiben. Weniger zu erwarten war, dass sich auch bezogen auf Fragekompetenz klar erkennbare Ausprägungsgrade unterscheiden lassen. Höhere Niveaus sind z.B. durch das aktive Herstellen von Zusammenhängen gekennzeichnet.

- Hasberg, Wolfgang/ Körber, Andreas (2003): Geschichtsbewusstsein dynamisch. In: Körber, Andreas (Hrsg.): Geschichte – Leben – Lernen. Bodo von Borries zum 60. Geburtstag. Schwalbach/ Ts., S. 179-202.
- Körber, Andreas/ Schreiber, Waltraud/ Schöner, Alexander (Hrsg.) (2010): Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Ge-schichtsdidaktik (2. Auflage). Neuried.
- Krieger, Rainer: Mehr Möglichkeiten als Grenzen – Anmerkungen eines Psychologen, in: Klaus Bergmann/ Rita Rohrbach (Hgg.): Kinder entdecken Geschichte. Theorie und Praxis historischen Lernens in der Grundschule und im frühen Geschichtsunterricht, Schwalbach/ Ts. 2001, S. 32-50.
- Kuckartz, Udo (2010): Einführung in die computergestützte Analyse qualitativer Daten (3. Auflage). Wiesbaden.
- Lamnek, Siegfried (2010): Qualitative Sozialforschung. Lehrbuch (5. Auflage). Weinheim/ Basel.
- Mayring, Philipp (2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken (11. Auflage). Weinheim und Basel.
- Schreiber, Waltraud et al. (2012): Kategoriale Inhalts- und Strukturanalysen. Grundlage empirischer Geschichtsdidaktik, erarbeitet an der Analyse von Schulbüchern. Stuttgart.
- Schreiber, Waltraud/ Schöner, Alexander (2008): Teilprojekt Schulbuchanalysen. Richtlinien zur Codierung. Eichstätt.
- Serwuschok, Stefanie (2013): Ausprägungen historischen Denkens vor der ersten Geschichtsstunde im Sachunterricht. Eine empirische Studie. In: Wannack, Evelynne/ Bosshart, Susanne/ Eichenberger, Astrid/ Fuchs, Michael/ Hardegger, Elisabeth/ Marti, Simone (Hrsg.): 4- bis 12-Jährige - ihre schulischen und außerschulischen Lern- und Lebenswelten. Münster, 232-240.
- Serwuschok, Stefanie (2013): Möglichkeiten und Grenzen eines strukturierten Kategorien-Rasters für die Analyse von unsystematischem, historischem Novizen-Denken. In: Schreiber, Waltraud/ Schöner, Alexander/ Sochatzky, Florian: Analyse von Schulbüchern als Grundlage empirischer Geschichtsdidaktik. Stuttgart, S. 171-193.
- Zabold, Stefanie (2013): Auswertungsmethoden und erste Ergebnisse der Studie „Ausprägungen historischen Denkens vor dem Geschichtsunterricht“. In: Hodel, Jan/ Waldis, Monika/ Ziegler, Béatrice (Hrsg.): Forschungswerkstatt Geschichtsdidaktik 12. Beiträge zur Tagung „geschichtsdidaktik empirisch 12“. Bern, S. 166-177.
- Serwuschok, Stefanie (2011): Erhebung zu Ausprägungen historischer Kompetenzen bei Grundschulern vor der unterrichtlichen Intervention im Heimat- und Sachunterricht. In: Hodel, Jan/ Ziegler, Béatrice (Hrsg.): Forschungswerkstatt Geschichtsdidaktik 09. Beiträge zur Tagung „geschichtsdidaktik empirisch 09“. Bern, S. 183-195.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G147



## ***Instruktionale Maßnahmen zur Unterstützung der Analyse von Propagandafilmen***

Filmemachern steht zur Beeinflussung von Zuschauerinnen und Zuschauern (ZuZ) eine Vielzahl von Stilmitteln zur Verfügung. Beispielsweise können Kameraperspektive (Kraft, 1987) oder Distanz der Kamera zum Geschehen (vgl. Williams & Bargh, 2008) die Wahrnehmung beeinflussen. Diese filmischen Mittel wirken meist unterbewusst. Daher ist es problematisch, dass sich filmische Gestaltungsmittel häufig der Aufmerksamkeit entziehen, selbst dann wenn die Probanden im Vorfeld eine weitere Aufgabe zur kritischen Analyse von Filmen durchliefen (Zahn et al., 2010). In der vorliegenden Studie wurde daher auf Cues zurückgegriffen, um die ZuZ auf das Vorliegen eines Stilmittels hinzuweisen. Zwar hat sich Cueing als effektives Mittel zur Lenkung der visuellen Aufmerksamkeit erwiesen (de Koning et al., 2010; Jarodzka et al., 2012), allerdings ist unklar, inwiefern Cues auch dann wirken, wenn sie Aufmerksamkeit nicht auf visueller, sondern auf kognitiver Ebene lenken sollen. Weiterhin sollte das Zusammenspiel der Cues mit einem multimedialen Methodentraining (Schreiber et al., 2012) untersucht werden.

Daher wurde eine Untersuchung mit 81 Neuntklässlern in einem unterrichtsnahen Klassenzimmersetting an zwei bayrischen Gymnasien durchgeführt. Die Schülerinnen und Schüler (SuS) wurden zufällig den vier Untersuchungsbedingungen zugewiesen, die sich aus den Faktoren Training und Cueing ergaben. Nachdem Vorwissen und Interesse erfasst wurden, durchliefen die SuS das Training. Dabei fokussierte das multimediale Training in der Experimentalgruppe auf die Wirkung filmischer Stilmittel, die durch Bilder und Videos veranschaulicht wurden. Das Training in der Kontrollgruppe enthielt Hinweise darauf, dass filmische Mittel zur Manipulation der ZuZ eingesetzt werden, ohne jedoch deren Wirkung zu spezifizieren. Auf die Trainings folgte ein Trainingsfilm, zu dem die SuS drei Fragen zum Erkennen der Stilmittel, zur Wirkung der Stilmittel und zur Intention des Filmemachers beantworteten. Die Fragen werden inhaltsanalytisch ausgewertet. Die Schüler wurden in der Experimentalgruppe durch Symbole (Cues) unterstützt, die oben links eingeblendet wurden, sobald das entsprechende Stilmittel eingesetzt wurde. In der Kontrollbedingung wurden keine Symbole eingeblendet. Auf den Trainingsfilm folgten drei Transferfilme, zu denen dieselben Fragen wie zum Trainingsfilm gestellt wurden. Bei den Transferfilmen wurden grundsätzlich keine Cues eingeblendet.

Es wurde angenommen, dass sowohl das Experimenttraining, als auch die Cues zu einer besseren Identifikation und Interpretation von Stilmitteln und einer besseren Interpretation der Intention der Filmemacher führen. Nachfolgend werden vorläufige Ergebnisse von 21 SuS vorgestellt. Bis zur Tagung liegen die Ergebnisse für alle SuS vor.

Hinsichtlich der Identifikation der Stilmittel zeigte sich für Trainings- und Transferfilme ein positiver Effekt des multimedialen Trainings. Zudem deutete sich ein noch nicht signifikanter negativer Effekt der Cues an. Dieser unerwartete Befund erklärt sich möglicherweise dadurch, dass die Cues zu einem oberflächlichen Aufzählen der Symbole verleiteten, ohne dabei auf die Ausprägung der jeweiligen Stilmittel einzugehen. Diese oberflächliche Strategie bei der Beantwortung der Frage muss jedoch nicht die tatsächliche Identifikationsleistung der SuS widerspiegeln.

Dieser Verdacht erhärtet sich durch die Befunde für die Interpretation der Stilmittel. Hier ergaben sich für Trainings- und Transferfilme positive Effekte der Cues. Vermutlich entlasteten diese die SuS bei der Identifikation der Stilmittel, sodass Ressourcen für deren Interpretation frei wurden. Auch für das multimediale Training deutete sich ein aufgrund der geringen Stichprobengröße noch nicht signifikanter positiver Effekt an.

Für die Interpretation der Intention des Filmemachers erbrachte die vorläufige Analyse eine marginal signifikante Interaktion von Training und Cueing beim Trainingsfilm. Hier profitierten die SuS nur in der Bedingung ohne Cues vom multimedialen Training. Möglicherweise waren die Probanden durch die Cues stärker auf die Interpretation der filmischen Mittel fokussiert und dadurch von der Interpretation der übergeordneten Intention des Filmemachers abgelenkt.

In der Gesamtschau deuten die vorläufigen Ergebnisse der Untersuchung auf einen positiven Effekt des multimedialen Trainings hin, während Cues vor allem bei der lokalen Interpretation der Stilmittel, nicht aber bei der globalen Interpretation der Intention des Filmemachers zu einer Verbesserung der Leistung beitragen.

de Koning, B., Tabbers, H., Rikers, R., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20(2), 111-122. doi:10.1016/j.learninstruc.2009.02.010

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement examples. *Instructional Science*, 40(5), 813-827.

Kraft, R. N. (1987). The influence of camera angle on comprehension and retention of pictorial events. *Memory & Cognition*, 15(4), 291-307. doi:10.3758/BF03197032

Schreiber, W., Schöner & F. Sochatzy, F. (2013). *Analyse von Schulbüchern als Grundlage empirischer Geschichtsdidaktik*. Stuttgart: Kohlhammer.

Williams, L., & Bargh, J. (2008). Keeping one's distance: The influence of spatial distance cues on affect and evaluation. *Psychological Science*, 19(3), 302-308. doi:10.1111/j.1467-9280.2008.02084.x

Zahn, C., Pea, R., Hesse, F. W., & Rosen, J. (2010). Comparing simple and advanced video tools as supports for complex collaborative design processes. *Journal of the Learning Sciences*, 19(3), 403-440.



## **Ein empirischer Zugang zum historischen Denken von Schüler/innen – Spielfilme über die Vergangenheit als Beispiel**

Das geschichtsdidaktische Forschungsprojekt, das aus Mitteln des österreichischen Bundesunterrichtsministeriums finanziert wurde, steht im unmittelbaren Zusammenhang mit der Debatte zur domänenspezifischen Kompetenzorientierung im österreichischen Unterrichtsfach „Geschichte und Sozialkunde/ Politische Bildung“ sowie in der internationalen und deutschsprachigen Geschichtsdidaktik. Im Zentrum der empirischen Studie steht die Frage nach den bei SchülerInnen der 7. Schulstufe (n=260) verfügbaren historischen Teilkompetenzen in Bezug auf die Wahrnehmung des Mediums „historischer Spielfilm“ als Rekonstruktion der Vergangenheit, wobei Jörn Rüsen's Überlegungen zur Historik als geschichtstheoretische Grundlagen dienen.

Die forschungsleitenden Fragen lauteten:

Welche Durchdringungstiefe der Unterscheidung zwischen „Geschichte“ und „Vergangenheit“ weisen Schüler/innen im Bezug auf einen historischen Spielfilm auf?

Anhand welcher Beispiele nehmen Schüler/innen – in diesem Zusammenhang – in ihren Begründungen Triftigkeitsprüfungen der historischen Darstellung vor?

Nehmen Schüler/innen die Normativität der vorgeführten historischen Darstellung und ihre Auswirkungen auf die Erzählung der Vergangenheit in ausreichendem Maße wahr? Welche Aspekte des Mediums „Spielfilm“ werden von den Schüler/innen als konstituierend in Hinblick auf die darin umgesetzte historische (Re-)Konstruktion aktiv wahrgenommen?

Der Untersuchungsaufbau berücksichtigte Daten verschiedener Provenienz, die aus mehreren Quellen des Untersuchungssettings stammten, um sie im Sinn einer Data-Triangulation für die Auswertung heranzuziehen: a) Essays der Schüler/innen; b) Interviews mit Schüler/innen; c) Befragung der Geschichtslehrer/innen der Klassenverbände; d) Analyse des Filmausschnittes. Ziel war es dabei, unterschiedliche Aspekte und Einflussfaktoren desselben Phänomens zu erfassen, um über deren Berücksichtigung eine pluriperspektivische Erfassung, Beschreibung und Erklärung der erhobenen historischen Denkstrukturen der Schüler/innen zu ermöglichen. Um bei der Analyse der einzelnen Datensätze subjektive Sichtweisen in Interpretationen zu erweitern, zu korrigieren bzw. zu überprüfen, kam bei der Auswertung zudem eine Investigator-Triangulation zum Einsatz.

Die dazu gewählten Erhebungsinstrumente versuchten daher einerseits besonders qualitative Aspekte in den Vordergrund zu stellen, andererseits Datenmaterial, das sowohl die schulische als auch die Erhebungssituation kontextualisiert. In der empirischen Phase der Studie wurden daher qualitative und an sie gebundene quantitative Erhebungsverfahren miteinander kombiniert, indem nämlich Schüler/innen nach dem Vorführen eines Filmausschnittes entlang einer vorgegebenen Aufgabenstellung je ein Essay aufgrund einer Impulsfrage verfassten, welches in einigen Fällen durch vertiefende Interviews mit ausgewählten Lernenden ergänzt wurde. Zudem wurde eine Befragung der Geschichtslehrer/innen der Schüler/innen hinsichtlich deren kritischen Beschäftigung mit Filmen im Unterricht der befragten Klassen durchgeführt, um auch eine Rückbindung an schulische Lernerfahrungen der Proband/innen zu erreichen. In Bezug auf den gezeigten Filmausschnitt wurde eine detaillierte Filmanalyse zu den zentralen Einflussgrößen des Mediums vorgenommen.

Um die Schüleressays auswerten zu können, wurde auf eine qualitative Inhaltsanalyse zurückgegriffen, die zum einen Teil in Anlehnung an Philipp Mayring entwickelt wurde, sich zum anderen an Konzeptionen von Jochen Gläser/Grit Laudel anlehnt, um die für die Forschungsfrage relevanten Aspekte erschließen zu können. Auf diese Weise konnte in dem verwendeten Erhebungsraster zur Analyse der Essays auch die methodische Herangehensweise des Concept-Mappings nach Judith Torney-Purta integriert werden, die ausgehend von den kognitionspsychologischen Konzepten der mentalen Modelle die konzeptionellen Denkstrukturen zu fassen sucht. Es handelt sich dabei in Summe um eine Methode, um kognitive Verstehensakte und ihre gedankliche Komplexität sichtbar zu machen.

Die Ergebnisse der Untersuchung zeigen einmal mehr ein heterogenes Bild der Schüler/innen im Zusammenhang mit historischen Denkleistungen und vor allem hinsichtlich der Unterscheidung von belegbaren Momenten und fiktionalen Anteilen im Umgang mit Geschichte. Besonders interessant sind dabei aber auch die signifikanten Zusammenhänge zwischen Mediennutzungsgewohnheiten und der inhaltlichen Argumentation, sowie die Einschätzung der Lehrpersonen hinsichtlich der erwarteten Schülerleistungen und der tatsächlich erbrachten.

Gläser, Jochen/ Laudel, Grit: Experteninterviews und qualitative Inhaltsanalyse. Weisbaden 2010.

Kühberger, Christoph (Hg.). Geschichte denken. Zum Umgang mit Geschichte und Vergangenheit von Schüler/innen der Sekundarstufe I am Beispiel „Spielfilm“. Empirische Befunde – Diagnostische Tools – Methodische Hinweise. Innsbruck – Wien 2013.

Kühberger, Christoph: Diagnose als Herausforderung. In: Geschichtsunterricht. De-Konstruktion in der Schulpraxis. In: 1968 – Vorgeschichte – Folgen. Bestandaufnahmen der österreichischen Zeitgeschichte. 7. Österreichischer Zeitgeschichtetag 2008. Hg. v. I. Böhler/E. Pfanzelter/Th. Spielbüchler/R. Steininger. Innsbruck – Wien 2010. S. 639–646.

Mayring, Philipp: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim 2003.

Rüsen, Jörn: Historische Vernunft. Göttingen 1983.

Turne-Purta, Judith: Adolescents' Schemata of Society, Self and Others. An Approach to Conceptualizing and Measuring Adolescents' Understanding of Public Good. (Paper presented at symposium on Adolescents' Conceptions of the Public and Private Good, Society for Research in Child Development, Kansas City, Missouri, April 1989).



Chair(s): Ilonca Hardy (Universität Frankfurt a.M.)

Diskutant/in: Thilo Kleickmann (IPN Kiel)

Es ist eine wichtige aktuelle Aufgabe für die vergleichsweise junge Disziplin „Didaktik des Sachunterrichts“, Erkenntnisse der Lehr-Lernpsychologie und der Unterrichtsforschung mit traditionellen und bewährten didaktischen Ansätzen des Sachunterrichts zu verbinden (vgl. in diesem Zusammenhang z.B. Lohrmann, Hartinger & Schwelle, 2013). In diesem Symposium werden die Ergebnisse aus vier Projekten zum naturwissenschaftlichen Sachunterricht vorgestellt, die ihre jeweilige Fragestellung und Perspektive mit Blick auf diese Vernetzung entwickelt haben.

Der inhaltlich verbindende Kern ist die Frage nach kognitiv anspruchsvollen Lernprozessen bzw. nach komplexen und anschlussfähigen Konzepten, die anspruchsvolle Abstrahierungsleistungen erfordern. In den verschiedenen, überwiegend DFG-geförderten Forschungsvorhaben wird in Labor- bzw. Interventionsstudien mit (quasi-)experimentellen Designs untersucht, wie Vorschul- und Grundschulkindern beim Aufbau naturwissenschaftlicher Konzepte unterstützt werden können. Dies geschieht anhand unterschiedlicher physikalischer Inhalte im naturwissenschaftlichen Sachunterricht. In den Vorträgen werden Theorie, Design sowie Ergebnisse der jeweiligen Studien vorgestellt. Die Studie von Haider, Keck, Haider und Fölling-Albers beschäftigt sich mit der Frage, inwieweit Schüler(innen) in der Lage sind, Analogiemodelle zu nutzen, um Stromverbrauchsvorstellungen abzubauen und ein stabiles Konzept des (nicht wahrnehmbaren) Stromkreislaufflusses zu entwickeln. Ebenfalls um Analogien geht es in der Studie von Schwelle, Lohrmann und Hartinger. In dieser geschieht die Übertragung der kognitionspsychologisch entwickelten Theorie des structural alignment auf die Didaktik des Sachunterrichts. Dabei steht die Frage im Mittelpunkt, inwieweit es für Schüler(innen) (mit verschiedenen Vorkenntnissen) günstiger ist, anhand von Phänomenen, die auf einer Oberflächenstruktur ähnlich oder unähnlich sind, zu lernen, um ein Verständnis der zugrundeliegenden naturwissenschaftlichen Gesetzmäßigkeit (hier das Hebelgesetz) zu entwickeln. Bei beiden Studien sind Analogien im Fokus und damit die Anforderung für die Schüler(innen), Verbindungen zwischen diesen zu ziehen durch schlussfolgerndes Denken. Dieses ist gezielt im Fokus der Studie von Robisch, Tröbst und Möller. Hier wird im mikrogenetischen Design überprüft, inwieweit verschiedene Scaffolding-Maßnahmen geeignet sind, Kinder bei der adäquaten Beurteilung unterschiedlicher Versuchsausgänge für die Gültigkeit allgemein formulierter Vermutungen zu unterstützen. Die Versuche und Vermutungen beziehen sich hierbei auf die Frage „Wie kommt es, dass ein Ball springt?“. Die Untersuchung von Ranger, Martschinke und Kopp erweitert die Fragestellung, indem neben kognitiven Scaffolds auch noch soziale Aktivierungsmaßnahmen eingeführt und auf ihre Wirksamkeit überprüft werden.

Das Symposium endet mit einer abschließenden Diskussion aller Beiträge durch Thilo Kleickmann, der in seinen Forschungsarbeiten Fragen des (frühen) naturwissenschaftlichen Lernens eng mit lehr-lernpsychologischen Ansätzen und Theorien vernetzt und somit die beiden das Symposium bestimmenden Perspektiven gut verbinden kann.

Michael Haider (Universität Erlangen-Nürnberg), Marika Keck (Universität Regensburg), Thomas Haider (Universität Regensburg), Maria Fölling-Albers (Universität Regensburg)

## **Lernen mit Modellen im Sachunterricht am Beispiel elektrischer Strom**

Der Aufbau belastbarer, flexibel anwendbarer Konzepte zum Thema elektrischer Strom ist aus didaktischer Sicht sehr schwierig, weil diese eine anspruchsvolle theoretische Modellierung erfordern. Zudem werden hier Begriffe in der Alltagssprache und Fachsprache nicht nur unterschiedlich verwendet, vielmehr werden (aus fachlicher Sicht) durch die Alltagssprache falsche Vorstellungen aufgebaut (z.B. Stromverbrauch). Somit sind im Unterricht neben einer Klärung von Fachbegriffen inadäquate Begriffe bzw. falsche Konzepte (hier z.B. Einwegzuführungsvorstellung) abzubauen und fachlich angemessene Vorstellungen aufzubauen (Konzeptwechsel) (u.a. Caravita, 2001). Analogiebildung kann als didaktisches Mittel zur Veranschaulichung von unanschaulichen und komplexen Inhalten genutzt werden. Analogien haben dann eine Mittler- bzw. eine Brückenfunktion zwischen „primärem“ und „sekundärem Lernbereich“ (vgl. u.a. Gentner 1988; Duit, Roth, Komorek & Wilbers, 2001). Vom analogen Denken wird ein vertieftes Verständnis über komplexe Sachverhalte (Kurtz, Miao, & Gentner, 2001) erwartet. Gentner, Loewenstein und Thompson (2003) betonen die Bedeutung des Vergleichs für den Lernprozess, d.h. dass der Lerngegenstand aus verschiedenen Perspektiven betrachtet werden kann. (Unterschiedliche) Analogiemodelle bieten die Möglichkeit, den Lerngegenstand unter verschiedenen Perspektiven wahrzunehmen. Anhand der (konkreten) Modelle kann auf spezifische, jeweils als lernrelevant erachtete Merkmale des primären Lerngegenstandes (Oberflächenanalogien und/oder strukturelle Analogien) fokussiert werden. Auf diese Weise kann die Komplexität des Lerngegenstandes reduziert werden; durch den (anschließenden) Vergleich mit dem primären Lerngegenstand kann dieser in einer für das Lernziel relevanten (neuen) Perspektive gesehen und interpretiert werden.

In der Studie stand die Frage im Fokus, ob und inwiefern bereits Grundschüler(innen) Analogiemodelle für den Aufbau flexibel anwendbarer und komplexer Konzepte nutzen (können). Dabei wurden die Effekte des Einsatzes von Analogiemodellen untersucht, und es wurde erhoben, ob und wie die Schüler(innen) diese Analogiemodelle für ihre Verstehensprozesse nutzen.

Die Forschungsfragen wurden anhand des Lehrplanthemas „Strom“ (3. Jahrgangsstufe; Unterrichtsziele: Aufbau eines Stromkreisbegriffs und Abbau von Stromverbrauchsvorstellungen) untersucht, da es sich hierbei um abstrakte und komplexe Konzepte handelt, die in der Regel den Alltagsvorstellungen widersprechen. Im Forschungsvorhaben wurden im Rahmen eines quasi-experimentellen Designs zwei Experimentalgruppen (je n = 4 Klassen), die mit verschiedenen Analogiemodellen (je zwei Wasseranalogiemodelle und zwei mechanische Modelle) arbeiteten, einer Kontrollgruppe ohne Einsatz von Analogiemodellen (4 Klassen) gegenüber gestellt. Die Modelle sollen jeweils weitergehende Perspektiven auf den Lerngegenstand ermöglichen. Einerseits unterscheiden sie sich äußerlich deutlich vom primären Lerngegenstand (Stromkreis) – alle Elemente sind von anderer Qualität als die Elemente des Stromkreises (z.B. Schläuche statt Kabel, Doppelwassersäule statt Batterie) –, andererseits haben sie entsprechende Funktionen, und anhand der Modelle können Prozesse sichtbar gemacht werden, die am primären Lerngegenstand nicht beobachtbar sind (Strömen des Wassers bzw. die Bewegung des Riemens – im Unterschied zum Elektronenfluss). Die Untersuchung der Frage, wie die Schüler(innen) die Analogiemodelle nutzen, erfolgte an ausgewählten Schülergruppen anhand qualitativer Verfahren (Interviews).

Im Vortrag werden der theoretische Hintergrund, der Aufbau der Studie und erste Ergebnisse vorgestellt, so z.B. Varianzanalysen zur unterschiedlichen Wirkung von Analogiemodellen in den Experimentalgruppen im Vergleich zur Kontrollgruppe. Verglichen werden die Daten zum Aufbau des Stromkreisbegriffs und zum Abbau von Verbrauchsvorstellungen. Es zeigt sich, dass die Intervention in allen Gruppen erfolgreich verlief, dass es tendenzielle Unterschiede zwischen den Gruppen gibt, die jedoch nicht signifikant werden. Ergebnisse der qualitativen Inhaltsanalyse der Interviews konkretisieren die durch die Intervention hervorgerufenen Vorstellungen und Konzepte in den jeweiligen Gruppen. Darüber hinaus geben differentielle Untersuchungen Aufschluss über etwaige Subgruppenunterschiede. So profitieren einzelne Schülergruppen (z.B. Schüler der mittleren Leistungsgruppe) in Teilbereichen der Fragestellung (z.B. beim Abbau der Verbrauchsvorstellung) besonders von den Wasseranalogiemodellen.

Caravita, S. (2001). A re-framed conceptual change theory? – Commentary. *Learning and Instruction*, 11, 421-449.

Duit, R., Roth, W.-M., Komorek, M. & Wilbers, J. (2001). Fostering Conceptual Change by Analogies – Between Scylla and Carybdis. *Learning and Instruction*, 36 (4-5), 283-303.

Gentner, D. (1988). Analogical inference and analogical access. In A. Prieditis. (Ed.), *Analogica* (pp.63-88). Los Altos: Morgan Kaufmann Publishers.

Gentner, D., Loewenstein, J. & Thompson, L. (2003). Learning and Transfer: A General Role for Analogical Encoding. *Journal of Educational Psychology*, 95 (2), 393-408.

Kurtz, K.J., Miao, Ch.-H., & Gentner, D. (2001). Learning by Analogical Bootstrapping. *The Journal of the Learning Sciences*, 10 (4), 414-446.



## **Kognitive Aktivierung im naturwissenschaftlichen Sachunterricht in der Grundschule**

Innerhalb des Forschungsprojekts „Kognitive Aktivierung in kooperativen Lernphasen des naturwissenschaftlichen Sachunterrichts in der Grundschule“ beschäftigt sich die vorgestellte Studie mit der Fragestellung, ob Kinder eingeführte Maßnahmen zur kognitiven und sozialen Aktivierung tatsächlich wahrnehmen und umsetzen können. Dazu wurde im Rahmen einer explorativen Interventionsstudie eine neunstündige Unterrichtssequenz zum Thema Magnetismus in der dritten Jahrgangsstufe mit je einer kooperativen Lernphase in jeder Unterrichtseinheit durchgeführt.

Speziell aus soziokonstruktivistischer Perspektive (in Anlehnung an Piaget, bei Renkl die sogenannte neo-piagetsche Perspektive) geht man davon aus, dass im sozialen Austausch soziokognitive Konflikte und Perturbationen ausgehandelt werden. Das Aufeinandertreffen verschiedener Perspektiven bei Interaktionen von Peers gleicher Altersstufe wird als besonders gewinnbringend angesehen, weil sich dabei die Lernenden auf ähnlichem Sprachniveau bewegen, aber meist unterschiedlich kompetent sind. Damit ist keine unkritische oberflächliche Perspektivenübernahme zu erwarten, sondern ein gemeinsames Aushandeln und Konvergieren des Wissens. Die Aufgabe des Lehrers sollte es sein, den Schülern herausfordernde Aufgaben zu stellen, die kognitive Konflikte provozieren und dazu führen, dass die Kinder ihr Vorwissen, ihre Ideen und Lösungswege verbalisieren und im Dialog und Austausch mit den anderen diskutieren (Klieme, Lipowsky, Rakoczy, & Ratzka, 2006; Lipowsky, 2006, 2009). Man vermutet, dass dadurch vertieftes Nachdenken, reiches Elaborieren sowie Vernetzen mit dem Vorwissen stattfinden können.

In der vorgestellten Hauptstudie erhielten 93 Schüler(innen) im Rahmen des Unterrichts sowohl Lernvorgaben zur kognitiven als auch zur sozialen Aktivierung, um die Gruppenarbeitsphasen zu intensivieren. Zur kognitiven Aktivierung wurden Gruppenarbeitsaufträge, ein Forschertagebuch und Lerntipps eingesetzt. Die soziale Aktivierung wurde durch ein spezifisches Kooperationskript mit wechselndem Laborchef gesteuert. Zur Erfassung des Lernerfolgs wurde sowohl vor als auch nach der Intervention ein Wissenstest durchgeführt. Zusätzlich wurden prozessbegleitend Daten über Videographie und Forschertagebücher gewonnen, um die Interaktionen während der kooperativen Lernphasen und damit Qualität und Quantität der kognitiven Aktivierung der Schüler(innen) zu erfassen.

Der Schwerpunkt des Vortrages liegt auf der kognitiven Aktivierung und den Möglichkeiten, diese Aktivierung der Schüler(innen) quantitativ und qualitativ zu erfassen. Nach der Beschreibung des Designs und der eingesetzten Erhebungsinstrumente (Wissenstest, Einschätzungsskala zur kognitiven Aktivierung, Videographie) wird auf die Ergebnisse der Hauptstudie eingegangen: Die Auswertung der quantitativen Daten zeigt, dass sich in der Gesamtstichprobe ein positiver Wissenszuwachs mit einer sehr hohen Effektstärke beobachten lässt. Dabei finden sich keine signifikanten Klassen- oder Gruppenunterschiede. Des Weiteren schätzen sich die Schüler(innen) selbst kognitiv hoch aktiviert in den einzelnen kooperativen Lernphasen ein. Die qualitativen Daten aus der Videoanalysen ergeben ein ähnliches Bild: Das Ergebnis des time-samplings auf Grundlage des Münchner Aufmerksamkeitsinventars (Helmke 1988) verweist auf ein hohes on-task Verhalten der einzelnen Gruppenmitglieder, d.h. die Kinder bringen sich aktiv in die Diskussionen ein und scheinen aufmerksam zuzuhören. Um herauszufinden, worüber die Kinder genau miteinander in der Gruppe sprechen, wurde in induktiver Vorgehensweise in Anlehnung an Mayring (2000) ein Kategoriensystem zu den Gesprächsinhalten gebildet. Mithilfe eines event-samplings wurden alle Redebeiträge der Kinder ausgewertet. Auch hier zeigt sich, dass sie sich intensiv mit den Inhalten auseinandersetzen. Hierbei nehmen Aussagen, die sich beispielweise damit beschäftigen, eigene Ideen einzubringen, schlussfolgerndes Denken aufzeigen und die kritische Auseinandersetzung mit den Ideen der anderen Gruppenmitgliedern widerspiegeln, einen großen Anteil ein. Alle Ergebnisse (Aufmerksamkeitsverhalten, inhaltliche Prozesse, Selbsteinschätzung) verweisen auf reichhaltiges Elaborieren in den kooperativen Lernphasen.

Helmke, A. (1988). Das Münchner Aufmerksamkeitsinventar (MAI): Manual für die Beobachtung des Aufmerksamkeitsverhaltens von Grundschulern während des Unterrichts (Paper 6). München: Max-Planck-Institut für Psychologische Forschung.

Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel & L. Allolio-Näcke (Hrsg.), Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms (S. 127-146). Münster: Waxmann.

Lipowsky, F. (2006). Auf den Lehrer kommt es an. In C. Allemann-Ghionda & E. Terhart (Hrsg.), Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf (S. 47-70). Zeitschrift für Pädagogik, 51. Bei-heft. Weinheim: Beltz.

Lipowsky, F. (2009). Unterricht. In E. Wild & J. Möller (Hrsg.), Pädagogische Psychologie (S.74-101 ). Berlin: Springer.

Mayring, Philipp (2000): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim: Deutscher Studien Verlag.

Renkl, A. (1997). Lernen durch Lehren: Zentrale Wirkmechanismen beim kooperativen Lernen. Wiesbaden: Deutsche Universitäts-Verlag.

## Wie können adäquate Schlussfolgerungen im scientific inquiry-orientierten Sachunterricht gefördert werden?

Adäquate Schlussfolgerungen gelten kognitionspsychologisch als bedeutsam für die Prüfung von Hypothesen (Moshman, 1979). Die fokussierte Schlussfolgerungsform ist definiert als die Evaluation eines gegebenen Ereignisses in Bezug auf die Bedeutung für die Prüfung des Wahrheitsgehaltes einer Hypothese (Vergauwe, Gauffroy, Mosanyi, Dagry, & Barrouillet, 2013) - Was zeigt ein Ereignis über die Gültigkeit der Hypothese? Schlussfolgerungen sind dann adäquat, wenn bekräftigende, widerlegende und für die Prüfung irrelevante Ereignisse korrekt unterschieden werden. Die korrekte Unterscheidung wird vor dem Hintergrund der dual-process mental models theory des konditionalen Schlussfolgerns mit der Bildung der erforderlichen mentalen Modelle im Arbeitsgedächtnis erklärt (Gauffroy & Barrouillet, 2009). Sowohl kognitionspsychologische (Barrouillet, Gauffroy, & Lecas, 2008; Gauffroy & Barrouillet, 2011; English, 1997) als auch fachdidaktische (Tröbst, Hardy, & Möller, 2011) Befunde weisen darauf hin, dass insbesondere die angemessene Evaluation irrelevanter Ereignisse eine Herausforderung (nicht nur) für Grundschulkindern ist.

Aus fachdidaktischer Perspektive wird die Prüfung von Hypothesen bereits in der Grundschule, als ein Prozess (natur-) wissenschaftlicher Denk- und Arbeitsweisen (scientific inquiry), gefordert (OECD, 2007). Bisherige Studien zur Förderung der Schlussfolgerungsfähigkeit haben sich primär auf Erwachsene beschränkt und naturwissenschaftliche Kontexte vernachlässigt (z.B. English, 1997; Klauer, Stegmaier, & Meiser, 1997). Unter Berücksichtigung der für angemessene Schlüsse als relevant identifizierten Faktoren – Arbeitsgedächtnis (Klauer et al., 1997), Inhibitionsfähigkeit (Handley, Capon, Beveridge, Dennis & Evans, 2004), Inhalt und Kontext (Thompson, 1994) sowie Wissenschaftsverständnis (Hardy, Stephan-Gramberg, Robisch, Tröbst, & Möller, 2013) – wurden daher zwei Fördermaßnahmen entwickelt, um adäquate Schlussfolgerungen bei Grundschulkindern in einem naturwissenschaftlichen Kontext zu fördern.

Im Rahmen einer experimentellen Laborstudie im mikrogenetischen Design (Siegler & Crowley, 1991) mit N = 101 DrittklässlerInnen wurde anhand des exemplarischen Themas „Wie kommt es, dass ein Ball springt?“ (Thiel, 1987) untersucht, inwiefern (1) prozessbegleitende Scaffolds mit unterschiedlichem Unterstützungsgrad (KG = keine Scaffolds, EG 1 = aufmerksamkeitslenkende Scaffolds, EG 2 = aufmerksamkeitslenkende und modellierende Scaffolds) angemessene Schlussfolgerungen fördern und (2) ob bei einer Rücknahme (Fading) der Unterstützung der etwaige Fördereffekt erhalten bleibt. Das Experiment wurde durch eine Sachbeihilfe (Kennz.: MO 941/5-1) der DFG ermöglicht. Die varianzanalytische Auswertung zeigte einen (1) signifikanten Haupteffekt des Faktors Bedingung. Kontrastanalysen spezifizierten den Befund, indem eine signifikante Überlegenheit der beiden EGs gegenüber der KG hinsichtlich der angemessenen Evaluation von Ereignissen nachgewiesen wurde, während sich die EG 1 und die EG 2 nicht überzufällig unterschieden. Weiterhin wurde (2) der erwartete Haupteffekt des Faktors Bedingung, bei vollkommener Rücknahme der Unterstützung, nachgewiesen. Kontrastanalysen ergaben, dass die zuvor geförderten EGs der nicht-geförderten KG auch ohne Unterstützung hinsichtlich angemessener Schlussfolgerungen signifikant überlegen waren. Überzufällige Unterschiede zwischen den Experimentalgruppen wurden nicht gefunden.

Eine Förderung angemessener Schlussfolgerungen im Kontext „Wie kommt es, dass ein Ball springt?“ ist folglich bereits bei Drittklässlern möglich. In weiteren Studien sollen die Fördermaßnahmen in einen Unterricht für ganz Schulklassen implementiert werden. Das Modeling, welches in der zweiten Experimentalgruppe über die aufmerksamkeitslenkenden Scaffolds hinaus angeboten wurde, kann dabei, je nach favorisierter Unterrichtsstruktur, flexibel in den Unterricht eingebettet werden, da es über die aufmerksamkeitslenkenden Scaffolds hinaus zwar keinen positiven, aber auch keinen negativen Effekt auf angemessene Schlussfolgerungen gab. Anschließend sollen die Forschungsbefunde in Handreichungen für Lehrkräfte überführt werden, um die Umsetzung eines inquiry-orientierten Sachunterrichts in der Grundschule zu unterstützen.

- Barrouillet, P., Gauffroy, C., & Lecas, J. F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review*, 115, 760–772.
- English, L.D. (1997). Interventions in children's deductive reasoning with indeterminate problems. *Contemporary Educational Psychology*, 22, 338-362.
- Gauffroy, C., & Barrouillet, P. (2011). The primacy of thinking about possibilities in the development of reasoning. *Developmental Psychology*, 47, 1000-1011.
- Gauffroy, C., & Barrouillet, P. (2009). Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory. *Developmental Review*, 29(4), 249-282.
- Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking and Reasoning*, 10(2), S. 175-195.
- Hardy, I., Stephan-Gramberg, S., Robisch, C., Tröbst, S., Möller, K. (2013). Die Schlussfolgerungsfähigkeit im naturwissenschaftlichen Kontext – Ergebnisse einer Querschnittstudie aus dem Elementar- und Primarbereich. Vortrag auf der 22. Jahrestagung der Gesellschaft für Didaktik des Sachunterrichts e.V., Solothurn, Schweiz.
- Klauer, K.C., Stegmaier, R., & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning*, 3, 9-47.
- Moshman, D. (1979). Development of formal hypothesis-testing ability. *Developmental Psychology*, 15, 104-111.
- OECD. (2007). PISA 2006. Schulleistungen im internationalen Vergleich: Naturwissenschaftliche Kompetenzen für die Welt von Morgen. Paris: OECD.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method. A direct means for studying cognitive development. *American Psychologist*, 46(5), S. 606-620.
- Thiel, S. (1987). Wie springt ein Ball? Grundschule, 1, 18–23.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22(6), S. 742-758.
- Tröbst, S., Hardy, I., & Möller, K. (2011). Die Förderung deduktiver Schlussfolgerungen bei Grundschulkindern in naturwissenschaftlichen Kontexten. *Unterrichtswissenschaft*, 39, 7-20.
- Vergauwe, E., Gauffroy, C., Morsanyi, K., Dagry, I., & Barrouillet, P. (2013). Chronometric evidence for the dual-process mental models theory of conditional. *Journal of Cognitive Psychology*, 25(2), S. 174-182.



## ***Vorwissen und kognitive Grundfähigkeiten als Einflussfaktoren auf den Aufbau von konzeptuellen Wissensstrukturen***

Die Gestaltung von Lernumgebungen zielt darauf, insbesondere den Aufbau von konzeptuellem Wissen zu ermöglichen (BLK/ Prenzel et al., 2004). Bezogen auf den naturwissenschaftlichen Sachunterricht der Grundschule wird hierfür die Auseinandersetzung mit Konkretionen, also naturwissenschaftlichen Phänomenen, als besonders hilfreich angesehen (GDSU, 2013). Forschungsbefunde konnten diesbezüglich zeigen, dass es dabei förderlich ist, wenn mit multiplen Konkretionen, denen eine gemeinsame Tiefenstruktur zugrunde liegt, gearbeitet wird (u.a. Gentner, Loewenstein, & Thompson, 2003; Sagi, Gentner, & Lovett, 2012) und die Konkretionen dabei miteinander verglichen werden (structure mapping; Gentner et al., 2003). Hinsichtlich der Frage, auf Basis welcher Kriterien die Konkretionen ausgewählt werden sollen, um den Lernprozess möglichst positiv anzuregen, liefern theoretische Überlegungen und empirische Befunde ein ambivalentes Bild: Einerseits sprechen sich Autoren für das Arbeiten mit oberflächlich ähnlichen Phänomenen eines Funktionsprinzips aus. Argumentiert wird hierbei, dass gerade Novizen bezogen auf einen Inhaltsbereich ansonsten überfordert seien (Chi, Feltovich, & Glaser, 1981; Vosniadou, 1989). Für das Arbeiten mit ähnlichen Phänomenen sprechen zudem Forschungsergebnisse, die zeigen, dass die Novizen ihren Blick eher auf die Oberflächenstruktur von Phänomenen richten, da ihnen die Tiefenstruktur noch unbekannt ist (Blanchette & Dunbar, 2001). Andererseits plädieren Autoren für das Arbeiten mit oberflächlich unähnlichen Phänomenen eines Funktionsprinzips, um so auf die Tiefenstruktur der Phänomene, d.h. auf das gemeinsame Funktionsprinzip, zu fokussieren. Eine ähnliche Oberflächenstruktur der Phänomene würde demnach den Blick auf die Tiefenstruktur verstellen (Gentner et al., 2003; Kurtz, Miao, & Gentner, 2001).

In dem DFG-geförderten Projekt „Phänomen(un)ähnlichkeiten“ wurden im Rahmen einer quasi-experimentell angelegten Interventionsstudie deshalb zwei didaktische Settings, das Arbeiten mit ähnlichen und das Arbeiten mit unähnlichen Phänomenen, in 22 dritten Klassen (N=397) untersucht. Im Mittelpunkt der Studie steht die Frage, ob unähnliche Phänomene stärker dazu beitragen, dass der Blick der Kinder auf die zugrunde liegende Gesetzmäßigkeit gerichtet wird. Dabei stellt sich im Hinblick auf ATI-Effekte die Frage, ob diese anspruchsvollen kognitiven Prozesse möglicherweise nur Schüler(innen) mit entsprechenden Voraussetzungen leisten können: Vermutet wird, dass Kinder mit hoch ausgeprägten kognitiven Grundfähigkeiten und hohem inhaltspezifischem Vorwissen beim Arbeiten mit unähnlichen Phänomenen einen höheren Lernzuwachs und bessere Fähigkeiten, das erworbene Wissen auf neue Phänomene anzuwenden, erlangen als Kinder mit gering ausgeprägten kognitiven Grundfähigkeiten und wenig inhaltspezifischem Vorwissen. Bei den Schüler(inne)n, die sich mit oberflächlich ähnlichen Phänomenen auseinandersetzen, werden diese Unterschiede als weniger stark ausgeprägt eingeschätzt.

Untersucht wurden daher mittels hierarchischer Regressionsanalysen in Mplus (Muthén & Muthén, 2010) Wechselwirkungseffekte zwischen den beiden didaktischen Settings und den individuellen Schülervoraussetzungen (kognitive Grundfähigkeiten, Vorwissen).

Im Vortrag werden Ergebnisse des Projekts präsentiert, die zeigen, dass sich sowohl die kognitiven Grundfähigkeiten als auch das Vorwissen signifikant auswirken, wobei das Vorwissen größere Effekte hervorbringt als die kognitiven Grundfähigkeiten. Des Weiteren werden gruppenspezifische Analysen vorgestellt, die zeigen, dass es zwar tendenziell Unterschiede zwischen den Versuchsgruppen gibt, welche jedoch nicht signifikant werden. Festgehalten werden kann jedoch, dass erwartungswidrig in der Treatmentgruppe „ähnlich“ über alle Messzeitpunkte hinweg die Bedeutung des Vorwissens größer war als in der Treatmentgruppe „unähnlich“.

Bezogen auf die kognitiven Grundfähigkeiten zeigt sich diese Tendenz nicht. Die Daten lassen allerdings den Schluss zu, dass, hypothesenkonform, bezogen auf das konzeptuelle Wissen ein etwas größerer Einfluss in der Treatmentgruppe „unähnlich“ festgestellt werden kann.

Blanchette, I. & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory & Cognition* 29 (5), 730–735.

BLK (Bund-Länder-Kommission)/ Prenzel, M. et al. (2004). Sinus – Transfer GS. Weiterentwicklung des mathematischen und naturwissenschaftlichen Unterrichts an Grundschulen (Heft 112). Kiel: IPN.

Chi, M., Feltovich, P. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5 (2), 121–152.

GDSU (Gesellschaft für Didaktik des Sachunterrichts) (2013). Perspektivrahmen Sachunterricht. Bad Heilbrunn: Klinkhardt.

Gentner, D., Loewenstein, J. & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology* 95 (2), 393–405.

Kurtz, K. J., Miao, C.-H. & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences* 10 (4), 417–446.

Muthén, L. & Muthén, B. (1998-2010). *Mplus user's guide* (sixth edition). Los Angeles: Muthén & Muthén.

Sagi, E., Gentner, D. & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science* 36 (6), 1019–1050.

Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition: A developmental perspective. In S. Vosniadou & A. Ortony (Hrsg.), *Similarity and analogical reasoning* (S. 413–437). New York: Cambridge University Press.



Christina Linninger (Goethe-Universität Frankfurt), Olga Kunina-Habenicht (Goethe-Universität Frankfurt), Hendrik Lohse-Bossenz (PädQUIS® gGmbH – Kooperationsinstitut der Freien Universität Berlin), Theresa Dicke (Universität Duisburg-Essen), Mareike Kunter (Goethe-Universität Frankfurt)

## **Wirkung von Reflexion auf die Entwicklung von Selbstregulation bei angehenden Lehrkräften**

**Chair(s):** Christina Linninger (Goethe-Universität Frankfurt)

Dieser Beitrag beleuchtet die Bedeutung von Reflexion für die Entwicklung von Selbstregulation als Teil professioneller Kompetenz und den Zusammenhang mit Belastungserleben bei angehenden Lehrkräften.

**THEORETISCHER HINTERGRUND:** Lehrkräfte sehen sich beim Berufseinstieg vielfältigen Beanspruchungen des Schulalltags ausgesetzt (Keller-Schneider & Hericks, 2011; Veenman, 1984). Es besteht daher allgemeiner Konsens über die Notwendigkeit von Unterstützungsmaßnahmen für Berufseinsteiger, auch wenn Struktur und Ausmaß dieser Angebote international stark variieren (EURYDICE, 2002; Deinum, 2011). In Deutschland werden angehende Lehrkräfte in der zweiten, praxisbegleitenden Phase der Lehramtsausbildung (Vorbereitungsdienst/Referendariat) während ihres Berufseinstiegs durch Ausbildungsangebote an der Schule und im begleitenden Studienseminar unterstützt (Terhart, 2004). Es bestehen vielerseits Bemühungen, die Qualität dieser Lerngelegenheiten im Vorbereitungsdienst durch strukturelle Veränderungen und die Implementation neuer Ausbildungselemente zu verbessern (z.B. Gerdes & Annas-Sieler, 2013). Als ein Ziel von Ausbildungsangeboten und als zentraler Prozess bei der professionellen Entwicklung wird unter anderem die Reflexion von praktischen Erfahrungen diskutiert (Etscheidt, Curran & Sawyer, 2012; Wood & Stanulis, 2009). Bezüglich der vermittelnden Prozesse bei der Entwicklung professioneller Kompetenz besteht allerdings noch Forschungsbedarf (Beijaard, Buitink & Kessels, 2010; Hascher, 2011), denn es ist wünschenswert, differenzierte Effekte auf einzelne Kompetenzaspekte zu identifizieren, um Lernprozesse im Vorbereitungsdienst zu verstehen und gezielt steuern zu können. Decker und Kunter (2013) konnten zeigen, dass die im Studienseminar angeregte Reflexion zu einer Veränderung der epistemologischen Überzeugungen von Referendarinnen und Referendaren beiträgt. Die vorliegende Untersuchung setzt bei der Exploration eines weiteren möglichen Effekts von Reflexion auf die professionelle Kompetenz an: Gegenstand der kritischen Auseinandersetzung sind im Referendariat häufig problematische, als belastend empfundene Situationen. Am Ende einer gelungenen Reflexion dieser Erlebnisse sollten sich veränderte Verhaltensdispositionen in Bezug auf künftige Situationen unter angemessener Berücksichtigung der eigenen Ressourcen auf Seiten der angehenden Lehrkräfte zeigen. In diesem Fall hätte die Reflexion also einen direkten Effekt auf die Selbstregulation. Diese wird als Aspekt professioneller Kompetenz verstanden, bezieht sich auf einen nachhaltigen Umgang mit den eigenen Ressourcen und weist z.B. querschnittliche Zusammenhänge mit emotionaler Erschöpfung als Indikator beruflichen Wohlbefindens auf (Baumert & Kunter, 2006; Klusmann, Kunter, Voss & Baumert, 2012). Martin, Nejad, Colmar und Liem (2012) formulieren als Facette der Selbstregulation, die sich im Speziellen auf den Umgang mit neuen, sich verändernden bzw. unsicheren Situationen bezieht, die Anpassungsfähigkeit.

**FRAGESTELLUNG:** Vor dem Hintergrund der vorangegangenen Überlegungen stellt sich also die Frage, ob angehende Lehrkräfte durch Reflexion von problematischen Erlebnissen, vermittelt über die Entwicklung von Anpassungsfähigkeit, ihr Beanspruchungserleben vermindern können.

**METHODE:** Die längsschnittliche Untersuchung der vorliegenden Fragestellung ist eingebettet in das Forschungsprogramm BiWiss (Terhart et al., 2012). In die Pfadanalyse gehen Daten von 219 Referendarinnen und Referendaren aus 21 Studienseminaren in Nordrhein-Westfalen ein, die nach einem Jahr im Vorbereitungsdienst (t1) und an dessen Ende (t2) zu ihrem Erleben an Ausbildungsschule und Studienseminar befragt wurden. Im Selbstbericht wurden theoriebasierte und handlungsbezogene Reflexion von problematischen Situationen (Eigenkonstruktion;  $\alpha=.68/.69$ ), Anpassungsfähigkeit (verändert nach Martin et al., 2012;  $\alpha=.77$ ) sowie emotionale Erschöpfung (verändert nach Enzmann & Kleiber, 1989;  $\alpha=.77$ ) erhoben.

**ERGEBNISSE:** Eine manifeste Pfadanalyse mit Berücksichtigung der Mehrebenenstruktur der Daten zeigt auf der Individualebene erwartungsgemäß keinen direkten Effekt von Reflexion auf emotionale Erschöpfung, aber einen signifikanten indirekten Effekt über Anpassungsfähigkeit: Handlungsbezogene und theoriebasierte Reflexion zu t1 haben positiven Einfluss auf die Anpassungsfähigkeit zu t2 ( $\beta=.24/.17$ ) und Anpassungsfähigkeit zu t2 hängt erwartungsgemäß negativ mit emotionaler Erschöpfung zu t2 zusammen ( $\beta=-.36$ ). Auch nach der Kontrolle von emotionaler Erschöpfung zu t1 bleibt der Zusammenhang zwischen Anpassungsfähigkeit und emotionaler Erschöpfung bestehen ( $\beta=-.19$ ).

**DISKUSSION:** Trotz insgesamt kleinen Effekten sprechen die Ergebnisse dafür, dass die Reflexion von Praxiserfahrungen, vermittelt über Anpassungsfähigkeit als Element professioneller Kompetenz, einen Beitrag zur Reduktion von emotionaler Erschöpfung bei angehenden Lehrkräften leisten kann. Die Datenanalyse auf latenter Ebene zur Berechnung messfehlerbereinigter Effekte steht noch aus.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9 (4), 469–520.

Beijaard, D., J. Buitink & Kessels, C. (2010). Teacher Induction. In B. McGaw, P. L. Peterson & E. Baker (Hrsg.), International Encyclopedia of Education (S. 563–568). Amsterdam: Elsevier.

Decker, A.-T. & Kunter, M. (2013, August). The influence of discourse quality in teacher induction programs on beginning teachers' beliefs. Vortrag gehalten auf der 15th Biennial Conference of European Association for Research on Learning and Instruction (EARLI), München.

Deinum, J. F. (2011). Innote Annual Report III.

Enzmann, D. & Kleiber, D. (1989). Helfer-Leiden: Streß und Burnout in psychosozialen Berufen. Heidelberg: Asanger.

Etscheidt, S., Curran, C. M. & Sawyer, C. M. (2012). Promoting Reflection in Teacher Preparation Programs: A Multilevel Model. Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children, 35 (1), 7–26.

EURYDICE (2002). The teaching profession in Europe: Profile, trends, and concerns. Report I: Initial training and transition to working life. General lower secondary education. Brüssel.

Gerdes, R. & Annas-Sieler, D. (2011). Der reformierte Vorbereitungsdienst. Neue Lehrerausbildung in NRW. Schule NRW, 454–457.

Hascher, T. (2011). Forschung zur Wirksamkeit der Lehrerbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), Handbuch der Forschung zum Lehrerberuf (S. 418–440). Münster: Waxmann.

Keller-Schneider, M. & Hericks, U. (2011). Forschungen zum Berufseinstieg. Übergang von der Ausbildung in den Beruf. In E. Terhart, H. Bennewitz und M. Rothland (Hrsg.), Handbuch der Forschung zum Lehrerberuf (S. 296–313). Münster: Waxmann.

Klusmann, U., Kunter, M., Voss, T. & Baumert J. (2012). Berufliche Beanspruchung von angehenden Lehrkräften: Die Effekte von Persönlichkeit, praktischer Erfahrung und professioneller Kompetenz. Zeitschrift für Pädagogische Psychologie, 26 (4), 275–290.

Martin, A. J., Nejad, H., Colmar, S., & Liem, G. A. D. (2012). Adaptability: Conceptual and empirical perspectives on responses to change, novelty and uncertainty. Australian Journal of Guidance and Counselling, 22 (1), 58–81.

Terhart, E. (2004). Struktur und Organisation der Lehrerbildung in Deutschland. In S. Blömeke, P. Reinhold, G. Tulodziecki & J. Wildt (Hrsg.), Handbuch Lehrerbildung (S. 37–59). Bad Heilbrunn/Braunschweig: Klinkhardt/Westermann.

Terhart, E., Schulze-Stocker, F., Kunina-Habenicht, O., Dicke, T., Förster, D., Lohse-Bossenz, H., Gößling, J., Kunter, M., Baumert, J. & Leutner, D. (2012). Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung – Eine Kurzdarstellung des BiWiss-Projekts. Lehrerbildung auf dem Prüfstand, 5 (2), 96–106.

Veenman, S. (1984). Perceived Problems of Beginning Teachers. Review of Educational Research 54 (2), 143–178.

Wood, A. L. & Stanulis, R. N. (2009). Quality Teacher Induction: "Fourth-Wave" (1997–2006) Induction Programs. The New Educator 5 (1), 1–23.





## **Intensive Lehrerkooperation unter Stress – Welche Ressourcen und Stressbewältigungsstrategien sind notwendig?**

Intensive Lehrerkooperation gilt in der Schulentwicklungsforschung als ein wesentlicher Erfolgsfaktor für eine effektive und langfristige Implementation von Innovationen in der Schule. Gleichzeitig wird in empirischen Studien deutlich, dass eine intensive Form der Lehrerkooperation, nur in geringem Maße von Lehrkräften ausgeübt wird (vgl. Fussangel, 2008; Pröbstel, 2008; Soltau, in Druck). Zahlreiche Studien befassten bzw. befassen sich folgerichtig mit den Bedingungen für intensive Lehrerkooperation (vgl. Huber & Ahlgrimm, 2008; Soltau, in Druck). Allerdings besteht bei der Benennung von förderlichen Faktoren von intensiver Lehrerkooperation bislang keine Einigkeit. So werden in bisherigen Studien unterschiedlichste Einflussfaktoren auf institutioneller, organisationsstruktureller und individueller Ebene aufgeführt, ohne einem einheitlichen theoretischen Modell zu folgen. Zudem wird oftmals vernachlässigt, dass Lehrkräfte zu einer der am stärksten an Burnout gefährdetsten Berufsgruppe gehören (vgl. Käser & Wasch, 2009; Poschkamp, 2008; Rothland, 2013). Insgesamt ist festzustellen, dass es an einem fundierten theoretischen Modell mangelt, welches die Annahmen notwendiger Bedingungen mit einer Definition von Lehrerkooperation, im Sinne von professionellem Handeln, verbindet. An dieser Stelle knüpft diese Studie an, indem sie versucht, mit der COR-Theorie (vgl. Hobfoll, 1998; Buchwald & Hobfoll, 2004) und dem multiaxialen Copingmodell (vgl. Dunahoo, Monnier, Hobfoll, Hulsizer, & Johnson, 1998; Schwarzer, Starke & Buchwald, 2004) eine signifikante theoretische Grundlage für die Beantwortung der Frage nach den notwendigen Bedingungen von Lehrerkooperation unter Stress zu schaffen. Intensive Lehrerkooperation wird in dieser Analyse in Anlehnung an Gräsel, Fussangel und Pröbstel (2006) als Kokonstruktion bezeichnet und beinhaltet damit u. a. intensive Absprachen über einen längeren Zeitraum hinweg, die eine konkrete Zielinterdependenz und auch eine Abstimmung der Interaktionspartner hinsichtlich des Arbeitsprozesses benötigen.

Ausgangspunkt der Studie ist die Überlegung, dass die Forderung der Schulentwicklung nach erfolgreichen Prozessen des Innovationstransfers möglicherweise zu Stress bei Lehrkräften führen kann, da damit eine Veränderung routinierter Handelns verbunden ist. Traditionelle stresstheoretische Ansätze gehen davon aus, dass in erste Linie das Individuum selbst durch kognitive Bewertungen bestimmt, wie stressreich eine Anforderung der Umwelt interpretiert wird (vgl. Lazarus & Folkman, 1984). Problematisch daran ist, dass die soziale Umgebung, in dem die Wahrnehmungen und das Verhalten des Individuums stattfinden, keine Berücksichtigung findet. Die COR-Theorie ermöglicht eine erweiterte Perspektive, indem neben individuumszentrierten auch soziale Aspekte Berücksichtigung finden. Dabei stehen nicht die subjektiven Einschätzungen eines stressreichen Erlebnisses im Vordergrund, sondern dessen Auswirkung auf die Ressourcen eines Individuums. Der Stress einer Lehrkraft wird als Ressourcenverlust, -bedrohung oder dem Ausbleiben notwendiger Ressourcengewinne definiert. Stress im Lehrerberuf betrifft in der Regel nicht nur eine Lehrkraft, sondern das gesamte Kollegium, dem Interaktionen zugrunde liegen. In Assoziation mit der COR-Theorie verdeutlicht das multiaxiale Copingmodell den interpersonalen Prozess des Ressourcenmanagements. Hierbei gelten - anstelle des emotions- oder problemzentrierten Copings der herkömmlichen Ansätze - vielmehr prosoziale bzw. aktiv-prosoziale Strategien als effektiver Weg einer langfristigen Stressbewältigung. Entsprechend lässt sich annehmen, dass die Stressentwicklung nicht allein durch Ressourcenverluste bzw. -gewinne beeinflusst wird, sondern vor allem durch die Stressbewältigung in Interaktion mit Anderen.

Eine gemeinsame Stressbewältigung geht allerdings nicht nur mit positiven Effekten einher, so dass insgesamt sechs Strategien berücksichtigt werden, denen eine unterschiedliche Wirkung in Bezug auf die Stressbewältigung und damit einhergehend für die Ausübung intensiver Lehrerkooperation zugesprochen wird: Selbstbehauptung, Vermeidung, Suche nach sozialer Unterstützung, rücksichtsvolles Handeln, aggressiv-antisoziales Handeln und indirektes Handeln.

Insgesamt zeigen die Befunde, die mittels eines Strukturgleichungsmodells gewonnen wurden, dass die soziale Unterstützung im Kollegium sowie die eigene Bereitschaft nach Unterstützung zu fragen - verbunden mit ausgeprägter fachlicher Selbstsicherheit - zentrale Aspekte für die intensive Lehrerkooperation unter Stress sind.

Die notwendigen Daten wurden im Rahmen des vom BMBF geförderten Projektes LIMA - Lehrerfortbildungen zur Innovationsunterstützung im Mathematikunterricht in der Grundschule (vgl. Gräsel, Buchwald, Selter & Bensen, 2009) erhoben. Die gültige Stichprobe dieser Analyse umfasst 118 Lehrkräfte aus 24 Grundschulen in NRW.

Dunahoo, C. L., Monnier, J., Hobfoll, S. E., Hulsizer, M. R. & Johnson, R. (1998). There's more than rugged individualism in coping. Part I: Even the lone ranger had tonto. *Anxiety, Stress and Coping*, 11, 137-165.

Fussangel, K. (2008). Subjektive Theorien von Lehrkräften zur Kooperation. Eine Analyse der Zusammenarbeit von Lehrerinnen und Lehrern in Lerngemeinschaften [Internetdokument]. Universität Wuppertal. Verfügbar unter <http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3A468-20080475> [24.11.2012].

Gräsel, C., Fussangel, K. & Pröbstel, C. (2006). Lehrkräfte zur Kooperation anregen - eine Aufgabe für Sisyphos? *Zeitschrift für Pädagogik*, 52 (2), 205-219.

Hobfoll, S. E. (1998). *Stress, culture, and community. The psychology and philosophy of stress*. New York: Plenum Press.

Huber, S. & Ahlgrimm, F. (2008). Was Lehrkräfte davon abhält zusammenzuarbeiten – Bedingungen für das Gelingen von Kooperation. In A. Bartz, J. Fabian, S. Huber, C. Kloft, H. Rosenbusch & H. Sassenscheidt (Hrsg.), *PraxisWissen Schulleitung* (Bd. 81.10). München: Wolters, Kluwer.

Käser, U. & Wasch, J. (2009). *Burnout bei Lehrerinnen und Lehrern. Eine Bedingungsanalyse im Schulformvergleich*. Berlin: Logos.

Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Poschkamp, T. (2008). *Lehrergesundheit. Belastungsmuster, Burnout und Social Support bei dienstunfähigen Lehrkräften* (Band 1). In P. Buchwald & T. Ringeisen (Hrsg.), *Angewandte Stress- und Bewältigungsforschung*. Logos Verlag.

Rothland, M. (2013). *Belastung und Beanspruchung im Lehrerberuf und die Modellierung professioneller Kompetenz von Lehrerinnen und Lehrern*. In M. Rothland (Hrsg.), *Belastung und Beanspruchung im Lehrerberuf* (S. 7-20). Wiesbaden: Springer.

Schwarzer, C., Starke, D. & Buchwald, P. (2004). Die Diagnose multiaxialer Stressbewältigung mit dem multiaxialen Stressbewältigungsinventar (SBI). In P. Buchwald, C. Schwarzer & S. E. Hobfoll (Hrsg.), *Stress gemeinsam bewältigen. Ressourcenmanagement und multiaxiales Coping* (S. 60-73). Göttingen: Hogrefe.

Soltau, A. (in Druck). *Isolation aus Unsicherheit? Berufliche Unsicherheit und deren Zusammenhang zur Lehrerkooperation*. Dissertation. Universität Bremen.

Mittwoch,  
05.03.2014,

09:00 - 10:45,

IG-Farben-Haus

Raum 0.251





## Kooperation von Lehrkräften im Unterrichtsentwicklungsprogramm SINUS an Grundschulen

Hintergrund: Empirische Studien belegen bedeutsame Zusammenhänge zwischen der Qualität des Lehrens und der Leistungszuwächse von Lernenden (Hattie, 2009). Internationale und nationale Vergleichsstudien wie TIMSS, PISA und PIRLS zeigen, dass die Qualität des Unterrichts an deutschen Schulen nicht ausreicht, um optimale Lernbedingungen für Schülerinnen und Schüler zu schaffen. In SINUS an Grundschulen soll der mathematische und naturwissenschaftliche Unterricht über die Professionalisierung von Lehrpersonen weiterentwickelt und verbessert werden. Die erfolgreiche Umsetzung eines solchen Professionalisierungsprogramms erfordert ein hohes Maß an Kooperation zwischen den Lehrpersonen (Putnam & Borko, 2000; Gräsel, 2010). Das bedeutet, Lehrkräfte brauchen die Möglichkeit Ideen auszutauschen, gemeinsame Unterrichtskonzepte zu entwickeln, zu erproben und zu reflektieren. In deutschen Schulen jedoch findet Kooperation eher auf der Ebene des Materialaustausches und der Besprechung von Alltagsproblemen statt (Gräsel, Fussangel & Parchmann, 2006; Harazd & Drossel, 2011). Intensivere Formen der Kooperation werden von Lehrpersonen zwar positiv bewertet, die Umsetzung scheitert jedoch häufig an Rahmenbedingungen und persönlichen Einstellungen (Soltau & Mienert, 2009). In SINUS an Grundschulen ist die Kooperation von Lehrpersonen sowohl an ihrer eigenen Schule als auch über die eigenen Schulgrenzen hinaus ein zentrales Element zu ihrer Professionalisierung. In diesem Beitrag soll der Frage nachgegangen werden, inwieweit die Kooperation in den SINUS-Schulen gelingt.

Forschungsfragen: Ein Ziel der wissenschaftlichen Begleitforschung ist es, herauszufinden, wie die Kooperation in den Schulen stattfindet und wie sie sich über die Laufzeit des Programms entwickelt. Damit sollen Faktoren identifiziert werden, die zum Erfolg eines Programms zur Professionalisierung von Lehrpersonen beitragen. Die zentralen Forschungsfragen sind dabei:

1. Welche Rolle spielt die Kooperation für die Wahrnehmung des Programms in der Schule?
2. In welcher Intensität arbeiten die Lehrpersonen einer Schule zusammen?
3. Welche Faktoren fördern oder behindern die Zusammenarbeit der Lehrpersonen?

Methoden: Das Programm SINUS an Grundschulen wurde bundesweit von 2009 bis 2013 mit 5500 Lehrpersonen an 850 Schulen durchgeführt. Die wissenschaftliche Begleitforschung umfasst multimethodische Ansätze der Datengenerierung und der Datenauswertung und erreicht damit einen möglichst umfassenden Blick auf die Wirkung der Programminhalte (Fischer et al., 2013). Für den vorliegenden Beitrag werden Daten aus zwei verschiedenen Studien miteinander verbunden. Zum einem werden Skalen zur Zusammenarbeit der Lehrpersonen aus der Onlinebefragung aller Beteiligten 2010 genutzt. An der Befragung beteiligten sich 332 Schulleitungen und 1662 Lehrpersonen. Zum anderen werden Daten von 50 Schulen aus der Studie zu dem in SINUS eingesetzten Dokumentationsinstrument herangezogen. Anhand des Instrumentes, mit dem die Schulgruppen ihren Entwicklungsprozess dokumentieren, wurde die Kooperation an der Schule eingeschätzt. Zur Beantwortung der Forschungsfragen wurden die quantitativen Daten aus der Online-Befragung (Regressionsanalysen) mit den Daten der qualitativen Analyse (Deskriptive Daten und Fallanalysen) aus der Einschätzung der Dokumentationsinstrumente miteinander verknüpft, um einen mehrperspektivischen Einblick in die Kooperation der Lehrpersonen zu erhalten.

Ergebnisse: Die Analysen der Online-Befragung zeigen, dass die Kooperation eine der wichtigsten Komponenten für die wahrgenommene Unterstützung, für Zufriedenheit und Entwicklung im Programm ist. Die Analyse der Dokumentationen zeigt, dass die SINUS-Gruppen der Stichprobe überwiegend auf einer Ebene kooperieren, die über den Austausch von Materialien hinausgeht. Einige Gruppen scheinen sich intensiv auszutauschen, ihren Unterricht gemeinsam zu gestalten und die SINUS-Arbeit innerhalb der Schule zu etablieren. Eine fallanalytische Untersuchung der Kooperationsarbeit der einzelnen Schulen zeigt, welche Einschränkungen (Einzelkämpferverhalten, Akzeptanzprobleme, Konkurrenz anderer Maßnahmen) und positive Entwicklungen der Kooperation (z.B. schulübergreifende Kooperation) in den Schulen stattfanden. Die Befunde aus den verschiedenen Instrumenten ergänzen sich. Auf diese Weise wird ersichtlich, welche Bedeutung die Kooperation für die Wahrnehmung des Programms hat (Befragung) und in welcher Qualität die Kooperation stattfindet (Dokumentationen). Schließlich geben identifizierte Probleme bei der Umsetzung der Kooperation (Fallanalysen) wichtige Hinweise und Ansatzpunkte für zukünftige Unterrichtsentwicklungsprogramme.

Fischer, C., Kobarg, M., Dalehefte, I. M., & Trepke, F. (2013). Ein Unterrichtsentwicklungsprogramm wissenschaftlich begleiten. Anlage und Hintergründe des Forschungsdesigns. *Psychologie in Erziehung und Unterricht*, 60(1), 26-31.

Gräsel, C., Fussangel, K., & Parchmann, I. (2006). Lerngemeinschaften in der Lehrerfortbildung. Kooperationserfahrungen und -überzeugungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 545-561.

Gräsel, C. (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 13(1), 7-20.

Harazd, B., & Drossel, K. (2011). Formen der Lehrkooperationen und ihre schulische Bedingungen - Empirische Untersuchung zu kollegialen Zusammenarbeit und Schulleitungshandeln. *Empirische Pädagogik*, 25 (2), 145-160.

Hattie, J. A. C. (2009). *Visible Learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Putnam, R. T., & Borko, H. (2000). What Do New Views of Knowledge and Thinking Have to Say About Research on Teacher Learning? *Educational Researcher*, 29(1), 4-15.

Soltau, A., & Mienert, M. (2009). Teamorientierung und Einstellung zu Formen der Lehrkooperation bei Lehrkräften. *Psychologie in Erziehung und Unterricht*, 56(3), 213-223.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
IG-Farben-Haus  
Raum 0.251



## Zur Erklärung der Resignationstendenz von Lehrer(inne)n

Durch ihre direkte Arbeit mit Kindern und Jugendlichen sowie die Ausführung politisch und bürokratisch beschlossener Bildungsabläufe sind Lehrer(innen) eine der wichtigsten Akteursgruppen des deutschen Bildungssystems. Hinsichtlich der Arbeitsbelastung zählt der Lehrerberuf jedoch zu den anstrengendsten und herausforderndsten Berufen. Betrachtet man die Anzahl von Lehrer(inne)n, die in ihrem beruflichen Alltag resignieren (Jehle & Schmitz, 2007; Klusmann et al., 2012) und sich vorzeitig aus dem Beruf zurückziehen (Schaarschmidt, 2005; Scharf & Rupprecht, 2011), erscheint es bedeutsam, Möglichkeiten zur Stärkung der physischen und psychischen Gesundheit von Lehrer(inne)n zu kennen.

Angelehnt an das von der Universität Hamburg und dem Deutschen Zentrum für Hochschul- und Wissenschaftsforschung durchgeführte BMBF-Projekt KomPaed - Tätigkeitsbezogene Kompetenzen in pädagogischen Handlungsfeldern (Schwippert et al., 2011), möchte der Beitrag mögliche Erklärungen für die Resignationstendenz von Lehrer(inne)n untersuchen. Diese bringt individuelle Gesundheitsressourcen und -risiken zum Ausdruck und kann Hinweise auf die Belastbarkeit, die Zufriedenheit und das Engagement von Lehrer(inne)n geben (Schaarschmidt & Kieschke, 2007). Es werden zum einen selbstberichtete generische berufliche Tätigkeiten und Anforderungen sowie selbsteingeschätzte generische Kompetenzen von Lehrer(inne)n allgemein- und berufsbildender Schulen herausgearbeitet. Zum anderen wird der Zusammenhang zwischen der individuellen Passung selbstberichteter Tätigkeiten, Anforderungen und Kompetenzen und der Resignationstendenz unter Kontrolle bildungspolitischer und schulischer Rahmenbedingungen untersucht.

Aus theoretischer Sicht verbindet der Beitrag vorhandene Arbeiten aus den Bereichen der Lehrer-, Arbeitsbelastungs- und Kompetenzforschung. Fokussiert wird dabei (1) der Person-Environment Fit Ansatz (Caplan, 1987; Kristof-Brown et al., 2005; Lauver & Kristof-Brown, 2001; Rosman & Burke, 1980; Sekiguchi, 2007), insbesondere die individuelle Passung zwischen den Charakteristika eines Individuums und denen der auszuübenden Tätigkeit (Person-Job Fit), (2) Arbeiten zur professionellen Kompetenz von Lehrer(inne)n (Baumert & Kunter, 2011; Baumert & Kunter, 2006; Kunter, 2010) und zur Motivation (Barnabé & Burns, 1994; Klusmann, 2011; Malmberg, 2006; Müller et al., 2008; Neves De Jesus & Lens, 2005; Pohlmann & Möller, 2010; Retelsdorf & Möller, 2012; Sylvia & Hutchison, 1985; Watt et al., 2012), (3) Literatur zur Arbeitsbelastung und Lehrergesundheit (BAuA, 2005; Doll, 2010; Freitag, 1998; Klusmann et al., 2012; Ksienzy & Schaarschmidt, 2005; Rothland, 2007; Schaarschmidt & Kieschke, 2007; Van Dick & Stegmann, 2007) sowie (4) Forschung zu Kompetenzdimensionen (Erpenbeck, 2010; Kultusministerkonferenz, 2005; OECD, 2005; Rychen & Salganik, 2003; Weinert, 2001) und Messverfahren (Braun et al., 2011; Braun et al., 2008; Frey, 2004; Gnahn, 2007; Heyse, 2010; Kaufhold, 2006).

Der Beitrag möchte Antworten auf die Fragen liefern, (1a) welche selbstberichteten generischen beruflichen Tätigkeiten Lehrer(innen) allgemein- und berufsbildender Schulen ausüben, (1b) welche Anforderungen an sie gestellt werden und (1c) welche Tätigkeiten und Anforderungen dabei aus individueller Sicht einen besonders großen Stellenwert haben. Außerdem ist von Interesse, (2a) welche selbsteingeschätzten generischen Kompetenzen Lehrer(innen) bei sich wahrnehmen und (2b) wie die individuelle Passung selbstberichteter Tätigkeiten, Anforderungen und Kompetenzen ist. Schließlich soll untersucht werden, (3) welche Rolle diese individuelle Passung sowie die bildungspolitischen und schulischen Rahmenbedingungen hinsichtlich der Resignationstendenz von Lehrer(inne)n spielen.

Die Forschungsfragen werden anhand von Befragungsdaten der KomPaed Onlinebefragung von Hochschulabsolvent(innen) der Prüfungsjahrgänge 2000/2001, 2004/2005, 2008/2009 bearbeitet (Feldphase: Jahreswechsel 2013/2014). In der Stichprobe befinden sich ca. 1860 Lehramtsabsolvent(innen), die zum Befragungszeitpunkt zwischen 5 und 13 Jahren Berufspraxis besitzen. Das Befragungsinstrument wurde vor dem Hintergrund einer Instrumenten- und Literaturrecherche sowie im Projekt KomPaed durchgeführter Arbeitgeberinterviews auf Basis des Job Requirement Ansatzes (Felstead et al., 2007; OECD, 2011; Peterson et al., 2001) erarbeitet. Die erhobenen Daten werden zum einen deskriptiv analysiert, zum anderen werden Konfirmatorische Faktorenanalysen und Strukturgleichungsmodelle mit MPlus berechnet. Vorgestellt werden soll die konzeptionelle Umsetzung des theoretischen Modells in Fragebogenform. Zudem erwarten wir erste empirische Ergebnisse zu den konzeptionellen Dimensionen und zum Ausmaß von Resignationstendenz bei Lehrer(inne)n, ggf. differenziert nach Schultyp und Fächergruppe, berichten zu können.

Barnabé, C. & Burns, M., 1994. Teachers' job characteristics and motivation. *Educational Research*, 36 (2), 171-185.

BAuA, 2005. Lehrergesundheit - Praxisrelevante Modelle zur nachhaltigen Gesundheitsförderung von Lehrern auf dem Prüfstand. Tagungsbericht Tb 141. Dortmund/Berlin/Dresden: BAuA, 1-135.

Baumert, J. & Kunter, M., 2011. Das Kompetenzmodell von COACTIV. In: Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). *Professionelle Kompetenz von Lehrkräften - Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann, 29-53.

Baumert, J. & Kunter, M., 2006. Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469-520.

Braun, E., Sheik, H. & Hannover, B., 2011. Self-rated competences and future vocational success: a longitudinal study. *Assessment & Evaluation in Higher Education*, 36 (4), 417-427.

Braun, E., Gusy, B., Leidner, B. & Hannover, B., 2008. Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BévaKomp). *Diagnostica*, 54 (1), 30-42.

Caplan, R. D., 1987. Person-Environment Fit Theory and Organizations: Commensurate Dimensions, Time Perspectives, and Mechanisms. *Journal of Vocational Behavior*, 31, 248-267.

Doll, J., 2010. Entwurf eines Hamburger Lehrpanels: Belastung, Gesundheit und professionelle Handlungskompetenz. ZUSE-Diskussionspapier Nummer 3.

Erpenbeck, J., 2010. Kompetenzen - eine begriffliche Klärung. In: Heyse, V., Erpenbeck, J. & Ortman, S. (Hrsg.). *Grundstrukturen menschlicher Kompetenzen - Praxiserprobte Konzepte und Instrumente*. Waxmann: Münster, 13-19.

Felstead, A., Gallie, D., Green, F. & Zhou, Y., 2007. *Skills at Work 1986-2006*. Oxford: Skope.

Freitag, M., 1998. Was ist eine gesunde Schule? Einflüsse des Schulklimas auf Schüler- und Lehrergesundheit. Weinheim: Juventa.

Frey, A., 2004. Die Kompetenzstruktur von Studierenden des Lehrberufes. Eine internationale Studie. *Zeitschrift für Pädagogik*, 50 (6), 903-925.

Gnahn, D., 2007. Kompetenzen - Erwerb, Erlangung, Instrumente. Gütersloh: Bertelsmann.

Heyse, V., 2010. Verfahren zur Kompetenzermittlung und Kompetenzentwicklung. KODE im Praxistest. In: Heyse, V., Erpenbeck, J. & Ortman, S. (Hrsg.). *Grundstrukturen menschlicher Kompetenzen - Praxiserprobte Konzepte und Instrumente*. Waxmann: Münster, 55-76/81/121-122.

Jehle, P. & Schmitz, E., 2007. Innere Kündigung und vorzeitige Pensionierung von Lehrpersonen. In: Rothland, M. (Hrsg.). *Belastung und Beanspruchung im Lehrerberuf. Modelle, Befunde, Interventionen*. Wiesbaden: VS, 160-184.

Kaufhold, M., 2006. Kompetenz und Kompetenzerfassung. Analyse und Beurteilung von Verfahren der Kompetenzerfassung. Wiesbaden: VS.

Klusmann, U., Kunter, M., Voss, T. & Baumert, J., 2012. Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 26 (4), 275-290.

Klusmann, U., 2011. Allgemeine berufliche Motivation und Selbstregulation. In: Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). *Professionelle Kompetenz von Lehrkräften - Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann, 277-294.

Kristof-Brown, A. L., Zimmerman, R. D. & Johnson, E. C., 2005. Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group fit. *Personnel Psychology*, 58, 281-342.

Ksienzy, B. & Schaarschmidt, U., 2005. Beanspruchung und schulische Arbeitsbedingungen. In: Schaarschmidt, U. (Hrsg.). *Halbtagsjobber? Psychische Gesundheit im Lehrerberuf - Analyse eines veränderungsbedürftigen Zustandes*. Weinheim/Basel: Beltz, 72-87.

Kultusministerkonferenz, 2005. *Qualifikationsrahmen für Deutsche Hochschulabschlüsse*. Bonn: KMK.

Kunter, M., 2010. Modellierung von Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 56, 307-312.

Lauver, K. J. & Kristof-Brown, A., 2001. Distinguishing between Employees' Perceptions of Person-Job and Person-Organization Fit. *Journal of Vocational Behavior*, 59, 454-470.

Malmberg, L.-E., 2006. Goal-orientation and teacher motivation among teacher applicants and student teachers. *Teaching and Teacher Education*, 22, 59-76.

Müller, F. H., Andreitz, I. & Palekic, M., 2008. Lehrermotivation - Ein vernachlässigtes Thema in der empirischen Forschung. *Ogdojne Znanosti*, 10 (1), 39-60.

Neves de Jesus, S. & Lens, W., 2005. An Integrated Model for the Study of Teacher Motivation. *Applied Psychology: An International Review*, 54 (1), 119-134.

OECD, 2011. *PIAAC Conceptual Framework of the Background Questionnaire Main Survey*. Paris: OECD.

OECD, 2005. *The Definition and Selection of Key Competencies. Executive Summary*. Paris: OECD.

Peterson, N. G., Borman, W. C., Jeanneret, P. R., Fleischman, E. A., Levin, K. Y., Campion, M. A., Mayfield, M. S., Morgeson, F. P., Pearlman, K., Gowing, M. K., Silver, M. B. & Dye, D. M., 2001. Understanding work using the Occupational Information Network (O\*NET): Implications for practice and research. *Personnel Psychology*, 54, 451-492.

Pohlmann, B. & Möller, J., 2010. Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie*, 24 (1), 73-84.

Retelsdorf, J. & Möller, J., 2012. Grundschule oder Gymnasium? Zur Motivation ein Lehramt zu studieren. *Zeitschrift für Pädagogische Psychologie*, 26 (1), 5-17.

Rosman, P. & Burke, R. J., 1980. Job Satisfaction, Self-Esteem, and the Fit between Perceived Self and Job on Valued Competencies. *The Journal of Psychology: Interdisciplinary and Applied*, 105 (2), 259-269.

Rothland, M., 2007. Belastung und Beanspruchung im Lehrerberuf. In: Rothland, M. (Hrsg.). *Belastung und Beanspruchung im Lehrerberuf. Modelle, Befunde, Interventionen*. Wiesbaden: VS, 7-10.

Rychen, D. S. & Salganik, L. H., 2003. A holistic model of competence. In: Rychen, D. S. & Salganik, L. H. (Hrsg.). *Key Competencies for a Successful Life and a Well-Functioning Society*. Bern: Hogrefe & Huber, 41-62.

Schaarschmidt, U. & Kieschke, U., 2007. Einführung und Überblick. In: Schaarschmidt, U. & Kieschke, U. (Hrsg.). *Gerüstet für den Schullauf - Psychologische Unterstützungsangebote für Lehrerinnen und Lehrer*. Weinheim: Beltz, 17-43.

Schaarschmidt, U., 2005. Potsdamer Lehrertude - Anliegen und Konzept. In: Schaarschmidt, U. (Hrsg.). *Halbtagsjobber? Psychische Gesundheit im Lehrerberuf - Analyse eines veränderungsbedürftigen Zustandes*. Weinheim/Basel: Beltz, 15-40/162-172.

Scharf, R. & Rupprecht, S., 2011. Schulstress: Jeder 5. Lehrer denkt an Frühpensionierung. DAK.

Schwippert, K., Braun, E., Prinz, D. & Schaeper, H., 2011. Projektantrag KomPaed - Tätigkeitsbezogene Kompetenzen in pädagogischen Handlungsfeldern. Hamburg/Hannover.

Sekiguchi, T., 2007. A contingency perspective of the importance of PJ fit and PO fit in employee selection. *Journal of Managerial Psychology*, 22 (2), 118-131.

Sylvia, R. D. & Hutchison, T., 1985. What Makes Ms. Johnson Teach? A Study of Teacher Motivation. *Human Relations*, 38 (9), 841-856.

Van Dick, R. & Stegmann, S., 2007. Belastung, Beanspruchung und Stress im Lehrerberuf - Theorien und Modelle. In: Rothland, M. (Hrsg.). *Belastung und Beanspruchung im Lehrerberuf. Modelle, Befunde, Interventionen*. Wiesbaden: VS, 34-51.

Watt, H. M. G., Richardson, P. W., Klusmann, U., Kunter, M., Beyer, B., Trautwein, U. & Baumert, J., 2012. Motivations for choosing teaching as a career: An international comparison using the FIT-Choice scale. *Teaching and Teacher Education*, 28, 791-805.

Weinert, F. E., 2001. Concept of Competence: A Conceptual Clarification. In: Rychen, D. S. & Salganik, L. H. (Hrsg.). *Defining and Selecting Key Competencies*. Bern: Hogrefe & Huber, 45-66.

## **Die Entwicklung der Qualität des Anfangsunterrichts im Modellprojekt „Bildungshaus 3 - 10“ und das berufliche Belastungserleben teilnehmender Pädagog/-inn/en - ein Zusammenhang?**

Im Modellprojekt „Bildungshaus 3 – 10“ des Landes Baden-Württemberg führen, im Rahmen einer engen Kooperation Erzieher/-innen und Lehrkräfte gemeinsam regelmäßig stattfindende altersgemischte und institutionsübergreifende Spiel- und Lerneinheiten für Kinder aus Kindergarten und Grundschule durch. Ziel ist dabei die individuelle Förderung der Kinder und deren bruchlose Bildungsbiographie mit reibungslosem Übergang vom Kindergarten in die Grundschule. Die wissenschaftliche Begleitung, gefördert vom Bundesministerium für Bildung und Forschung und dem Europäischen Sozialfond der Europäischen Union, geht auf Ebene der Einrichtungen u. a. folgenden Fragen nach: (1) Mit Blick auf die Unterrichtsqualität im Anfangsunterricht stellt sich die Frage, ob Lehrkräfte der Modellstandorte von der engen Zusammenarbeit mit Erzieher/-innen für ihre Unterrichtsgestaltung profitieren und sich dadurch einzelne Qualitätsmerkmale ihres Anfangsunterrichts im Verlauf der Projektlaufzeit verbessern. (2) Hinsichtlich der Pädagog/-inn/en wird die Frage untersucht, ob und wie sich die Teilnahme am Modellprojekt auf ihre berufliche Belastung im Projektverlauf auswirkt und ob sich dabei Unterschiede im Belastungserleben zwischen Erzieher/-innen und Lehrkräften zeigen. Das Forschungsdesign in der wissenschaftlichen Begleitung des Modellprojekts „Bildungshaus 3 – 10“ ist als Experimental-Kontrolldesign mit Modell- und Vergleichsstandorten konzipiert.

Zu (1). Als theoretischer Hintergrund zur Qualität des Anfangsunterrichts dient das dreidimensionale Modell von Klieme et al. (2006) und die Qualitätsmerkmale von Helmke (2009). Berücksichtigung findet hier das schülerorientierte Unterrichtsklima, die Strukturierung, die effiziente Klassenführung, die Motivierung, die individuelle Unterstützung und die kognitive Aktivierung der Lernenden. Auf Grund der altersgemischten Arbeit im Bildungshaus und des Anspruches einer individuellen Förderung der Kinder, werden außerdem Merkmale des selbständigen Lernens und der inneren Differenzierung im Unterricht fokussiert.

Die Unterrichtsqualität wurde anhand eines eigens entwickelten Unterrichtsbeobachtungsbogens (UBB) durch hoch-inferente Beobachtung in vivo eingeschätzt. Der Anfangsunterricht wurde dabei in jeweils einer ersten Klasse während mindestens drei Unterrichtsstunden in 32 Modell- und 24 Vergleichsschulen im Längsschnitt 2009 - 2011 und 2013 von jeweils eine/-r/-m Rater/-in beobachtet. Um eine ausreichende Interraterreliabilität zu gewährleisten, wurden alle Rater/-innen vorab mit Hilfe eines ausführlichen Manuals und Videosequenzen zu Unterrichtssituationen intensiv geschult. Für die Analyse der Daten des Anfangsunterrichts wurde für jedes Merkmal eine zweifaktorielle ANOVA mit Messwiederholung mit den Faktoren Modelltyp (ME, VE) und Zeit (2009-2013) gerechnet.

Die Ergebnisse weisen auf eine positive Entwicklung des Anfangsunterrichts in allen Qualitätsmerkmalen hin. Modellschulen werden dabei in der individuellen Unterstützung der Kinder und im Unterrichtsklima etwas besser beurteilt als Vergleichsschulen. Außerdem gibt ein Ergänzungsfragebogen für Lehrkräfte zur eigenen Einschätzung ihres Unterrichts (Jerusalem et al., 2009), der 2013 parallel zur Unterrichtsbeobachtung eingesetzt wurde Hinweise auf die Validität des Unterrichtsbeobachtungsbogens (UBB).

Zu (2). Die Teilnahme an einem Modellprojekt, das z. T. veränderte pädagogische Konzepte im Alltag nach sich zieht, könnte zu Stress und Belastung für Pädagog/-innen beitragen und sich ungünstig auf deren pädagogische Arbeit auswirken. Dieses legt die transaktionale Stresstheorie nahe, die Stress als Wechselwirkung von Personen- und Umweltbedingungen beschreibt (Lazarus et al., 1984).

Anhand des Fragebogens zur beruflichen Belastung (Schwarzer et al., 2001; Enzmann et al., 1989) wurden daher die Pädagog/-inn/en aller Modell- und Vergleichseinrichtungen im jährlichen Rhythmus, bisher zu drei Erhebungszeitpunkten 2011-2013, nach ihrer Arbeitsüberforderung, dem Kontrolliertheitserleben, ihrer Arbeitszufriedenheit und der Selbstwirksamkeit gefragt.

Dabei zeigen bisherige Ergebnisse (2011 und 2012, für beide Erhebungszeitpunkte gilt N = 593) keinen negativen Einfluss der Teilnahme auf das berufliche Belastungserleben, im Gegenteil: in den Modelleinrichtungen fühlen sich die Pädagog/-inn/en weniger überfordert und die Arbeitszufriedenheit ist höher. Allerdings zeigt sich auch ein Unterschied in den Professionen: Lehrkräfte fühlen sich zufriedener in ihrer Arbeit als Erzieher/-innen, ein Befund, der sich bisher im Verlauf sogar verstärkt. Inwiefern diese Ergebnisse mit der erhobenen Unterrichtsqualität zusammenhängen, werden Analysen mit gemischten Modellen zeigen.

Enzmann, D., Kleiber, D. (1989). Helfer-Leiden. Streß und Burnout in psychosozialen Berufen. Heidelberg: Asanger.

Helmke, A. (2009). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. 2. Aufl., Seelze: Kallmeyer.

Jerusalem, M., Drössler, S., Kleine, D., Klein-Heßling, J., Mittag, W., Röder, B. (2009). Förderung von Selbstwirksamkeit und Selbstbestimmung im Unterricht. Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Berlin: Universität.

Klieme, E., Lipowsky, F., Rakoczy, K., Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. In: M. Prenzel, L. Allolio-Näcke (Hrsg.), Untersuchungen zur Bildungsqualität von Schule, Münster: Waxmann, 128-146.

Lazarus, R. S., Folkman, S. (1984). Stress, appraisal, and coping. New York: Springer.

Schwarzer, R., Jerusalem, M. (Hrsg.) (2001). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Berlin: Universität.

## ***Rumination und negativer Affekt bei Grundschulkindern***

Chair(s): Anja Leonhardt (DIPF, IDeA)

Die Rolle von Rumination bei der Entwicklung und Aufrechterhaltung depressiver Symptome wurde bereits im Zusammenhang mit der Response Style Theory ausgiebig untersucht. Rumination (d.h. wiederkehrende Gedanken über die eigenen negative Gefühle und deren Ursprung) kann depressive Symptome und negativen Affekt aufrechterhalten und verlängern. Belege für diese Annahmen stammen aus zahlreichen quer- und längsschnittlichen Studien (Rood, Roelofs, Bögels, Nolen-Hoeksema & Schouten, 2009). So nimmt negativer Affekt an Tagen zu, an denen ruminieren wird (Brose, Schmiedek, Lövdén & Lindenberger, 2011; Dickson, Ciesla & Reilly, 2012; Silk, Steinberg & Morris, 2003). Das ist besonders dann der Fall, wenn an diesen Tagen auch negative Ereignisse aufgetreten sind (Genet & Siemer, 2012; Moberly & Watkins, 2008). Auch werden mehr ruminative Gedanken berichtet, wenn negative Ereignisse erlebt werden, (Brose et al., 2011; Schneiders et al., 2006). Auch wenn bisher bekannt ist, dass Rumination bereits im Kindes- und Jugendalter mit depressiven Symptomen zusammenhängt, ist in diesem Alter nichts über intraindividuelle Zusammenhänge zwischen negativen Ereignissen, ruminativen Gedanken und negativem Affekt bekannt. Wegen möglicher negativer Auswirkungen von Rumination und negativem Affekt sowohl auf das allgemeine Wohlbefinden als auch auf die Schulleistung (Pekrun, 1992), ist die Betrachtung der Zusammenhänge zwischen Rumination und negativem Affekt bei Kindern im tagtäglichen Schulkontext relevant. Mit der vorliegenden Studie sollen daher drei zentrale Fragen beantwortet werden: Ruminieren Grundschul Kinder im Alltag? Stehen negative Ereignisse im Zusammenhang mit dieser Rumination? Und führen negative Ereignisse und Rumination zu verstärktem negativem Affekt? In der FLUX-Studie haben 110 Grundschul Kinder (8-11 Jahre) vier Wochen lang viermal täglich ihren negativen Affekt auf einem Smartphone bewertet. Jeden Abend wurden die Kinder außerdem zu negativen Ereignissen und ruminativen Gedanken an diesem Tag befragt. Momentaner negativer Affekt und tägliche ruminative Gedanken und Gefühle wurden jeweils mit drei und sechs Items gemessen. Das Vorliegen eines negativen Ereignisses wurde dichotom erfasst. Ergebnisse von Mehrebenenanalysen für Messwiederholungsdaten zeigen, dass Grundschul Kinder zwar im Mittel wenig ruminieren, es aber reliable Unterschiede zwischen den Kindern gibt. Kinder berichteten mehr Rumination an Tagen mit negativen Ereignissen, wobei auch der vorhergehende negative Affekt ruminative Gedanken vorhersagte. Sowohl tägliche negative Ereignisse als auch tägliche ruminative Gedanken führten zu mehr negativem Affekt. Diese Zusammenhänge zwischen negativen Ereignissen, Rumination und negativem Affekt waren auch dann noch bedeutsam, wenn für den vorhergehenden negativen Affekt kontrolliert wurde. Zusammenfassend lässt sich festhalten, dass es zwischen Kindern beträchtliche Unterschiede im Ausmaß ruminativer Gedanken gibt. Ruminative Gedanken führen auch bei Kindern zu stärkerem negativem Affekt, vor allem im Zusammenspiel mit negativen Ereignissen. Die bei Erwachsenen und Jugendlichen gefundenen Zusammenhänge lassen sich also auch für Grundschul Kinder zeigen. Studien zum Vergleich von Rumination mit anderen Copingstrategien sind im Hinblick auf die Relevanz dieser Variablen für die Schulleistung wünschenswert.

- Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2011). Normal aging dampens the link between intrusive thoughts and negative affect in reaction to daily stressors. *Psychology and Aging*, 26, 488–502. doi:10.1037/a0022287
- Dickson, K. S., Ciesla, J. A., & Reilly, L. C. (2012). Rumination, worry, cognitive avoidance, and behavioral avoidance: Examination of temporal effects. *Behavior Therapy*, 43, 629–640.
- Genet, J. J., & Siemer, M. (2012). Rumination moderates the effects of daily events on negative mood: Results from a diary study. *Emotion*, 12, 1329–1339. doi:10.1037/a0028070
- Moberly, N. J., & Watkins, E. R. (2008). Ruminative self-focus, negative life events, and negative affect. *Behaviour Research and Therapy*, 46, 1034–1039. doi:10.1016/j.brat.2008.06.004
- Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, 41, 359–376. doi:10.1111/j.1464-0597.1992.tb00712.x
- Rood, L., Roelofs, J., Bögels, S. M., Nolen-Hoeksema, S., & Schouten, E. (2009). The influence of emotion-focused rumination and distraction on depressive symptoms in non-clinical youth: A meta-analytic review. *Clinical Psychology Review*, 29, 607–616. doi:10.1016/j.cpr.2009.07.001
- Schneiders, J., Nicolson, N. A., Berkhof, J., Feron, F. J., van Os, J., & deVries, M. W. (2006). Mood reactivity to daily negative events in early adolescence: Relationship to risk for psychopathology. *Developmental Psychology*, 42, 543–554. doi:10.1037/0012-1649.42.3.543
- Silk, J. S., Steinberg, L., & Morris, A. S. (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. *Child Development*, 74, 1869–1880. doi:10.1046/j.1467-8624.2003.00643.x



## **Emotionales Erleben im Französischunterricht im Verlauf der Sekundarstufe I**

Obwohl Emotionen als relevante Bedingungen für schulisches Lernen und Leisten vermutet wurden, finden sich bislang kaum Arbeiten die das Emotionserleben im schulischen Kontext a) systematisch fokussieren, b) ein längsschnittliches Design verfolgen und c) sich auf den Fremdsprachenunterricht konzentrieren. Emotionen wird jedoch eine bedeutsame Wirkung auf die Qualität von Lernen und Leisten im Unterricht zugesprochen (Götz, 2004). In Bezug auf die Erforschung des Fremdsprachenunterrichts wurde bereits häufiger das Desiderat nach einer Untersuchung von Emotionen im Fremdsprachenunterricht im Speziellen geäußert, die Forschung befindet sich jedoch noch am Anfang (vgl. Hu, 2003; Schwerdtfeger, 1997). Erkenntnisse insbesondere aus der Forschung zum naturwissenschaftlich-mathematischen Unterricht zeigen, dass eine einseitige Konzentration auf negative Emotionen nicht gerechtfertigt ist, sondern die gesamte Bandbreite an Emotionen im Unterricht erlebt wird. Außerdem wird deutlich, dass Unterrichtsemotionen domänenspezifisch organisiert sind, eine fachspezifische Betrachtung ist also notwendig. (vgl. Götz, Frenzel, Pekrun & Hall, 2006). Im sozial-kognitiven Kontroll-Wert-Modell der Emotionsentstehung von Pekrun (2000) werden Schülermerkmale (z.B. fachspezifisches Selbstkonzept, Fachwertschätzung) sowie Unterrichtsmerkmale wie bspw. das Fehlerklima und die Relevanz des korrekten Sprachgebrauchs als Auslöser von emotionalem Erleben verstanden.

Im Vortrag werden insbesondere zwei Fragestellungsbereiche untersucht:

1. Verändert sich das Erleben von Freude, Angst, Ärger und Langeweile von Schülerinnen und Schülern im gymnasialen Französischunterricht im Verlauf der Sekundarstufe I? Und wenn ja inwiefern?

2. Können ausgewählte Schüler- und Unterrichtsmerkmale als Prädiktor für Ausprägung und Entwicklung von Emotionen angenommen werden?

Mittels Fragebogen wurden N = 548 Schülerinnen und Schüler aus 35 Lerngruppen zu 6 Messzeitpunkten (Klasse 6 bis Klasse 9) mit einem standardisierten Fragebogen befragt. Dabei konnte an Daten aus einem vorigen Forschungsprojekt von Cronjäger (2009) angeknüpft werden. Als abhängige Variablen wurden die Emotionen Freude, Angst, Ärger und Langeweile anhand von erprobten 4-Item-Kurzskalen untersucht, die hinreichend gute interne Konsistenzwerte aufwiesen. Die Entwicklung der betrachteten Emotionen wurde mittels latenter Wachstumskurvenmodelle (LGCM) analysiert. Die Veränderungen lassen sich dabei nicht als lineare Prozesse beschreiben, sondern es ist von zwei Entwicklungsphasen auszugehen. Die resultierenden Modelle konnten die Daten gut abbilden (CFI=.95-.99; RMSEA=.06).

Es zeigte sich – unter Berücksichtigung der Fachleistung – dass im Französischunterricht das Freudeerleben über die Zeit abnimmt, während das Erleben negativer Emotionen ansteigt. Allerdings schwächt sich die Entwicklung im Verlauf der Sekundarstufe I ab. Dies unterstreicht die Wichtigkeit des Lernbeginns bzw. des ersten Lernjahres für die Entwicklung von Emotionen. Hinsichtlich der Schülermerkmale stehen insbesondere die Fachwertschätzung und das fachspezifische Selbstkonzept in einem deutlichen Zusammenhang mit den untersuchten Emotionen. Hinsichtlich der Unterrichtsmerkmale konnten insbesondere das Fehlerklima und die Motivierungsfähigkeit der Lehrkraft als wichtige Einflussfaktoren identifiziert werden.

Diese Ergebnisse verdeutlichen die Relevanz der persönliche Bedeutsamkeit für das Erlernen von Fremdsprachen und die Wichtigkeit eines positiven Fehlerklimas (vgl. Kleppin & Mehlhorn, 2008) im Unterricht auf der anderen Seite. Gleichzeitig unterstreichen sie die Bedeutsamkeit des ersten Lernjahres für die weitere Emotionsentwicklung.

Cronjäger, H. (2009). Emotionen im schulischen Fremdsprachenunterricht: Bedingungen, Wirkungen und Veränderungen im ersten Lernjahr Französisch. (Unveröffentlichte Dissertation, Universität Jena).

Götz, T. (2004). Emotionales Erleben und selbstreguliertes Lernen bei Schülern im Fach Mathematik. München: Utz

Götz, T., A. C. Frenzel, R. Pekrun & N. C. Hall (2006). „The Domain Specificity of Academic Emotional Experiences“. In: The Journal of experimental Education 75.1, S. 5–29

Hu, A. (2003). Schulischer Fremdsprachenunterricht und migrationsbedingte Mehrsprachigkeit. Tübingen: Gunter Narr Verlag.

Kleppin, K. & G. Mehlhorn (2008). „Zum Stellenwert von Fehlern. Am Beispiel des Französischen und Russischen“. In: Praxis Fremdsprachenunterricht 4, S. 17–20 (siehe S. 73).

Pekrun, R. (2000). A Social-Cognitive, Control-Value Theory of Achievement Emotions. In J. Heckhausen (Hrsg.). Motivational psychology of human development: developing motivation and motivating development (S. 143–163). Oxford: North Holland.

Schwerdtfeger, I. C. (1997). Der Unterricht Deutsch als Fremdsprache: Auf der Suche nach den verlorenen Emotionen. Informationen Deutsch als Fremdsprache, 24.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
IG-Farben-Haus  
Raum 0.254





## ***Einfluss des akademischen Fähigkeitsselbstkonzeptes auf den Stress in einer Prüfungssituation***

Schülerinnen und Schüler empfinden starken Stress in der Schule (Cramer et al., 2011), über die beeinflussenden Faktoren ist allerdings nur wenig bekannt. Einige Untersuchungen deuten zwar darauf hin, dass bestimmte, eher übergeordnete Persönlichkeitsmerkmale mit Stressreaktionen in Verbindung stehen (z.B. Spangler et al., 2002), andere weisen jedoch keinen solchen Zusammenhang nach (z.B. Schoofs et al., 2008). Auch wenn die Befunde z.T. unklar sind, so liegt die Vermutung nahe, dass spezifischere Merkmale die Stressreaktion und das Stressempfinden in assoziierten Stresssituationen beeinflussen können. Ein solches Merkmal im Kontext von Schule ist das akademische Fähigkeitsselbstkonzept (FSK) in Bezug auf spezifische Unterrichtsfächer, welches die Einschätzung einer Person gegenüber den eigenen akademischen Fähigkeiten in dem jeweiligen Fach umfasst (Shavelson et al., 1976; Schöne et al., 2003). In der vorliegenden Studie untersuchten wir daher, ob die Höhe des akademischen FSKs in dem Fach Biologie das Stressempfinden und die physiologische Stressreaktion während eines Biologietestes beeinflusst.

Dazu untersuchten wir 52 Teilnehmer eines eintägigen molekularbiologischen Projektes im Schülerlabor der Ruhr-Universität Bochum. Es handelte sich hierbei um Oberstufenschüler aus Biologiekursen von 10 verschiedenen Schulen. Zunächst wurde das akademische FSK in Bezug auf das Fach Biologie und Deutsch erfasst. Dazu beantworteten die Schüler acht Items (z.B. "Ich gehöre in Biologie zu den Guten.") auf einer sechststufigen Likert Skala (DISK-Gitter; Rost et al., 2007). Zur Messung der physiologischen Stressreaktion wurde eine Speichelprobe entnommen, aus der die Konzentration des Stresshormons Cortisol analysiert wurde. Das subjektive Stressempfinden wurde mit einer Visual Analogue Scale (Folstein & Luria, 1973) erfasst. Dies ist eine 100 mm lange, horizontale Linie, die an ihrem linken Ende mit "kein Stress" und an dem rechten mit "maximaler Stress" übertitelt ist. Darauf kreuzten die Schülerinnen und Schüler an, wie gestresst sie sich gerade fühlten. Zur Erfassung des Basis-Stresspegels wurde sowohl die erste Speichelprobe, als auch die erste Visual Analogue Scale (VAS) vor der Konfrontation mit dem eigentlichen Stressor gesammelt. Danach wurden die Versuchspersonen einem Stressor in Form eines kurzen (10 Minuten) Biologietests ausgesetzt. Die Kontrollgruppe erhielt ebenfalls ein Aufgabenblatt, mit der Aufforderung, dass sie irgendetwas auf das Blatt schreiben sollen, z. B. über das letzte Buch, welches sie gelesen haben. Nach dem Test erfolgten die Sammlung der zweiten Speichelprobe und das Ausfüllen der zweiten VAS. Die Cortisol-Konzentration beider Speichelproben wurde mithilfe eines enzymgekoppelten Immunadsorptionstest (IBL Kit, Multiscan FC) analysiert.

Die Ergebnisse bestätigen zunächst, dass die Versuchspersonen mit einem signifikanten Anstieg des Cortisol auf den Test reagieren ( $t(27) = -3,78, p < 0,01$ ). Zudem gibt es in Bezug auf das akademische FSK in dem Fach Biologie eine signifikante Korrelation mit der Cortisolkonzentration ( $r = -0,531, p < 0,01$ ). Schüler mit höheren FSK Werten zeigen einen geringeren Cortisolanstieg als Schülern mit niedrigem FSK. Das subjektive Stressempfinden korreliert jedoch nicht mit der Höhe des akademischen FSKs ( $p > 0,05$ ). Zwischen dem subjektiven Stressempfinden, sowie der in dem Test erreichten Leistung und der Cortisolreaktion gibt es jedoch keinen signifikanten Zusammenhang ( $ps > 0,05$ ). Tendenziell zeigt sich jedoch eine Korrelation zwischen dem akademischen FSK und der Testleistung ( $p = 0,09$ ). Die Höhe des akademischen FSKs kann somit direkt das Ausmaß der physiologischen Stressreaktion beeinflussen.

Cramer, I., Paulus, P., Schumacher, L., Sieland, B. (2011). Depressive Stimmungen bei Schü-lerinnen und Schülern. Personale und schulische Risikofaktoren und Ansatzpunkte zur Prävention und Intervention. Zentrum für Angewandte Gesundheitswissenschaften (ZAG), Leuphana Universität, Lüneburg.

Folstein, M.F. & Luria, R. (1973). Reliability, validity, and clinical application of the visual analogue mood scale. *Psychological Medicine*, 3, 479–486.

Rost, D., Sparfeldt, J., & Schilling, S. (2007). DISK-GITTER mit SKSLF-8. Differentielles schulisches Selbstkonzept-Gitter mit Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten. Göttingen: Hogrefe.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2003). Das Fähigkeits-selbstkonzept und seine Erfassung. In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), *Diagnostik von Motivation und Selbstkonzept* (pp. 3-14). Göttingen: Hogrefe.

Schoofs, D., Hartmann, R., & Wolf, O.T. (2008). Neuroendocrine stress response to an oral academic examination: No strong influence of sex, repeated participation and personality traits. *Stress*, 11 (1), 52–61.

Shavelson, R.J., Hubner, J.J., & Stanton, G.C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46 (3), 407–441.

Spangler, G., Pekrun, R., Kramer, K., & Hofmann, H. (2002). Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress & Coping*, 15 (4), 413-432.

Mittwoch,

05.03.2014,

09:00 - 10:45,

IG-Farben-Haus

Raum 0.254



## **Gibt es geschlechtsspezifische Determinanten von Prüfungsangst?**

### **Eine Analyse handlungsregulativer und kognitiver Prädiktoren auf Basis der Kontroll-Wert-Theorie**

Theoretischer Hintergrund: Unter den kognitiven Emotionstheorien spezifiziert die Kontroll-Wert-Theorie die Qualität und Reihenfolge der Variablen, die leistungsbezogene Emotionen in Lehr- und Lernsettings vorhersagen (Pekrun, 2006; Pekrun et al., 2007). Prüfungsangst repräsentiert das mit Abstand am besten untersuchte Konstrukt; beispielsweise liegen überzeugende empirische Belege für die differenzielle Bedeutung unterschiedlicher Angstprädiktoren vor. Bisher fehlen jedoch Studien, welche die Annahmen der Kontroll-Wert-Theorie zur Wirkungskette dieser Prädiktoren näher untersuchen.

Fragestellung: Basierend auf den Annahmen der Kontroll-Wert-Theorie wurde in der vorliegenden Studie untersucht, ob geschlechtsspezifische Unterschiede in den Zusammenhangsmustern zwischen Prüfungsangst und ausgewählten Prädiktoren bestehen. Es wurde angenommen, dass (1) dispositionale Kontrollüberzeugungen, (2) antizipierter Misserfolg, und (3) eine hohe Wichtigkeit schulischen Erfolgs in der genannten Reihenfolge als Wirkungskette Prüfungsangst (Besorgtheit und Aufgeregtheit) vorhersagen. Für beide Geschlechter wurden strukturell äquivalente Muster erwartet, doch sollten die Zusammenhänge bei Mädchen stärker ausgeprägt sein (vgl. Frenzel et al., 2007; Pekrun et al., 2007; Selkirk et al., 2011).

Methode: Eine deutsche Stichprobe jugendlicher Schüler (N = 783; 419 Mädchen, 364 Jungen; Mage = 15.3; SD = .50) füllte Fragebögen zur schulbezogenen Selbstwirksamkeit, zu prüfungsbezogenen Kognitionen und zur Prüfungsangst aus. Die Schüler stammten aus 23 weiterführenden Schulen im Land Brandenburg, die per Zufall aus einem Pool von 124 Schulen ausgewählt worden waren. Auf Basis eines Multigruppenvergleichs wurden anhand von Strukturgleichungsmodellen mögliche Geschlechtsunterschiede im Zusammenhangsmuster der Variablen untersucht.

Ergebnisse: Die Ergebnisse bestätigen die Annahmen der Kontroll-Wert-Theorie zur sequentiellen Bedeutsamkeit der Prüfungsangstprädiktoren für Mädchen und Jungen: Wie erwartet waren die Beziehungsmuster zwischen Selbstwirksamkeit und prüfungsbezogenen Kognitionen nahezu äquivalent für beide Geschlechter. Besorgtheit und Aufgeregtheit wurden bei Mädchen und Jungen jedoch leicht unterschiedlich vorhergesagt. Über alle Variablen waren Zusammenhänge stärker ausgeprägt für Mädchen. Implikationen für die weitere Forschung zur sequentiellen Bedeutsamkeit von Emotionsprädiktoren werden diskutiert.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics – A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22, 497–514. doi: 10.1007/BF03173468

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315–341. doi: 10.1007/s10648-006-9029-9 .

Pekrun, R., Frenzel, A., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotions in education* (pp. 13–36). San Diego, CA: Academic Press.

Selkirk, L., Bouchey, H., & Eccles, J. (2011). Interactions among domain-specific expectancies, values, and gender: Predictors of test anxiety during early adolescence. *Journal of Early Adolescence*, 31, 361–389. doi: 10.1177/0272431610363156 .

## ***A Social-Emotional Enterprise: Positive Emotions in Higher Education Teaching***

Based on studies by Trigwell (2012) and Postareff and Lindblom-Ylänne (2011), the study investigates relations between university teachers' emotions and their teaching. More specifically, we hypothesized that their positive and negative emotions are elicited by the interaction with students. In addition, we researched consequences of these emotions for their teaching.

Study 1 investigates how the teachers themselves describe the sources of their emotions during teaching. N = 45 university teachers indicated in open-ended questions which situations during their teaching elicited emotions in them. The answers were analyzed using the qualitative content analysis and instances of each category were counted. Consistent with the prediction, the interaction with students was the main source of both positive emotions (71,8%) and negative emotions (61,5%) during teaching.

Study 2 investigates on the use of interactive, student-oriented teaching methods. We hypothesized that these teaching methods would be related to teachers' positive emotions. N = 37 university teachers indicated the frequency of specific (positive and negative) emotions and of student-oriented and teacher-centered teaching methods in their teaching. In addition, participants indicated their teaching approach (Approaches to Teaching Inventory, ATI, German: Braun & Hannover, 2008). Results show that only the

interactive, student-oriented, but not the teacher-centered methods predict the frequency of teachers' positive (but not negative) emotions. The more frequent teachers use student-oriented methods, the more often they experience positive emotions in their teaching. Partial correlations indicated that this relation is mediated by their teaching approach: The stronger their student-oriented teaching approach, the more often they experienced positive emotions.

These results confirm the importance of the interaction with

students for university teachers' (positive) emotions and the importance of positive emotions for student-oriented teaching.

Study 3 investigates correlates of university teachers' social skills. Based on studies that showed that the emotion regulation via "reappraisal" (versus emotion suppression) relates to their social skills (cf. Gross & John, 2003), we hypothesized that this emotions regulation strategy also related to their teaching approach. N = 145 university teachers indicated their emotions, their teaching approach (ATI) and their emotion regulation strategies (Emotion Regulation Questionnaire, German: Abler & Kessler, 2009). Results

show positive relations between reappraisal and a student-oriented teaching approach: The stronger the teachers focus on student learning in their teaching, the stronger their reappraisal strategy. There were no such relations to either emotion suppression nor to teacher-oriented approaches. These results confirm the close relations between how university teachers deal with their emotions relating to their social skills, and their student-orientation.

Taken together, the studies show that university teachers' emotions are elicited by the teacher-student interaction (studies 1 and 2) and that positive emotions as well as emotion regulation have strong implications for their teaching (studies 2 and 3). The author discusses implications for both educational development and further research.

Abler, B., & Kessler, H. (2009). Emotion Regulation Questionnaire – Eine deutschsprachige Fassung des ERQ von Gross und John. *Diagnostica*, 55(3), 144-152.

Braun, E., & Hannover, B. (2008). Zum Zusammenhang zwischen Lehr-Orientierung und Lehrgestaltung von Hochschuldozierenden und subjektivem Kompetenzzuwachs bei Studierenden. *Zeitschrift für Erziehungswissenschaft [Sonderheft 9: Perspektiven der Didaktik]*, 277-291.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348-362.

Postareff, L., & Lindblom-Ylänne, S. (2011). Emotions and confidence within teaching in higher education. *Studies in Higher Education*, 36(7), 799-813.

Trigwell, K. (2012). Relations between teachers' emotions in teaching and their approaches to teaching in higher education. *Instructional Science*, 40, 607-621.

## ***Wirtschaftswissen von deutschen und schweizerischen Lernenden in der Sekundarstufe II und in der Hochschule***

Chair(s): Doreen Holtsch (Universität Zürich), Esther Kaufmann (Universität Zürich / Universität Konstanz)

Diskutant/in: Susan Seeber (Universität Göttingen)

In der Literatur wird vielfach darauf hingewiesen, dass Wirtschaftswissen angesichts der wachsenden Bedeutung und der zunehmenden Komplexität globalisierter wirtschaftlicher Prozesse eine der Grundlagen für die soziale und wirtschaftliche Teilhabe des Einzelnen und damit für die Entwicklung der Gesellschaft als Ganzes darstellt (Eberle, 2006; Reetzmann et al., 2010).

Dabei herrscht weitgehender Konsens, dass der Schule – sei es im allgemein bildenden oder berufsbildenden Bereich – die zentrale Rolle beim Erwerb von Wirtschaftskenntnissen zukommt. Unter Rückgriff auf die Überlegungen von Dubs (1989) kann dabei mit Blick auf die schulisch vermittelte Wirtschaftsbildung zwischen einer wirtschaftsberuflichen Bildung und einer Bildung des allgemeinen Wirtschafts- und Gesellschaftsverständnisses unterschieden werden.

Während die wirtschaftsberufliche Bildung primär auf die Erlangung von Berufsfähigkeit und -fertigkeit abzielt, ist der zweite Zielaspekt auf den Aufbau von Wirtschaftswissen im Sinne einer Allgemeinbildung ausgerichtet. Typischerweise ist deren Vermittlung im allgemein bildenden Bereich angesiedelt, während die Vermittlung wirtschaftsberuflicher Kenntnisse in der kaufmännischen Ausbildung und im akademischen Bereich in wirtschaftswissenschaftlichen Studiengängen erfolgt. Allerdings existieren curriculare „Hybridaufträge“, welche insbesondere im Bereich der beruflichen Ausbildung vorzufinden sind.

Vor diesem knapp skizzierten Hintergrund widmet sich das Symposium der Anlage und den Befunden aktueller Studien aus der Schweiz und Deutschland zum Wirtschaftswissen von Lernenden in der kaufmännischen Ausbildung, im Gymnasium und in der Universität. Bei allen Gemeinsamkeiten zwischen dem schweizerischen und dem deutschen Bildungssystem werden bei genauerem Hinsehen vielfältige Unterschiede deutlich, die sich auch in der Vermittlung von Wirtschaftskenntnissen widerspiegeln. Das Symposium greift diese Unterschiede gezielt auf und liefert zugleich erste empirische Hinweise über den Kenntnisstand der Lernenden in beiden Ländern. Das Symposium thematisiert neben der Perspektive auf zwei Länder und verschiedene Bildungsstufen unterschiedliche methodische Herangehensweisen. Es werden neu entwickelte Paper-Pencil-Tests zur wirtschaftsbürgerlichen Kompetenz (Kaufmann et al.), zum betriebswirtschaftlichen Wissen (Jähmig) sowie technologiebasierte Zugänge mittels computergestützter Testungen (Holtsch et al.) vorgestellt.

Zuerst werden erste Befunde zur Entwicklung eines Tests zur Erfassung der wirtschaftsbürgerlichen Kompetenz basierend auf Erhebungen bei deutschen und schweizerischen Lernenden im kaufmännischen Bereich vorgestellt (Kaufmann et al.). Beim Vergleich deutscher und schweizerischer Gymnasiastinnen und Gymnasiasten lassen sich substantielle Bildungsgangunterschiede beobachten, die u.a. auf curriculare Effekte rückführbar sind (Schumann). Gleichwohl sind Eingangsvoraussetzungen in Mathematik, Deutsch und die kognitiven Fähigkeiten für das Wirtschaftswissen zu Beginn der Ausbildung von Bedeutung (Holtsch et al.). Abschließend wird in einem Beitrag zur Wirtschaftsbildung im Hochschulbereich in Deutschland gezeigt, dass sich Studierende der Wirtschaftspädagogik und Studierende der BWL zwar hinsichtlich ihres betriebswirtschaftlichen Wissens kaum unterscheiden, dass aber bspw. berufliche Erfahrungen förderlich wirken (Jähmig).

Eberle, F. (2006). Zur Bedeutung von Wirtschaft und Recht in der gymnasialen Bildung. *Gymnasium Helveticum*, 60(3), 16-23.

Dubs, R. (1989). Kaufmännische Berufsbildung für die Zukunft. Studie erstellt im Auftrage der Schweizerischen Bankgesellschaft und des Schweizerischen Kaufmännischen Verbandes. St.Gallen: IWP.

Retzmann, T., Seeber, G., Remmele, B. & Jongebloed, H.-C. (2010). Ökonomische Bildung an allgemeinbildenden Schulen. Bildungsstandards. Standards für die Lehrerbildung. Gutachten im Auftrag vom Gemeinschaftsausschuss der Deutschen Gewerblichen Wirtschaft unter Vorsitz des ZDH (2009/2010). Retrieved from <http://bankenverband.de/downloads/102010/gutachten-okonomische-bildung-an-allgemeinbildenden-schulen.pdf>.

Mittwoch,

05.03.2014,

09:00 - 10:45,

IG-Farben-Haus

Raum 1.411



Esther Kaufmann (Universität Zürich, Universität Konstanz), Stephan Schumann (Universität Konstanz),  
Nicole Ackermann (Universität Zürich), Andreas Jüttler (Universität Konstanz), Franz Eberle (Universität Zürich)

## ***Wirtschaftsbürgerliche Kompetenz von Lernenden in der kaufmännischen Ausbildung: Erste Befunde zur psychometrischen Güte und Dimensionalität eines neu entwickelten Erhebungsinstruments***

Im wirtschaftsbildenden Unterricht in der kaufmännischen Ausbildung wird traditionell zwischen einer wirtschaftsberuflichen Bildung und der Bildung eines allgemeinen Wirtschafts- und Gesellschaftsverständnisses unterschieden (Dubs, 1989). Mit Blick auf diesen zweiten Aspekt steht damit die Förderung einer so genannten wirtschaftsbürgerlichen Kompetenz im Zentrum. Die Referenzfigur ist dabei der kompetente „Wirtschaftsbürger“, der sich seiner Handlungsmöglichkeiten und gesellschaftlichen Verantwortung auf der Basis seiner Kenntnis gesamtwirtschaftlicher Zusammenhänge bewusst ist. In der kaufmännischen Ausbildung in der Schweiz erfährt die Förderung wirtschaftsbürgerlicher Kompetenz im Vergleich zu Deutschland eine vergleichsweise stärkere Beachtung. Bislang liegen jedoch keine Studien vor, welche sich der Struktur und Ausprägung wirtschaftsbürgerlicher Kompetenz intensiver genähert und diese zugleich empirisch geprüft haben. Hier setzt der vorliegende Beitrag an: Im Rahmen eines in das deutsche Verbundvorhaben CoBALIT (Winther, 2011) eingegliederten Forschungsprojekts wird der Modellierung und Messung wirtschaftsbürgerlicher Kompetenz nachgegangen (Eberle & Schumann, 2011). Ausgehend von den Überlegungen von Winther (2010) wird der Untersuchung ein Kompetenzstrukturmodell zugrunde gelegt, welches a) wirtschaftsbürgerliche und kaufmännische Kompetenz gemeinsam abbildet und b) in Anlehnung an Winther (2010) zwischen einer domänenspezifischen und einer domänenverbundenen Kompetenz unterscheidet. Im Falle der domänenspezifischen wirtschaftsbürgerlichen Kompetenz handelt es sich um das Vermögen, private, betriebswirtschaftliche und wirtschaftspolitisch-gesamtwirtschaftliche Problemstellungen in staatlichen und nichtstaatlichen sozialen Systemen zu verstehen, vorgeschlagene Lösungen zu beurteilen und für einfachere Probleme selbst Lösungen zu entwickeln. Bei der Lösung von derartigen domänenspezifischen Problemstellungen sind u.a. allgemeine sprachliche, mathematische und technische Fähigkeiten im Zusammenspiel mit betriebs- und volkswirtschaftlichem Wissen erforderlich. Dies wird in der Studie als domänenverbundene wirtschaftsbürgerliche Kompetenz verstanden.

Die Fragestellungen, denen im Beitrag nachgegangen wird, richten sich vor diesem Hintergrund im ersten Schritt auf der Überprüfung der psychometrischen Güte der Items eines neu entwickelten Erhebungsinstruments. Im zweiten Schritt werden die Befunde zur Prüfung der angenommenen Kompetenzstruktur vorgelegt. Die verwendeten Daten beruhen auf den beiden in der Schweiz und Deutschland durchgeführten Pilotuntersuchungen (N = 562). In der Schweiz wurden dabei Auszubildende im zweiten Lehrjahr untersucht. In Deutschland erfolgte die Untersuchung in verschiedenen kaufmännischen Ausbildungsgängen des zweiten und dritten Lehrjahrs sowie ergänzend in Klassen des Wirtschaftsgymnasiums. Zur Messung der wirtschaftsbürgerlichen Kompetenz wurden sechs Anforderungssituationen entwickelt, welche sich auf authentische und aktuelle private, betriebswirtschaftliche und wirtschaftspolitisch-gesamtwirtschaftliche Problemstellungen beziehen (z.B. Altersvorsorge, Jugendverschuldung). Jede dieser Problemsituationen wird textuell mit einem modifizierten Zeitungsartikel eingeleitet. Dem schließen sich geschlossene und offene Items an. Die Erhebung erfolgte mittels eines Multi-Matrix-Designs mit randomisierter Zuteilung der Testhefte.

Die Auswertungen erfolgen auf Grundlage der klassischen und probabilistischen Testtheorie. Die Analysen beinhalten ebenfalls Befunde zum Differential Item Functioning (DIF) für das jeweilige Land, die Bildungsgangzugehörigkeit sowie für das Geschlecht. Die Prüfung der Dimensionalität wird mittels des Programms ConQuest (Wu, Adams, Wilson, & Haldane, 2007) sowie mittels Strukturgleichungsmodellierungen vorgenommen. Gegenwärtig befinden wir uns noch in der Auswertungsphase, so dass aktuell noch keine Befunde berichtet werden können.

Dubs, R. (1989). Kaufmännische Berufsbildung für die Zukunft. Studie erstellt im Auftrage der Schweizerischen Bankgesellschaft und des Schweizerischen Kaufmännischen Verbandes. St.Gallen: IWP.

Eberle, F. & Schumann, S. (2011). Modellierung und Messung wirtschaftsbürgerlicher Kompetenz. Unv. Antrag an das BBT. Zürich/Fribourg: Universität Zürich/Universität Fribourg.

Winther, E. (2010). Kompetenzmessung in der beruflichen Bildung. Bielefeld: Bertelsmann.

Winther, E. (2011). Kompetenzen in der kaufmännischen Berufsbildung (Competencies in the Field of Business and Administration, Learning, Instruction, and Transition – CoBALIT). Unv. Antrag an das BMBF. Paderborn: Universität Paderborn.

Wu, M. I., Adams, R. J., Wilson, M., & Haldane, S. (2007). ConQuest Version 2. [Computer software]. Assessment Systems Corporation, St. Paul, MN.

**GEBF** TAGUNG  
2014

E13 a

Symposiums  
beitrag

Mittwoch,  
05.03.2014,

09:00 - 10:45,  
IG-Farben-Haus

Raum 1.411

Wirtschaftswissen von deutschen und schweizerischen Lernenden





## **Wirtschaftskenntnisse von deutschen und schweizerischen Gymnasiastinnen und Gymnasiasten im Vergleich**

Angesichts der wachsenden Bedeutung und zunehmenden Komplexität globalisierter wirtschaftlicher Prozesse wird vielfach angenommen, dass ökonomische Kompetenzen Jugendlicher eine der Grundlagen für deren soziale und berufliche Teilhabe und damit langfristig für die Entwicklung der Gesellschaft als Ganzes darstellen (Seeber et al., im Druck; Schumann & Eberle, im Druck). Vor diesem Hintergrund mehren sich in den letzten Jahren in Deutschland Initiativen und öffentliche Forderungen nach einer stärkeren curricularen Berücksichtigung ökonomischer Bildung in den Schulen beider Sekundarstufen. Im Unterschied zu den meisten Bundesländern in Deutschland ist wirtschaftsbildender Unterricht in der Schweiz zumindest im Gymnasium ein seit 1995 etabliertes und für alle Schülerinnen und Schüler geltendes Obligatorium, welches je nach Kanton mit zumindest einer Wochenstunde (sog. Grundlagenfach) und im Falle einer Profilbildung in diesem Bereich mit bis zu 5 Lektionen pro Woche unterrichtet wird (sog. Schwerpunktfach). Zugleich existiert in der Schweiz mit der Berufsmaturität ein Bildungsgang, welcher im Rahmen von sechs möglichen Richtungen und in Kombination mit einer beruflichen Ausbildung zu einer fachgebundenen Hochschulreife führt. Auch in den Berufsmaturitätsschulen ist wirtschaftsbildender Unterricht obligatorisch, wobei der größte Lektionsumfang in der kaufmännischen Richtung anzutreffen ist.

Verschiedene Studien liefern Hinweise darauf, dass Schülerinnen und Schüler an Gymnasien insbesondere in Deutschland substanzielle Defizite im Bereich ökonomischer Kompetenzen aufweisen (Beck & Krumm, 1998; Müller et al., 2007). Allerdings ist der wissenschaftliche Kenntnisstand bezüglich der Kompetenzausprägungen jenseits der von Beck und Krumm (1998) im Rahmen der Anfang der 1990er Jahre in Deutschland, Österreich und der Schweiz durchgeführten Normierung des Wirtschaftskundlichen Bildungstests (WBT) keinesfalls als ausreichend zu bezeichnen. Anders gestaltet sich die Situation für die Schweiz, für die Eberle und Schumann (2013) eine für die Deutschschweiz repräsentative Untersuchung ökonomischer Kompetenzen durchgeführt haben. Neben einem eigens konstruierten Leistungstest wurden in der Untersuchung 11 WBT-Items eingesetzt, welchen einen Vergleich mit den Ergebnissen anderer Studien ermöglichen.

Im Beitrag steht vor diesem Hintergrund die Frage im Mittelpunkt, welche Unterschiede sich zwischen deutschen und schweizerischen Lernenden an Gymnasien und Berufsmaturitätsschulen (BMS) zeigen und inwieweit mögliche Unterschiede in den Leistungsausprägungen durch Profilbildungen innerhalb der jeweiligen Schulformen erklärt werden können. Zur Beantwortung dieser Frage wurden die Datensätze aus der WBT-Normierung (Beck & Krumm, 1998), der Studie von Müller et al. (2007) zu sächsischen Gymnasiastinnen und Gymnasiasten sowie aus der Studie von Eberle & Schumann (2013) zusammen geführt und für die oben erwähnten 11 WBT-Items ausgewertet. Die Stichprobe umfasst demnach  $N = 4.178$  Lernende. Berücksichtigt werden muss die eingeschränkte Messgenauigkeit des 11 Items umfassenden Konstrukts (Cronbach's Alpha = 0.64). Die Gruppenvergleiche erfolgen auf Rohwertebene mittels varianzanalytischer Verfahren. Die Ergebnisse zeigen, dass die Deutschschweizer Lernenden mit dem Schwerpunkt im wirtschaftsbildenden Bereich im Durchschnitt die höchsten Werte erzielen ( $M = 7.38$ ), während die Lernenden im Grundlagenfach mit einem Mittelwert von 6.00 signifikant niedrigere Werte erzielen. Die BMS-Lernenden der kaufmännischen Richtung liegen mit 6.66 Punkten in der Mitte zwischen diesen beiden Gruppen, während BMS-Lernenden der anderen Richtungen auf dem Niveau der Gymnasiastinnen und Gymnasiasten im Grundlagenfach verortet werden können. Der Vergleich mit Deutschland zeigt einen deutlichen Rückstand der sächsischen Gymnasiastinnen und Gymnasiasten ( $M = 5.17$ ). Die Daten der Normierungstichprobe des WBT zeigen für Lernende in allgemein bildenden Gymnasien in Hessen, Bayern, NRW und im Saarland leicht höhere Werte ( $M = 6.18$ ) und für Lernende im Fachgymnasium Wirtschaft Testwerte, die sich im Bereich schweizerischer BMS-Lernender der kaufmännischen Richtung bewegen ( $M = 6.53$ ). Insgesamt lassen sich – auch unter Kontrolle der Geschlechtszugehörigkeit – substanzielle Bildungsgangunterschiede beobachten, die in augenscheinlich auf curriculare Effekte rückführbar sind.

Die Ergebnisse des vorgelegten Vergleichs sind zweifellos von diversen methodischen Limitationen begleitet (Messfehler, Vergleichbarkeit der Stichproben, Kohorteneffekte). Dennoch dürfen sie als empirischer Hinweis auf die Wirksamkeit des Wirtschaftsunterrichts interpretiert werden. Zugleich eröffnen sie das Feld für notwendige weitere Studien zum Gegenstand.

Beck, K. & Krumm, V. (1998). Wirtschaftskundlicher Bildungs-Test (WBT). Handanweisung. Göttingen, Bern, Toronto, Seattle: Hogrefe Verlag für Psychologie.

Eberle, F. & Schumann, S. (2013). Ökonomische und weitere Kompetenzen von Deutschschweizer Berufsmaturanden und Gymnasiasten im Vergleich. In: *Gymnasium Helveticum*, 67(1), 18-21.

Müller, K., Fürstenau, B. & Witt, R. (2007). Ökonomische Kompetenz sächsischer Mittelschüler und Gymnasiasten. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 103, 227- 247.

Schumann, S. & Eberle, F. (im Druck). Ökonomische Kompetenzen von Lernenden am Ende der Sekundarstufe II. In: E. Winther & M. Prenzel (Hrsg.), *Perspektiven der empirischen Berufsbildungsforschung - Kompetenz und Professionalisierung*. Beiheft der Zeitschrift für Erziehungswissenschaft.

Seeber, S., Schumann, S. & Nickolaus, R. (im Druck). Ökonomische Kompetenzen: Konzeptuelle Grundlagen und empirische Befunde. Erscheint in Weißeno, G. & Schelle, C. (Hrsg.): *Was geschieht im Fachunterricht Politik, Geschichte, Geographie, Wirtschaft?* VS-Verlag.

## Zusammenhänge zwischen Eingangsvoraussetzungen und Wirtschaftswissen zu Beginn der kaufmännischen Berufsausbildung

In der Schweiz absolvieren ca. zwei Drittel der Jugendlichen eine duale Ausbildung (Staatssekretariat für Bildung, Forschung und Innovation, 2013). Ungefähr 11'000 Lernende beginnen jährlich eine kaufmännische Ausbildung, die somit zu den meist gewählten gehört. Die kaufmännischen Lernenden können ihre Ausbildung unter anderem im E-Profil mit dem Eidgenössischen Fähigkeitszeugnis (EFZ) oder im M-Profil mit der Berufsmaturität abschliessen. Das EFZ ist vergleichbar mit der dualen Berufsausbildung in Deutschland. Die auf den ersten Blick mit Deutschland vergleichbare Ausbildung unterscheidet sich allerdings im Detail. Erstens, die kaufmännische Ausbildung folgt in der Schweiz dem Allbranchen-Konzept, bei dem Lernende der 21 kaufmännischen Branchen in einer Berufsschulklasse lernen. Zweitens, in der Berufsschule soll im Lernbereich resp. Fach „Wirtschaft und Gesellschaft“ das domänenspezifische kaufmännische Wissen und Können der Lernenden gefördert werden. Neben der Vorbereitung auf die Berufstätigkeit im kaufmännischen Bereich verfolgt „Wirtschaft und Gesellschaft“ auch die allgemeine Zielsetzung des wirtschaftsbürgerlichen Wissens. Dies bedeutet, dass Wirtschaftswissen bei Lernenden hinsichtlich einer kaufmännischen und einer wirtschaftsbürgerlichen Dimension modelliert und erfasst werden muss.

Im Leading House-Projekt LINCA wird das kaufmännische und das wirtschaftsbürgerliche Wissen der Lernenden im E- und M-Profil längsschnittlich zu Beginn, während und am Ende der Ausbildung erfasst. Darüber hinaus werden kognitive und motivationale Eingangsvoraussetzungen erhoben. Es wird davon ausgegangen, dass unter anderem mathematische Kenntnisse und Deutschkenntnisse sowie kognitive Grundfähigkeiten die Bewältigung von domänenspezifischen kaufmännischen und allgemeinen Wirtschaftsaufgaben unterstützen (vgl. Seeber et al., 2010, p. 4). Daher werden in der Schweiz in LINCA empirische Analysen zum Zusammenhang zwischen Eingangsvoraussetzungen in Deutsch, Mathematik und kognitiver Grundfähigkeit und dem Wirtschaftswissen im obigen Sinne von Auszubildenden durchgeführt.

Im Beitrag werden zwei Forschungsfragen verfolgt:

1. Welche Zusammenhänge lassen sich zwischen kognitiven Eingangsvoraussetzungen und Wirtschaftswissen zu Beginn der kaufmännischen Berufsausbildung beobachten?
2. Inwieweit existieren hinsichtlich der Eingangsvoraussetzungen und dem Wirtschaftswissen bereits zu Beginn der Ausbildung Unterschiede in Bezug auf die Branchen- und Profilverzweigung?

Zur Erfassung des domänenspezifischen kaufmännischen Wissens wurden ausgewählte und für die Schweiz adaptierte Aufgaben der Computersimulation ALUSIM (Achtenhagen & Winther) genutzt. Die Aufgaben umfassen einen Bestellvorgang, die in Kenntnissetzung eines Kunden über eine Lieferverzögerung sowie das Einholen eines Angebots. Um das wirtschaftsbürgerliche Wissen zu erfassen, wurden OEKOMA-Items von Schumann et al. (2010) genutzt, die ebenfalls im Symposium vorgestellt werden. Die Eingangsvoraussetzungen in Mathematik, Deutsch und hinsichtlich kognitiver Grundfähigkeiten wurden mittels elaborierter Instrumente aus ULME (Lehmann & Seeber, 2007) und EVAMAR II (Eberle, 2008), TIMSS II und TIMSS III (z. B. Baumert et al., 1998) sowie KFT (Heller & Perleth, 2000) erfasst.

Die im Beitrag präsentierten Daten basieren auf der Befragung von 1.761 Auszubildenden, die im August 2012 ihre kaufmännische Ausbildung begannen. Die erste Befragung fand vom November bis Dezember 2012 in 36 Schulen und insgesamt 85 Klassen in der Deutschschweiz statt. Erste Analysen zeigen, dass die Eingangsvoraussetzungen in Mathematik, Deutsch sowie die kognitiven Grundfähigkeiten mit dem kaufmännischen sowie wirtschaftsbürgerlichen Wissen positiv korrelieren. Die höchste Korrelation zeigt sich für das wirtschaftsbürgerliche Wissen und Deutschkenntnisse ( $r = .467^{**}$ ). Erste Analysen zeigen sowohl Branchen- als auch Profilverzweigung Unterschiede hinsichtlich der Eingangsvoraussetzungen. So lässt sich sagen, dass Lernende im M-Profil durchweg bessere Testleistungen zeigen als Lernende des E-Profiles. Zudem schneiden Lernende der Bank-Branche in den einzelnen Leistungstest signifikant besser ab im Vergleich zu den restlichen Lernenden. Dies deutet auf eine branchenspezifische Selektion bereits zu Beginn der Ausbildung hin. Im Beitrag werden diese Unterschiede systematisch und im Detail anhand von t-Tests und ANOVA vorgestellt. Abschliessend werden die Zusammenhänge zwischen den Eingangsvoraussetzungen und dem Abschneiden bei kaufmännischen und wirtschaftsbürgerlichen Wissensfragen anhand eines Strukturgleichungsmodells dargestellt. Im Projektfortgang folgen drei weitere Messzeitpunkte, die über die Wirkung der Eingangsvoraussetzungen auf die Entwicklung des Wirtschaftswissens Aufschluss geben können.

Achtenhagen, F., & Winther, E. (2009). Konstruktvalidität von Simulationsaufgaben: Computergestützte Messung berufsfachlicher Kompetenz - am Beispiel der Ausbildung von Industriekaufleuten. Bericht an das Bundesministerium für Bildung und Forschung. Göttingen: Georg-August-Universität Göttingen, Seminar für Wirtschaftspädagogik.

Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., ... Günther, W. (1998). Testaufgaben Mathematik. TIMSS 7./8. Klasse (Population 2). Materialien aus der Bildungsforschung, Band 60. Berlin: Max-Planck-Institut für Bildungsforschung.

Eberle, F., Gehrler, K., Jaggi, B., Kottonau, J., Oepke, M., & Pflüger, M. (2008). Evaluation der Maturitätsreform 1995 (EVAMAR). Schlussbericht der Phase II. Zürich: Universität Zürich, Institut für Gymnasial- und Berufspädagogik.

Heller, K. & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision (KFT 4-12+ R). Göttingen: Hogrefe.

Lehmann, R. & Seeber, S. (2007). ULME III. Untersuchungen von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Hamburg: HIBB.

Schumann, S., Eberle, F., Oepke, M., Pflüger, M., Gruber, C., Stamm, P., & Pezzotta, D. (2010). Inhaltsauswahl für den Test zur Erfassung ökonomischen Wissens und Könnens im Projekt „Ökonomische Kompetenzen von Maturandinnen und Maturanden (OEKOMA)“. Zürich: Universität Zürich, Institut für Gymnasial- und Berufspädagogik.

Seeber, S., Nickolaus, R., Winther, E., Achtenhagen, F., Breuer, Klaus, Frank, Irmgard, ... Zöllner, Arnulf. (2010). Kompetenzdiagnostik in der Berufsbildung: Begründung und Ausgestaltung eines Forschungsprogramms. Berufsbildung in Wissenschaft und Praxis. Beilage, 1-15.

Staatssekretariat für Bildung, Forschung und Innovation (SBFI). (2013). Berufsbildung in der Schweiz – Fakten und Zahlen. Download unter: <http://www.berufsbildung.ch/dyn/2777.aspx> (Stand: 31. Oktober 2013)



## **Betriebswirtschaftliches Wissen von Studierenden in Deutschland: Determinanten der Testleistung und Subgruppenvergleiche**

Kompetenzorientierung“ und „Outcomeorientierung“ sind zwei zentrale Begriffe, die durch den Bologna-Prozesse (Bologna Declaration, 1999) Einzug in die deutsche Hochschullandschaft gehalten haben. Die Frage, was an deutschen Universitäten gelehrt und gelernt wird, rückt insbesondere wirtschaftswissenschaftliche Studiengänge immer mehr in das Interesse der Öffentlichkeit. Im Gegensatz zum öffentlichen Interesse liegen aus wissenschaftlicher Perspektive zu Fragen des Wissenserwerbs und der Wissensentwicklung im tertiären Bildungssektor nur wenige belastbare Ergebnisse vor. Universitäre Lehr-Lernprozesse und deren Ergebnisse gelten in der deutschen Bildungsforschung als „Black Box“, mit deren intensiven Betrachtung erst jüngst begonnen wurde (Bülow-Schramm & Braun, 2013). Die Zurückhaltung bei der Erforschung von Bildungsergebnissen im Hochschulsektor ist nicht zuletzt einem Mangel an validen Instrumenten und einem damit einhergehenden Mangel an grundlegenden Konzepten der Testentwicklung und -auswertung in diesem Bereich geschuldet (Blömeke, 2013).

Das Ziel des Beitrags ist, die Entwicklung und die Validierung eines Tests zu beschreiben, der betriebswirtschaftliches Wissen auf Bachelorniveau erfasst. Dafür wurden in Anlehnung an Arbeiten aus dem Bereich der Berufs- und Wirtschaftspädagogik und aus der Personalauswahl Items entwickelt und unter Bezugnahme auf item-response-theoretische Modelle quantitativ-empirisch analysiert und ausgewertet. Im Rahmen der Präsentation wird zwei zentralen Forschungsfragen nachgegangen:

- (1) Welche Determinanten der Testleistung in betriebswirtschaftlichen Wissenstests können identifiziert werden?
- (2) Wie unterscheidet sich die Testleistung von Studierenden je nach Studiengang, Geschlecht und bildungsbiographischen Merkmalen?

Das Vorgehen bei der Erstellung der Items orientierte sich am Ideal der deduktiven Testentwicklung (Hartig & Jude, 2007). Der Abgleich zwischen den theoretisch intendierten und den tatsächlichen Testeigenschaften erfolgte quantitativ-empirisch. Dafür wurden die Daten von ca. 500 Studierenden erhoben und nach aktuellen psychometrischen Standards unter Nutzung des Rasch-Modells ausgewertet. Potenzielle Determinanten der Testleistung wurden an einer Substichprobe von ca. 350 Studierenden durch eine multivariate Regressionsanalyse auf Signifikanz geprüft. Der Vergleich der Testleistung zwischen verschiedenen Studierendengruppen erfolgte unter Nutzung von ANOVA und t-Test. Die aus der Literatur bekannten Determinanten der Testleistung in wirtschaftswissenschaftlichen Wissenstests (kaufmännische Ausbildung, Note der Hochschulzugangsberechtigung, Geschlecht, Anzahl- und Note der besuchten Grundlagenmodule) erwiesen sich in einer multiplen OLS-Regression als signifikant in vorhersagekonformer Richtung. Besonders hervorzuheben ist dabei der Leistungsvorsprung von Studierenden mit kaufmännischer Ausbildung und der männlichen Studierenden. Von den als Prädiktor ausgewählten individuellen studienbezogenen Faktoren, die potenziell einer Intervention von Seiten der Lehrenden an Hochschulen zugänglich wären (fachliches Interesse, Lernstrategien, akademisches Selbstkonzept und Leistungsmotivation), zeigte sich nur das Interesse an betriebswirtschaftlichen Fachinhalten als bedeutsamer Prädiktor für die Testleistung. Mit dem fachspezifischen Studieninteresse wurde zusätzlich eine Variable mit der Testleistung in Verbindung gebracht, die potenziell durch Interventionen von Seiten der Hochschule zum Positiven veränderbar ist. Der Versuch, durch veränderbare personenbezogene Konstrukte wie Leistungsmotivation und Lernstrategien weitere Erklärungsfaktoren für die Testleistung zu identifizieren, gelang nicht. Hier müssten möglicherweise alternative Konstruktoperationalisierungen vorgenommen werden.

Die Items des Tests wurden gezielt so entwickelt, dass sie sowohl für Studierende der Betriebswirtschaftslehre als auch für Studierende der Wirtschaftspädagogik ein geeignetes Maß betriebswirtschaftlichen Wissens darstellen sollten. Wie von einem in diesem Sinne konzipierten Test zu erwarten war, schnitten Studierende der Betriebswirtschaftslehre und Studierende der Wirtschaftspädagogik in dem Test etwa gleich gut ab. Studierende der Volkswirtschaft und anderer Studiengänge wiesen allerdings signifikant schlechtere Testleistungen auf. Zwischen Studierenden im Bachelorstudium und Studierenden im Masterstudium lag wie erwartet kein signifikanter Unterschied bezüglich der Testleistung vor. Im Rahmen der Testauswertung wurden Erkenntnisse über die Struktur und die Determinanten betriebswirtschaftlichen Wissens gewonnen. Ein weiteres Ziel des Vortrages ist es, methodische Besonderheiten der Erfassung von Lernergebnissen im Hochschulsektor herauszuarbeiten und somit die Grundlage für weitere Forschung in diesem Bereich auszubauen.

Blömeke, S. (2013). Ja, aber ... – Lehren aus TEDS-M und KoKoHs für eine Teilnahme an AHELO. In Hochschul Informations System GmbH (Hrsg.), AHELO goes Germany? Dokumentation des GfHf- & HIS-HF-Workshops (S. 13–20). Zugriff am 12.09.2013 [http://www.his.de/pdf/pub\\_fh/fh-201302.pdf](http://www.his.de/pdf/pub_fh/fh-201302.pdf)

Bologna Declaration. (1999). The Bologna Declaration of 19 June 1999. Joint declaration of the European Ministers of Education (European Union, Hrsg.), Brussels. Zugriff am 12.09.2013 [http://www.bologna-berlin2003.de/pdf/bologna\\_declaration.pdf](http://www.bologna-berlin2003.de/pdf/bologna_declaration.pdf)

Bülow-Schramm, M. & Braun, E. (2013). Einleitung. In Hochschul Informations System GmbH (Hrsg.), AHELO goes Germany? Dokumentation des GfHf- & HIS-HF-Workshops (S. 1-8). Zugriff am 12.09.2013 [http://www.his.de/pdf/pub\\_fh/fh-201302.pdf](http://www.his.de/pdf/pub_fh/fh-201302.pdf)

Hartig, J. & Jude, N. (2007). Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In J. Hartig & E. Klieme (Hrsg.), Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik (Bildungsforschung, Bd. 20) (S. 17–36). Berlin: BMBF.



## **Zur (Nicht-)Modellierung lokaler Abhängigkeiten in Messmodellen: Weshalb der Modellfit kein geeignetes Kriterium für die Modellwahl ist**

**Chair(s):** Alexander Robitzsch (Bundesinstitut für Bildungsforschung, Innovation & Entwicklung)

In Kompetenztests werden Items häufig in sogenannten Itembündeln (Testlets) administriert. Im Allgemeinen sind dann die Items innerhalb von Testlets stärker voneinander abhängig als Items zwischen Testlets. In der psychometrischen Literatur wird argumentiert (Wainer, Bradlow & Wang, 2007), dass durch die Schachtelung der Items in Testlets die Eigenschaft der lokalen stochastischen Unabhängigkeit in Messmodellen (Item-Response-Modelle oder konfirmatorische Faktorenanalysen) verletzt sei und daher diese Abhängigkeit in einem psychometrischen Modell berücksichtigt werden müsse. Erst wenn die Items lokal stochastisch unabhängig sind (d.h. das Messmodell gilt), würden unverzerrte Schätzungen von Itemparametern und der Reliabilität resultieren. Im vorliegenden Beitrag soll zunächst mit Hilfe eines kleinen Zahlenbeispiels herausgearbeitet werden, dass die lokale stochastische Unabhängigkeit zwar eine hinreichende, aber keine notwendige Eigenschaft zur Gewinnung unverzerrter Schätzer darstellt (siehe z. B. Lucke, 2005). Dabei wird darauf hingewiesen, dass sich bei einer eindimensionalen Skalierung eines Tests positive und negative lokale Abhängigkeiten (residuale Korrelationen) näherungsweise aufheben müssen. D.h. das Auftreten positiver Abhängigkeiten bedeutet immer das gleichzeitige Auftreten negativer Abhängigkeiten. Ohne eine nicht empirisch bestimmbare Annahme hinsichtlich der mittleren lokalen Abhängigkeit (d. h. mittlere Residualkorrelation) ist das Messmodell nicht identifiziert. Es wird gezeigt, dass diese Annahme starke Konsequenzen für die Bestimmung der Itemparameter und Reliabilität besitzt (Westfall et al., 2012).

Anhand des HAMLET-Tests (Lehmann, Peek & Poerschke, 2006) zur Erfassung der Lesekompetenz sollen vor dem Hintergrund dieser Überlegungen verschiedene Skalierungsvarianten bei Vorliegen von Testlets illustriert werden. Wir unterscheiden dabei vier verschiedene Modellklassen.

1. Testletmodelle oder Bifaktormodelle spezifizieren für Testlets weitere unkorrelierte Faktoren (Wainer et al., 2007; Reise, 2012). Bei der Reliabilitätsbestimmung kann dann unterschieden werden, ob Testlet-Faktoren als Teil der Konstruktvarianz oder der Fehlervarianz anzusehen sind.
2. Marginale Modelle nehmen korrelierte Messfehler an, die lokale Abhängigkeiten modellieren (z. B. Braeken, 2011).
3. In der Generalisierbarkeitstheorie stellen Items eine (ggf. hypothetische) Stichprobe eines Behavior Domains dar (Brennan, 2001; McDonald, 2003). Dabei spielt der Modellfit keine Rolle, da die Annahme einer (mittleren) Unabhängigkeit des Samplings von Items ausschließlich substanztheoretisch zu legitimieren ist.
4. Die Ignorierung lokaler Abhängigkeiten führt zu einer eindimensionalen Skalierung mit unzureichender Modellpassung, so dass Modelle mit „approximativer lokaler Unabhängigkeit“ entstehen (Bai & Li, 2012). Es können allerdings Bedingungen identifiziert werden, in denen trotz Verletzung der Modellannahme der lokalen Unabhängigkeit nicht mit (deutlich) verzerrten Parameterschätzungen zu rechnen ist.

Anhand des HAMLET-Tests wird illustriert, dass mit diesen vier Modellen teilweise sehr unterschiedliche Reliabilitätsschätzungen resultieren (zwischen .59 und .73). Es wird argumentiert, dass Messmodelle nicht aufgrund eines Modellfits ausgewählt werden sollten, sondern die interessierende Parameterinterpretation oder die prädiktive Validität das Hauptkriterium für die Modellwahl darstellen sollten (Ballou, 2009). Da eindimensionale Messmodelle mit lokalen Abhängigkeiten (marginale Modelle) äquivalent zu mehrdimensionalen Messmodellen (Testletmodelle) sind, lassen sich somit für denselben Test äquivalente Messmodelle mit gleichem Modellfit schätzen, die aber verschiedene Interpretationen von Itemparametern und verschiedene Reliabilitäten besitzen (Ip, 2010; Westfall et al., 2012). Daher impliziert ein guter Modellfit nicht, dass mit keiner Verzerrung von Modellparametern zu rechnen ist. Im vorliegenden Beispiel würden wir für den Zweck der psychometrischen Beurteilung von Items das marginale Modell präferieren, da bei Testletmodellen nicht mehr dieselbe Interpretation von Itemparametern möglichst ist wie in eindimensionalen Modellen. So sind z.B. die Itemparameter von Items verschiedener Testlets nicht mehr direkt miteinander vergleichbar (Ip, 2010). Abschließend wird auf das Verhältnis von Reliabilität und Validität eingegangen. Es wird festgestellt, dass typischerweise bei Einbeziehung von Testlets in Lesekompetenztests aufgrund (möglicher) lokaler Abhängigkeiten die Reliabilität sinkt, allerdings die Validität ansteigt (siehe auch Feldt, 1997; Kane, 1982). Aufgrund dieser Beobachtung würde das Konzept der Generalisierbarkeit ein besseres Kriterium zur Beurteilung der psychometrischen Qualität von Tests mit Testlets darstellen, da es sowohl die Reliabilität als auch die Validität eines Tests umfasst (Kane, 1982; McDonald, 2003).

- Bai, J., & Li, K. (2012). Statistical analysis of factor models of high dimension. *The Annals of Statistics*, 40, 436-465.
- Ballou, D. (2009). Test scaling and value-added measurement. *Education Finance and Policy*, 4, 351-383.
- Braeken, J. (2011). A boundary mixture approach to violations of conditional independence. *Psychometrika*, 76, 57-76.
- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer.
- Feldt, L. S. (1997). Can validity rise when reliability declines? *Applied Measurement in Education*, 10, 377-387.
- Ip, E. H. (2010). Empirically indistinguishable multidimensional IRT and locally dependent unidimensional item response models. *British Journal of Mathematical and Statistical Psychology*, 63, 395-416.
- Kane, M. T. (1982). A sampling model for validity. *Applied Psychological Measurement*, 6, 125-160.
- Lehmann, R.H., Peek, R., & Poerschke, J. (2006). *HAMLET 3-4. Hamburger Lesetest für 3. und 4. Klassen*. Göttingen: Hogrefe.
- Lucke, J. F. (2005). "Rassling the hog": The influence of correlated item error on internal consistency, classical reliability, and congeneric reliability. *Applied Psychological Measurement*, 29, 106-125.
- McDonald, R. P. (2003). Behavior domains in theory and in practice. *Alberta Journal of Educational Research*, 49, 212-230.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667-696.
- Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. Cambridge University Press.
- Westfall, P. H., Henning, K. S., & Howell, R. D. (2012). The effect of error correlation on interfactor correlation in psychometric measurement. *Structural Equation Modeling*, 19, 99-117.



## Ein längsschnittlicher erklärender Mehrebenen-IRT Ansatz zur Modellierung von Instruktionssensitivität

Testergebnisse von Schülerinnen und Schülern werden regelmäßig auf einen mehr oder weniger guten Unterricht zurückgeführt (Creemers & Kyriakides, 2008). Dieses Vorgehen impliziert, dass die eingesetzten Instrumente sensitiv auf den Unterricht reagieren, also dazu in der Lage sind, Effekte des Unterrichts aufzufangen. Jedoch wird diese Annahme selten empirisch überprüft.

Die psychometrische Eigenschaft eines Tests beziehungsweise eines Items, sensitiv auf Unterricht zu reagieren, wird als Instruktionssensitivität bezeichnet (Polikoff, 2010). Doch obwohl zahlreiche statistische Operationalisierungsansätze vorgeschlagen wurden, ist der Zusammenhang mit Unterrichtsmerkmalen nach wie vor unklar. Unsere Studie zielt darauf ab, einen modellbasierten Ansatz zur Evaluation der Instruktionssensitivität von Leistungstestaufgaben vorzuschlagen, um die statistische Sensitivität von Aufgaben zu quantifizieren und deren Zusammenhang mit empirischen Unterrichtsmaßen zu untersuchen. Sensitivität wird zunächst als statistische Variabilität in Itemschwierigkeiten aufgrund des Unterrichtskontextes begriffen. Instruktionssensitivität entspricht dann dem Anteil an der statistischen Variabilität, der auf Unterrichtsmerkmale zurückgeht.

In einem längsschnittlichen Mehrebenen-IRT Modell mit Antworten genestet in Zeitpunkten, Items, Schülern und Klassen werden für jeden Messzeitpunkt die mittlere Klassenfähigkeit, die individuelle Abweichung der Fähigkeit einer Person vom Klassenmittelwert sowie die klassenspezifische Schwierigkeit eines Items zum ersten Messzeitpunkt geschätzt. Für Messzeitpunkte  $t > 1$  wird die Veränderung der Itemschwierigkeit vom vorhergehenden Messzeitpunkt als klassenspezifische Prätest-Posttest-Differenzen (PPD) spezifiziert. Sowohl klassenspezifische Itemschwierigkeiten als auch PPDs werden als normalverteilt angenommen. Der Mittelwert der Verteilung der klassenspezifischen Itemschwierigkeiten zu  $t=1$  beschreibt die mittlere Schwierigkeit eines Items über Klassen, während die Varianzkomponente die Streuung der Schwierigkeit eines Items über Klassen hinweg beschreibt (Mehrebenen-DIF; Robitzsch, 2009). Der Mittelwert der Verteilung der PPDs eines Items zeigt, ob die Veränderung dessen Schwierigkeit zwischen zwei Messzeitpunkten über Klassen im Mittel gerichtet oder ungerichtet verlief (globale Sensitivität). Die zugehörige Varianzkomponente gibt die Streuung der PPDs über Klassen an (differenzielle Sensitivität).

Die Verteilungen der Itemschwierigkeiten und PPDs können um Regressionsterme ergänzt werden. Im daraus resultierenden längsschnittlichen erklärenden IRT-Modell (vgl. van den Noortgate & De Boeck, 2005) können empirische Unterrichtsmerkmale als Prädiktoren für die Streuungen der Itemschwierigkeiten zu  $t=1$  sowie der PPDs eingebracht werden, um die Sensitivität zu erklären.

Exemplarisch wurde das Modell auf Daten des IGEL-Projekts (Hardy et al., 2011) angewendet. IGEL ist eine quasi-experimentelle Interventionsstudie zu individueller Förderung in der Grundschule. Teilnehmende Lehrpersonen erhielten eine Fortbildung in adaptiven Unterrichtsmethoden und setzten diese in einer standardisierten Unterrichtseinheit um. Der Datensatz umfasste 901 Kinder in 53 Klassen (Jahrgangsstufe drei).

Die Analysen fokussierten auf sieben Items eines Wissenstests, der unmittelbar vor und nach der Unterrichtseinheit durchgeführt wurde. Aufgaben wurden hinsichtlich des konzeptuellen Verständnisses der Schülerinnen und Schüler abgestuft kodiert. Drei Unterrichtsmaße wurden als Prädiktoren verwendet, die Umsetzung der Unterrichtsinhalte, die Zugehörigkeit zur formativen Assessment-Gruppe als Indikator für die Art der Vermittlung der Unterrichtsinhalte, sowie kognitive Aktivierung als ein etabliertes Maß für Unterrichtsqualität (Klieme, Pauli, & Reusser, 2009).

Alle Analysen wurden Bayesianisch mittels MCMC-Schätzung durchgeführt. Die Konvergenz war zufriedenstellend (Gelman-Rubins- $R=1.0$ ). Die Streuung der Itemschwierigkeiten zu  $t=1$  reichte von 0.00 bis 0.16 ( $M=0.06$ ,  $SD=0.07$ ). Die mittleren PPDs zeigten eine gerichtete Veränderung für alle Items an ( $M=-2.22$ ,  $SD=1.15$ ). PPD-Varianzen reichten von 0.22 bis 0.86 ( $M=0.52$ ,  $SD=0.25$ ). Drei PPD-Varianzen waren nicht signifikant. Zusammenfassend waren drei Items global sensitiv und vier Items waren global und differenziell sensitiv. Es zeigte sich kein Zusammenhang mit der Umsetzung. Die PPDs zweier Items hingen mit dem formativen Assessment zusammen. Zwei weitere Items zeigten einen Zusammenhang mit Unterrichtsqualität für deren Itemschwierigkeiten zu  $t=1$  und für deren PPDs. Vier Items waren demnach instruktionssensitiv hinsichtlich der Unterrichtsmethode und der Unterrichtsqualität. Zusammenfassend funktionierte das Modell gut in der Anwendung auf empirische Daten. Der Ansatz erlaubt es, die statistische Sensitivität von Items zwischen Messzeitpunkten und Klassen zu quantifizieren. Die Veränderung der Itemschwierigkeiten stand im Zusammenhang mit Unterrichtsmaßen. So konnte von einer Instruktionssensitivität der Aufgaben ausgegangen werden.

Creemers, B. P. M. & Kyriakides, L. (2008). The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools. London and New York: Routledge.

Hardy, I., Hertel, S., Kunter, M., Klieme, E., Warwas, J., Büttner, G., & Lühken, A. (2011). Adaptive Lerngelegenheiten in der Grundschule: Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrkompetenzen [Adaptive learning environments in primary school]. Zeitschrift für Pädagogik, 57 (6), 819–833.

Klieme, E., Pauli, C., & Reusser, K. (2009). The pythagoras study: Investigating effects of teaching and learning in swiss and german mathematics classrooms. In T. Janik & T. Seidel (Eds.), The power of video studies in investigating teaching and learning in the classroom (pp. 137–160). Münster, Waxmann.

Polikoff, M. S. (2010). Instructional sensitivity as a psychometric property of assessments. Educational Measurement: Issues and Practice, 29 (4), 3–14. doi: 10.1111/j.1745-3992.2010.00189.x

Robitzsch, A. (2009). Methodische Herausforderungen bei der Kalibrierung von Leistungstests [Methodical challenges in the calibration of performance tests]. In D. Granzer, O. Köller, & A. Bremerich-Vos (Eds.), Bildungsstandards Deutsch und Mathematik (pp. 42–106). Weinheim, Basel: Beltz.

van den Noortgate, W., & De Boeck, P. (2005). Assessing and Explaining Differential Item Functioning Using Logistic Mixed Models. Journal of Educational and Behavioral Statistics, 30 (4), 443–464.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
IG-Farben-Haus  
Raum 251





## **Schuleffekte bei der Erfassung analytischer Problemlösekompetenz:**

### **Eine Analyse der Messinvarianz im Mehrebenenfall**

Die Fähigkeit Problemstellungen zu lösen, die über den schulischen Kontext hinausgehen und in lebensnahen Situationen auftreten, ist eine der Kernkompetenzen, die Schüler/innen im Laufe ihrer Schulzeit erwerben sollen (OECD, 2004). Dabei kann die Anwendung von Problemlösestrategien in analytischen und komplexen Problemen differieren (Leutner et al., 2012). Im schulischen Kontext wird besonders auf den Erwerb analytischer Problemlösekompetenz fokussiert, der kognitive Prozesse des Verstehens, Repräsentierens, systematischen Kontrollierens und Evaluierens umfasst. Doch je nach curricularer Ausrichtung bzw. Orientierung des Unterrichts treten Unterschiede in den Ausprägungen der Problemlösefähigkeit zwischen Schulen auf, die bislang überwiegend auf Ebene von Personenfähigkeiten betrachtet und nicht auf Itemebene untersucht worden sind (OECD, 2004). Die aus einer Itemanalyse gewonnenen Erkenntnisse können jedoch genutzt werden, um detailliertere Informationen darüber zu erhalten, warum Items über Schulen hinweg unterschiedlich funktionieren können.

Ziel der Studie ist es somit, systematische Varianzquellen der analytischen Problemlösefähigkeit auf Item-, Schüler- und Schulebene zu analysieren und dabei mögliche Erklärungsansätze für Schulunterschiede zu liefern.

Basierend auf den Daten der deutschen PISA-2003-Stichprobe (N=2547) wurden 19 Items zur Erfassung analytischen Problemlösens auf differenzielles Itemfunktionieren zwischen insgesamt 215 Schulen untersucht. Die dichotom kodierten Antworten der Schüler/innen wurden dabei mit Hilfe von Mehrebenen-IRT-Modellen beschrieben und anhand von explanatorischen IRT-Modellen auf Schuleffekte geprüft (z.B. De Boeck & Wilson, 2004). Zudem konnten kognitive und nicht-kognitive Kovariaten auf Item-, Schüler- und Schulebene in die Modelle eingeführt werden. In den Analysen kamen unter anderem bayesianische Verfahren zur Anwendung, in denen die Schuleffekte für Itemschwierigkeiten und Itemdiskriminationen explizit modelliert wurden.

Die Ergebnisse zeigen Itemeffekte an, die unter anderem durch das assignierte PISA-Niveau der Items sowie die Art der Problemstellung aufgeklärt werden. Auf Schulebene wurden zudem Variablen wie das Lehrer- und Schülerverhalten, die unterrichtliche Ausstattung der Schulen sowie Variablen auf Seiten der Lehrer/innen identifiziert. Curriculare Unterschiede, die anhand der Schulautonomie gemessen wurden, zeigten dagegen keine bedeutsamen Effekte. Dennoch kann von Messäquivalenz zwischen Schüler- und Schulebene ausgegangen werden, die eine Voraussetzung zur Interpretation der Schulmittelwerte darstellt. In dieser Studie ist es erstmals gelungen, Schuleffekte der analytischen Problemlösekompetenz nicht nur auf Performanz- sondern auch auf Itemebene zu quantifizieren und substantiell aufzuklären. Die Ergebnisse deuten aber an, dass die Validität und Generalisierbarkeit des Konstrukts über Schulen hinweg in Frage zu stellen ist. Implikationen für das Assessment sowie die psychometrische Beschreibung von Schuleffekten werden im Vortrag diskutiert.

De Boeck, P., & Wilson, M. (Eds.) (2004). Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach. New York, NY: Springer.

Leutner, D., Fleischer, J., Greiff, S., Wirth, J., & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsstudien. Psychologische Rundschau, 63(1), 34-42.

OECD (2004). Problem solving for tomorrow's world. Paris: OECD.

Mittwoch,

05.03.2014,

09:00 - 10:45,

IG-Farben-Haus

Raum 251



## **Items zur Erfassung „funktionaler Lesekompetenz“ - Analyse von Itemparameter in Abhängigkeit von der Testdarbietungsform**

Im Rahmen eines von der Forschungsinitiative ASCOT (BMBF) geförderten Verbundprojekts (Messung allgemeiner Kompetenzen – adaptiv (Mak-adapt))\* wurden am Standort Darmstadt Leseitems im MC-Format entwickelt. Die Itementwicklung basiert auf einem Modell des „funktionalen Lesens“, das sich an Leseanforderungen im beruflichen Handlungskontext orientiert. Lesen im beruflichen Handlungskontext unterscheidet sich in einigen Merkmalen vom Lesen im Lernkontext. So wird primär mit dem Ziel gelesen Informationen zur Ausführung einer beruflichen Handlung zu gewinnen, und der Leseprozess ist in der Regel zeitlichen Restriktionen unterworfen. Zudem wird angenommen, dass sowohl Lesesituation als auch Texte andere mentale Modellierungsprozesse erfordern als beim Lesen mit primär epistemologischen Zielen (Ziegler, Balkenhol, Keimes, & Rexing, 2013). Die nach diesem Modell genierten Items wurden in zwei Studien getestet. In einer ersten Studie (Kalibrierungsstudie mit N = 1632) wurden die computerisierten Leseitems kombiniert mit computerisierten Items zur Erfassung der mathematischen Kompetenz und der naturwissenschaftlichen Kompetenz in einem Balanced Block Design dargeboten. In einer zweiten Studie (Pilotierungsstudie mit N = 1660) wurden dieselben Items einem Teil der Stichprobe (N = 528) in nach Domänen separierten Testheften in einer Paper & Pencil-Fassung vorgelegt. Die Testzeiten waren jeweils gleich, die Probanden der Teilstichprobe bearbeiteten jedoch nur ein Testheft, entweder zum Lesen, zur Mathematik oder zu Naturwissenschaften. In beiden Studien wurden Schülerinnen und Schüler in einer fortgeschrittenen Phase der Berufsausbildung in gewerblich-technischen oder kaufmännischen Berufen oder Berufen aus dem Berufsfeld Gesundheit und Pflege getestet.

Die Itemanalysen (Rasch-Modell) ergaben vor allem für die Domäne Lesen teilweise signifikante Abweichungen in den Schwierigkeitsparametern zwischen den am PC dargebotenen und im P&P-Testheft dargebotenen Items. Nachdem die Unterschiede vor allem in der Domäne Lesen auftreten, können sie nicht ausschließlich durch die Variation der Durchführungsmodalitäten verursacht sein, sondern basieren vermutlich auch auf Unterschiede in der Darbietungsform (PC versus P&P). Entsprechende Analysen sollen Aufschluss über die Einflussfaktoren bringen. Die Ergebnisse dieser Analyse sollen im Vortrag berichtet werden.

\*Es handelt sich um ein Verbundprojekt der Universitäten Göttingen (Susan Seiber), Jena (Andreas Frey) und Darmstadt (Birgit Ziegler).

## „( ) ja, ( ) nein, ( ) weiß nicht“? –

### Zur Reliabilität von Aufgaben mit unterschiedlichen Antwortformaten in Kompetenztests

Schulische Leistungen können sowohl mittels Aufgaben mit offenem Antwortformat (z.B. Antwort selbstständig notieren) als auch mit Aufgaben mit geschlossenem Antwortformat (z.B. „richtige“ Antwort ankreuzen bei Multiple-Choice-Aufgaben [MC]) erfasst werden. Bei Kompetenztests in großen Bildungsstudien werden häufig beide Antwortformate eingesetzt, wobei in der Regel alle Aufgaben auf einen varianzstarken (Haupt-)Faktor laden (z.B. Foy et al., 2012; Rauch & Hartig, 2010). Wenn MC und offene Antwortformate psychometrisch gleichwertig sind (vgl. Kastner & Stangl, 2011), bieten MC-Aufgaben beispielsweise eine höhere Auswertungsobjektivität und stellen die ökonomischere Wahl dar (Lukhele et al., 1994; Wainer & Thissen, 1993). Bei US-amerikanischen Advanced Placement (AP) Tests zeigte sich, dass zusätzliche Aufgaben mit offenem Antwortformat die Reliabilität von MC-basierten Skalen nicht wesentlich verbessern und nachfolgend keine inkrementelle Validität bzgl. des Studienerfolgs haben (Wainer & Thissen, 1993). Befunde aus einer kleinen türkischen TIMSS-Stichprobe deuten jedoch darauf hin, dass ein Verzicht auf Aufgaben mit offenem Antwortformat der Messgenauigkeit durchaus abträglich sein kann (Gültekin & Demirtasli, 2012).

Im Vortrag betrachten wir die Reliabilität der Aufgaben von IGLU 2006 (Lesen) und TIMSS 2007 (Mathematik) in Abhängigkeit vom Antwortformat. Die Parameterschätzungen basieren auf den Originaldatensätzen dieser beiden Studien. Als Reliabilitätsmaß der Item Response Theory wird die Fläche unter der Informationskurve berechnet – und zwar für (1) eine „Unterskala mit Aufgaben mit offenem Antwortformat“, (2) eine „Unterskala mit MC-Aufgaben“ und (3) eine Gesamtskala mit allen Items. Weiterhin wird exploriert, wie die Kriteriumsvalidität (Korrelation mit Schulnoten) in Abhängigkeit vom Aufgabenformat variiert.

Im mittleren Leistungsbereich zeigen beide Antwortformatspezifischen Skalen eine hohe Messgenauigkeit, während MC-Aufgaben im unteren Leistungsbereich nur sehr unpräzise Messungen liefern. Bei sehr hohen Kompetenzlevels liefern MC-Items zumindest bei den TIMSS-Aufgaben ebenfalls relativ ungenaue Leistungsschätzungen. Entgegen den US-amerikanischen AP-Befunden erhöhen Aufgaben mit offenem Antwortformat die Reliabilität von MC-Skalen, wobei erwartungsgemäß die Zahl der hinzugenommenen Items eine zentrale Rolle spielt.

Diese Ergebnisse sprechen für die Verwendung von MC-Aufgaben bei der Diagnostik im mittleren Fähigkeitsbereich (plus/minus eine Standardabweichung) – und zwar aufgrund der hohen Messgenauigkeit sowie der praktischen Vorteile bei Durchführung und Auswertung. Bei Messungen im niedrigen oder hohen Leistungsbereich empfiehlt sich dagegen die (zumindest ergänzende) Verwendung von Aufgaben mit offenem Antwortformat, zumal diese gelegentlich neben einem Hauptfaktor noch weitere Informationen erfassen (in Form eines methodenspezifischen zweiten, genesteten Faktors), welche vermutlich höhere Kompetenzen (mit-)abbilden (vgl. Rauch & Hartig, 2010). Insgesamt stehen die Ergebnisse im Einklang mit der aktuell in Bildungsstudien verfolgten Strategie, beide Formate zu verwenden, um eine möglichst präzise, valide und effiziente Kompetenzmessung zu erreichen.

Foy, P., Brossman, B., & Galia, J. (2012). Scaling the TIMSS and PIRLS 2011 Achievement Data. [http://timssandpirls.bc.edu/methods/pdf/TP11\\_Scaling\\_Achievement.pdf](http://timssandpirls.bc.edu/methods/pdf/TP11_Scaling_Achievement.pdf)

Gültekin, S., & Demirtaşlı, N. Ç. (2012). Comparing the Test Information Obtained through Multiple-Choice, Open-Ended and Mixed Item Tests Based on Item Response Theory. *Elementary Education Online*, 11(1), 251-263.

Kastner, M., & Stangl, B. (2011). Multiple Choice and Constructed Response Tests: Do Test Format and Scoring Matter? *Procedia Social and Behavioral Sciences*, 12, 263-273.

Lukhele, R., Thissen, D., & Wainer, H. (1994). On the Relative Value of Multiple-Choice, Constructed Response, and Examinee-Selected Items on Two Achievement Tests. *Journal of Educational Measurement*, 31(3), 234-250.

Rauch, D., & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items: A two-dimensional IRT analysis. *Psychological Test and Assessment Modeling*, 52(4), 354-379.

Wainer, H., & Thissen, D. (1993). Combining Multiple-Choice and Constructed-Response Test Scores: Toward a Marxist Theory of Test Construction. *Applied Measurement in Education*, 6(2), 103-118.

## **Chronotyp und Schule: Zusammenhänge zu Leistung und affektiv-motivationalen Variablen**

**Chair(s):** Franzis Preckel (Universität Trier)

Theoretischer Hintergrund: Auch wenn der Schulalltag bestimmte Vorgaben macht, wann Schüler/innen aufstehen müssen, gibt es dennoch eine sogenannte „innere Uhr“, die einen Einfluss darauf hat, wann wir bevorzugt schlafen oder aktiv sind. Bei diesem sogenannten Chronotyp werden vier Typen nach ihrer Ausprägung auf den unabhängigen Dimensionen der Morgenorientierung (MO) und Abendorientierung (AO) unterschieden. Morgentypen (MO+, AO-) fühlen sich morgens ausgeruht und leistungsstark und haben kein Problem damit, früh aufzustehen. Abendtypen (MO-, AO+) präferieren deutlich spätere Schlafens- und Aktivitätszeiten, gehen spät zu Bett und fühlen sich morgens unausgeruht. Lethargische Typen (MO-, AO-) gehen früh zu Bett und präferieren spätere Aufstehzeiten; sie fühlen sich in der Tagesmitte am leistungsstärksten. Energetische Typen (MO+, AO+) hingegen sind sowohl in den Morgen- wie den Abendstunden aktiv. Der Chronotyp ist mit Persönlichkeit, Intelligenz und akademischer Leistung korreliert und damit von großem Interesse für die Bildungsforschung (Preckel, Lipnevich, Schneider & Roberts, 2011).

Fragestellung: Es soll die Konzeptualisierung des Chronotypen in vier verschiedene Typen mit einem konfirmatorischen Ansatz (latente Klassenanalyse) überprüft werden. Weiterhin werden Unterschiede in Leistung und motivational-affektiven akademischen Variablen untersucht. Abschließend wird die inkrementelle Validität des Chronotypen für die Vorhersage von Leistungsmaßen nach Kontrolle der motivational-affektiven Variablen überprüft.

Methode: Untersucht wurde eine für Luxemburg repräsentative Teilstichprobe von insgesamt 1022 Neuntklässler/innen (50 % weiblich), die 2011 am Luxemburgischen Schulmonitoring (ÉpStan; EMACS, 2012) teilnahmen. Die Schüler/innen besuchten drei unterschiedliche Schulformen (12 % unterer, 59 % mittlerer, 29 % oberer Zweig).

Chronotyp: Mit der Kurzversion des Lark-Owl Chronotype Indicator (LOCI; Roberts, 1998) wurden die Morgen- und Abendorientierung der Schülerinnen mit jeweils 6 Items erfasst (z. B. „Mir fällt es leicht, morgens aufzustehen.“).

Die akademische Leistung wurde über vom Ministerium zur Verfügung gestellte Schulnoten in den Fächern Mathematik, Deutsch und Französisch sowie über im ÉpStan-Projekt entwickelte standardisierte Leistungstests in diesen Fächern erfasst. Ein weiteres Leistungsmaß war die Information, ob ein(e) Schüler/in bereits einmal sitzen geblieben war, welche vom Ministerium zur Verfügung gestellt wurde.

Als affektiv-motivationale Variablen erfassten wir das akademische Selbstkonzept (allgemein, Mathematik, Deutsch, Französisch) mit jeweils 3 Items aus dem Self-Description Questionnaire (z. B. Marsh & O'Neill, 1984). Die Skalen zur Erfassung des akademischen Interesses und der Prüfungsangst (jeweils allgemein, Mathematik, Deutsch, Französisch) wurden im ÉpStan Projekt entwickelt (jeweils 3 Items pro Skala). Die Zufriedenheit mit der Schule wurde mit drei Items aus PISA 2000 erfasst (alle Reliabilitäten > .80).

Ergebnisse: Die latente Klassenanalyse bestätigte die vier-Klassen-Chronotyp-Lösung. Bei Kontrolle von Schulform und Geschlecht hatten Schüler/innen mit einer hohen Abendorientierung signifikant schlechtere Noten als Lethargische oder Morgentypen. Chronotyp war zudem ein signifikanter Prädiktor für die Testleistung: Hier zeigten Abend- und Morgentypen die besten Werte. Im Hinblick auf die Versetzungsraten fand sich, dass Schüler/innen mit einer hohen Morgenorientierung häufiger versetzt wurden, während Schüler/innen mit einer hohen Abendorientierung häufiger sitzen blieben.

Für die akademischen Selbstkonzepte, das akademische Interesse und die Schulzufriedenheit zeigten sich Ergebnisse zu Gunsten der morgenorientierten bzw. zu Ungunsten der abendorientierten Chronotypen (keine Unterschiede für Prüfungsängstlichkeit).

Der Chronotyp erklärte inkrementell über Schultyp, Geschlecht und die affektiv-motivationalen Variablen hinaus 2 % der Notenvarianz (keine weitere Aufklärung der Testvarianz).

Diskussion und Implikationen: Abendorientierte Chronotypen zeigen ungünstigere Werte bezüglich Noten, Versetzungsraten und zentraler schulrelevanter affektiv-motivationaler Variablen. Keine entsprechenden Unterschiede ergeben sich für objektive Schulleistungstests. Dieser Befund korrespondiert damit, dass sich für Intelligenzmaße entweder keine oder Unterschiede zugunsten der Abendtypen finden (Preckel et al., 2011). Eine Abendorientierung kann daher als Risikofaktor für schulische Entwicklung interpretiert werden. Die Befunde werden zudem vor dem Hintergrund diskutiert, dass im Jugendalter generell eine Verschiebung des Chronotypen in Richtung Abendorientierung zu beobachten ist (z. B. Carskadon & Acebo, 2002).

Carskadon, M.A. & Acebo, C. (2002). Regulation of sleepiness in adolescence: Update, insights, and speculation. *Sleep*, 25, 606–616.

EMACS (2012). *Épreuves Standardisées. Nationaler Bericht 2011–2012*. Luxembourg: University of Luxembourg, EMACS.

Marsh, H. W. & O'Neill, R. (1984). Self Description Questionnaire Questionnaire III (SDQIII): The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, 21, 153–174.

Preckel, F., Lipnevich, A. A., Schneider, S. & Roberts, R. D. (2011). Chronotype, cognitive abilities, and academic achievement: A meta-analytic investigation. *Learning and Individual Differences*, 21, 483–492.

Roberts, R. D. (1998). *The lark-owl (chronotype) indicator (LOCI)*. Sydney, Australia: Entelligent Testing Products.

## ***Eine Vignettenstudie zum Einfluss von Persönlichkeit und situativen Anreizen – sowie deren Interaktion – auf prosoziales und kontraproduktives Schülerverhalten***

Prosoziales und kontraproduktives Verhalten von Schüler/inne/n ist für deren eigene schulische und soziale Entwicklung von großer Bedeutung (Wardle, Hunter & Warden, 2011). Außerdem stellen entsprechende Verhaltensweisen wichtige Qualitätsmerkmale einer Klassengemeinschaft dar (McEvoy & Welker, 2000). Entsprechend wichtig sind Kenntnisse über Antezedenzien oder Korrelate prosozialen und kontraproduktiven Verhaltens. Dabei stehen, in Anlehnung an allgemeine Erklärungsansätze zur Vorhersage von Verhalten, oftmals interindividuelle oder situative Faktoren im Fokus des Forschungsinteresses.

Hinsichtlich interindividueller Unterschiede scheint insbesondere die grundlegende Persönlichkeitseigenschaft Ehrlichkeit-Bescheidenheit aus dem HEXACO-Modell der Persönlichkeit (Ashton & Lee, 2007), ein Alternativmodell zum Fünf-Faktoren-Modell, bedeutsam zu sein. Ehrlichkeit-Bescheidenheit umfasst interindividuelle Unterschiede mit Bezug zu Aufrichtigkeit, Fairness oder Gier und konnte dementsprechend Kriterien wie Kooperationsverhalten (Hilbig, Zettler, Leist, & Heydasch, 2013) oder kontraproduktives Arbeitsverhalten (z.B. Zettler & Hilbig, 2010) vorhersagen. In den entsprechenden Studien besitzt Ehrlichkeit-Bescheidenheit in der Regel die stärkste Vorhersagekraft aller (grundlegenden) Persönlichkeitseigenschaften. Außerdem zeigte sich wiederholt eine Interaktion mit Situationsvariablen (Zettler, Hilbig, & Heydasch, 2013): Während Personen mit hohen Ausprägungen in Ehrlichkeit-Bescheidenheit generell viel prosoziales und wenig kontraproduktives Verhalten zeigen, neigen vor allem Personen mit niedrigen Ausprägungen in Ehrlichkeit-Bescheidenheit dazu, ihr Verhalten unterschiedlichen Bedingungen anzupassen. So zeigen sie vermehrt prosoziales Verhalten, wenn die Anreize dafür hoch sind.

Die bisherige Forschung zum HEXACO-Modell der Persönlichkeit im Allgemeinen und zu Ehrlichkeit-Bescheidenheit im Besonderen weist allerdings (wenigstens) eine größere Einschränkung auf: So gibt es bislang nahezu ausschließlich Befunde zu Erwachsenen, nicht aber zu Schulkindern. Dabei scheint Ehrlichkeit-Bescheidenheit auch für wichtige Variablen der Bildungsforschung wie eben prosoziales oder kontraproduktives Schülerverhalten bedeutsam zu sein. Genau dies wurde in der vorgestellten Studie untersucht. Konkret wurde getestet, ob Schüler/innen mit höheren Ausprägungen in Ehrlichkeit-Bescheidenheit bei der Vorstellung konkreter Schulsituationen eher angeben, sich prosozial bzw. nicht kontraproduktiv zu verhalten als solche mit niedrigen Ausprägungen. Zudem wurde untersucht, inwiefern die Anreizstruktur der beschriebenen Schulsituationen die Verhaltensangaben beeinflusst. Vermutet wurde dabei, dass sich Schüler/innen mit hohen Ausprägungen in Ehrlichkeit-Bescheidenheit nicht von der Anreizstruktur beeinflussen lassen, Schüler/innen mit niedrigen Ausprägungen dies jedoch tun.

Insgesamt nahmen N = 306 Schüler/innen (53 % weiblich) aus 17 Klassen (Klassenstufen 8-10) aus einer Realschule und einem Gymnasium an zwei Messzeitpunkten der Studie teil. Die Teilnehmer/innen bearbeiteten zunächst die aktuelle Version des HEXACO-Mittelstufenfragebogens (Gnisci et al.; in prep.). Zum zweiten Messzeitpunkt bearbeiteten sie verschiedene Vignetten zu einer typischen Schulalltagsituation. Dabei wurde für je ein Vignettenpaar manipuliert, ob die Situation einen positiven oder negativen Anreiz enthält, sich prosozial (bzw. in anderen Vignettenpaaren kontraproduktiv) zu verhalten. Ein positiver Anreiz bedeutete, dass dieser das prosoziale bzw. kontraproduktive Zielverhalten nahelegt. Hingegen legte ein negativer Anreiz nahe, dass Zielverhalten eher nicht zu zeigen. Beispielsweise handelte es sich in einer (prosozialen) Vignette um einen guten Freund, der Hilfe benötigt, und in der anderen Vignette des Vignettenpaares um einen Mitschüler, der einen zuvor geärgert hat. Abgesehen von dieser Manipulation wurden die Vignetten innerhalb eines Vignettenpaares identisch gehalten. Für jede Vignette sollten die Schüler/innen angeben, wie wahrscheinlich sie das beschriebene prosoziale (bzw. kontraproduktive) Zielverhalten in der jeweiligen Situation zeigen würden.

Die Ergebnisse lassen sich wie folgt zusammenfassen: Erwartungskonform gaben Personen mit hohen Ausprägungen in Ehrlichkeit-Bescheidenheit an, mehr prosoziales sowie weniger kontraproduktives Verhalten zu zeigen. Hinsichtlich des prognostizierten Interaktionseffekts zeigte sich ein differentielles Ergebnismuster: Für den Bereich des kontraproduktiven Verhaltens wurde die Hypothese bestätigt: Personen mit hohen Ausprägungen in Ehrlichkeits-Bescheidenheit besaßen relativ durchgängig eine niedrige Tendenz, kontraproduktives Verhalten zu zeigen. Personen mit niedrigen Ausprägungen hingegen zeigten vermehrt kontraproduktives Verhalten bei positiven Anreizen. Sie gaben aber bei Vignetten mit negativen Anreizen an, wenig kontraproduktives Verhalten zu zeigen. Für die Prosozialitäts-Vignetten ergab sich kein Interaktionsmuster.

Ashton, M.C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11, 150-166.

Gnisci, A., Perugini, M., Sergi, I., Conza, A., Ashton, M.C. & Lee, K. (in preparation). Development and validation of a HEXACO-personality questionnaire for middle school students (HEXACO-MSI).

McEvoy, A. & Welker, R. (2000). Antisocial behavior, academic failure, and school climate: A critical review. *Journal of Emotional and Behavioral Disorders*, 8, 130-140.

Wardle, G. A., Hunter, S. C. & Warden, D. (2011). Prosocial and antisocial children's perceptions of peers' motives for prosocial behaviours. *British Journal of Developmental Psychology*, 29, 396-408.

Zettler, I., Hilbig, B. E. & Heydasch, T. (2013). Two sides of one coin: Honesty-Humility and situational factors mutually shape social dilemma decision making. *Journal of Research in Personality*, 47, 286-295.

Zettler, I. & Hilbig, B. E. (2010). Honesty-Humility and a person-situation-interaction at work. *European Journal of Personality*, 24, 569-582.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G092





Nimisha Mahajan (Ruhr-Universität Bochum, Graduiertenkollegs Naturwissenschaftlicher Unterricht (NWU) Essen),  
Hubertina Thillmann (Ruhr-Universität Bochum),  
Joachim Wirth (Ruhr-Universität Bochum, Graduiertenkollegs Naturwissenschaftlicher Unterricht (NWU) Essen)

## ***The Impact of Goal Orientation on the Goal-Type Effect***

According to the “goal-type effect”, learning outcome was observed to be higher when learners were provided with learning goals which aim at modifying the “internal knowledge state” as compared to problem solving goals which aim at modifying the “external’ problem state” (Wirth, Künsting & Leutner, 2009). Though, the observed effects were small but significant ( $\eta^2=.015$ ). One reason could be that learners’ individual goal orientation affected their perception of the provided goals and learning outcome. To test this assumption, we assumed that there existed an aptitude-treatment interaction (ATI) among goal-type effect (learning vs. problem solving goal), mastery goal orientation (high vs. low) and performance goal orientation (high vs. low). Learners with the mastery goal orientation focus on the development of competencies while learners with performance goal orientation focus on the demonstration of competencies (Ames & Archer, 1987). The study followed a 2 x 2 x 2 experimental between-subjects design with mastery and performance goal orientations as quasi-experimental factor and goal type (learning vs. problem solving goal) as experimental factors. 429 students (from 8th and 9th grade of German schools) did self-regulated learning by conducting experiments within a computer-based learning environment.

The results could not replicate the “goal-type effect” which was observed in Wirth et al. (2009) study but supported the results from normative goal theory point of view (Ames, 1992; Dweck & Leggett, 1988) which suggests that a mastery goal orientation is most adaptive as compared to performance goal orientation. More precisely, we could show that high-mastery learners learned better as compared to low-mastery learners, whereas there was no difference in learning outcome with respect to performance goal oriented learners (high vs. low). Similar to other research findings (e.g., Barron & Harackiewicz, 2001; Darnon, Dompnier, Gillie’ron & Butera, 2010; Harackiewicz, Barron, Carter, Lehto & Elliot, 1997; Wolters, Yu & Pintrich, 1996), in this study also mastery and performance goal orientations were found to be positively correlated to each other, pointing to the possibility that both goal orientations were independently activated, and the learners could pursue different levels of these two goal orientations simultaneously (DeShon & Gillespie, 2005; Pintrich, 2000). We found that the highest learning outcome was displayed by the learners who were high on both mastery and performance goal orientations at the same time. The interaction effect between mastery goal orientation and goals provided was also analyzed, over and above their main effects. It was found that high-mastery goal oriented learners showed higher learning outcome, performed better on knowledge application test and also depicted effective use of strategies for learning as compared to low-mastery goal oriented learners when both were given learning goals. However, a three-way interaction of mastery goal orientation by performance goal orientation by goal-type was not found to be significant.

The most important generalization that a mastery goal orientation is the most adaptive is still valid. Therefore, emphasis should be on structuring classrooms in order to facilitate and foster mastery goal orientation among learners. But on a practical level, it is difficult to make schools free from performance goals (Urdu, 2004; Urdu & Turner, 2005); so perhaps in the meanwhile encouraging learners to pursue both types of goal orientations simultaneously could be the most adaptive pattern. In the present scenario, convincing learners to achieve performance via mastery and not via cheating, guessing or use of surface-level strategies should be emphasized. The present research also specifies the link between mastery goal orientation and learning goals. That is, if mastery goal oriented learners are provided with learning goals from the teachers, the overall net effect would be an “acquisition of new knowledge state” with a variety of positive outcomes.

- Ames, C. & Archer, J. (1987). Mothers’ beliefs about the role of ability and effort in school learning. *Journal of Educational Psychology*, 79, 409-414.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- Barron, K.E. & Harackiewicz, J.M. (2001). Achievement goals and optimal motivation; Testing multiple goal models. *Journal of Personality and Social Psychology*, 80, 706-722.
- Darnon, C., Dompnier, B., Gillie’ron, O. & Butera, F. (2010). The interplay of mastery and performance goals in social comparison: A multiple-goal perspective. *Journal of Educational Psychology*, 102, 212-222.
- DeShon, R.P. & Gillespie, J.Z. (2005). A motivated action theory account of global orientation. *Journal of Applied Psychology*, 90(6), 1096-1127.
- Dweck, C.S. & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Harackiewicz, J.M., Barron, K.E., Carter, S.M., Lehto, A.T. & Elliot, A.J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, 73, 1284-1295.
- Pintrich, P.R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544-555.
- Urdu, T. (2004). Can achievement goal theory guide school reform? In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement: Vol. 13. Motivating students, improving schools: The legacy of Carol Midgley* (pp. 361-392). Stamford, CT: JAI Press.
- Urdu, T. & Turner, J.C. (2005). Competence motivation in the classroom. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 297-317). New York, NY: Guilford Press.
- Wirth, J., Künsting, J. & Leutner, D. (2009). The impact of goal specificity and goal type on learning outcome and cognitive load. *Computers in Human Behavior*, 25, 299-305.
- Wolters, C.A., Yu, S.L. & Pintrich, P.R. (1996). The relation between goal orientation and students’ motivational beliefs and self-regulated learning. *Learning and Individual Differences*, 8, 211-238.

## Negativer Zusammenhang von Self-Efficacy und Lernerfolg

Das Konzept „Selbstwirksamkeitsüberzeugungen“ (Self-Efficacy, SE) bezeichnet Erwartungen, bestimmte Handlungen aufgrund eigener Fähigkeiten erfolgreich ausführen zu können (Bandura, 1977; Zimmerman, 2000). Ein positiver Zusammenhang von akademischen SE-Überzeugungen und akademischem Erfolg konnte in Meta-Analysen belegt werden (Multon, Brown, & Lent, 1991; Robbins, Lauver, Le, Davis, Langley, & Carlstrom, 2004; Richardson, Abraham, & Bond, 2012). Allerdings liegen auch gegenteilige Befunde vor (z.B. Vancouver & Kendall, 2006), wonach höhere SE-Überzeugungen zu verminderter Anstrengung und schlechterer Leistung führen.

Fragestellung: Ziel dieser Studie ist zu untersuchen, ob bereichsspezifische (auf die akademische Informationskompetenz bezogene) SE einen Einfluss auf den Lernerfolg in einem Informationskompetenztraining für Psychologiestudierende besitzt. Primäre Inhalte des Blended-Learning-Trainings waren die Nutzung von Fachdatenbanken und von Google Scholar zum Finden psychologischer Fachliteratur sowie die Beschaffung und Bewertung von Fachliteratur. Im Speziellen soll geprüft werden, (1) ob bereichsspezifische SE mit dem Niveau der Informationskompetenz vor Beginn des Trainings zusammenhängt und, (2) ob bereichsspezifische SE sich auf den Lernerfolg im Training auswirkt.

Methode: An dem Training nahmen N = 67 Bachelor-Psychologiestudierende (davon n = 52 [78%] weiblich, mittleres Alter: M = 21.67 Jahre) teil. Zur Erfassung der Informationskompetenz wurde ein Wissenstest mit 35 Multiple Choice-Items verwendet (Leichner, Peter, Mayer, & Krampen, 2013), der neben weiteren Instrumenten vor und nach dem Training (t1 bzw. t2) zu bearbeiten war. Die SE bei der Informationsrecherche und -bewertung wurde zu t1 mit einem aus k = 10 Items bestehenden Fragebogen erhoben.

Ergebnisse: Die Daten wurden mithilfe von Mehrebenenmodellen ausgewertet, da solche Modelle ermöglichen, die Veränderung der Kriteriumsvariablen auf Personenebene zu untersuchen (Langer, 2009). Kriteriumsvariable war die Leistung im Wissenstest zu den beiden Messzeitpunkten. Erste Analysen zeigten, dass ein Modell mit zwischen den Probanden variierender Steigung (interpretierbar als Hinweis auf interindividuelle Unterschiede im Lernerfolg) und Achsenabschnitt die Daten signifikant besser beschreiben konnte ( $X^2 [1] = 9.96, p < .05$ ) als ein Modell ohne diese Annahmen.

Der durchschnittliche Punktwert bei t1 (Intercept) betrug  $b = 0.45$  ( $t[66] = 29.66, p < .001$ ). Die Testwerte der Probanden stiegen im Verlauf des Trainings signifikant an ( $b = 0.14, t[66] = 17.29, p < .001$ ). SE-Überzeugungen zu t1 wurde als zentrierter Prädiktor auf Personenebene hinzugefügt, wodurch sich die Modellanpassung verbesserte ( $X^2 [1] = 14.88, p < .001$ ). Die Parameter für Intercept und Slope blieben nahezu unverändert. Es zeigte sich ein positiver Einfluss von SE auf die Leistung im Informationskompetenztest zu t1 ( $b = 0.007, t[65] = 3.81, p < .001$ ). Allerdings wirkte sich SE negativ auf den Lernerfolg aus ( $b = -0.003, t[65] = 2.95, p < 0.01$ ).

Diskussion: Probanden mit höheren Selbstwirksamkeitsüberzeugungen erreichten höhere Werte in Wissenstest, was dafür spricht, dass Selbsteinschätzungen von Studierenden durchaus valide sein können, was anderen Befunden (z.B. Kennedy, Lawton, & Plumlee, 2002) widerspricht. Allerdings wirkte sich SE negativ auf den Lerngewinn aus. Da der Informationskompetenztest keine Deckeneffekte aufwies, kann der Effekt nicht ausschließlich als methodisches Artefakt interpretiert werden. Stattdessen wird vermutet, dass die Effekte vermittelt über geringere Anstrengung der Probanden mit höherer SE zu Stande kamen. Ein solcher Zusammenhang kann mithilfe der Kontrolltheorie erklärt werden (vgl. Vancouver & Kendall, 2006). Implikationen für die Gestaltung von Trainingsprogrammen für Studierende werden diskutiert.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191–215.

Kennedy, E. J., Lawton, L., & Plumlee, E. L. (2002). Blissful Ignorance: The Problem of Unrecognized Incompetence and Academic Performance. *Journal of Marketing Education*, 24(3), 243–252. doi:10.1177/0273475302238047

Langer, W. (2009). Mehrebenenanalyse: Eine Einführung für Forschung und Praxis (2. Auflage). Studienskripten zur Soziologie. Wiesbaden: VS Verlag für Sozialwissenschaften.

Leichner, N., Peter, J., Mayer, A.-K., & Krampen, G. (2013). Assessing information literacy among German psychology students. RSR, accepted for publication.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30–38. doi:10.1037/0022-0167.38.1.30

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387. doi:10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. *Psychological Bulletin*, 130(2), 261–288. doi:10.1037/0033-2909.130.2.261

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology*, 91(5), 1146–1153. doi:10.1037/0021-9010.91.5.1146

Zimmerman, B. J. (2000). Self-Efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82–91. doi:10.1006/ceps.1999.1016



Silke Schneider (GESIS – Leibniz-Institut für Sozialwissenschaften),  
Natascha Massing (GESIS – Leibniz-Institut für Sozialwissenschaften)

## ***Degrees of Competency: The Relationship Between Educational Qualifications and Adult Competencies Across Countries***

Most surveys collect information on education by asking about respondents' highest educational qualification achieved. Only few specialized studies also collect data on individual competencies and skills, and often this is the case only in studies with specific (usually student) samples. As this information is missing, many studies focusing on „human capital“ theory assume that educational qualifications achieved validly indicate individuals' competencies and skills, at least in an ordinal sense of “more means more”: Higher levels of qualification require longer education times and passing more demanding examinations, both of which should be positively correlated with competencies. Furthermore, persons with higher competencies are more likely to remain in the educational system for a longer time due to the lower costs associated with learning. However, there are two comparability problems involved: Firstly, quality of education and effectiveness of learning is known to differ across countries, at the same educational level (see results from the PISA studies, e.g. OECD, 2010). Secondly, educational systems are not purely hierarchical but also characterised by horizontal differentiation, e.g. between general and vocational, or university preparatory and labor-market preparatory education. It can be assumed that the speed and level of skill production differs across levels as well as types of education within educational levels, and this may differ across countries. So how do the adult competencies measured for individuals with equivalent educational qualifications compare across countries, controlling for factors related to post-educational skill acquisition (e.g. work experience)? Is it valid to interpret levels of qualifications as proxies for competencies and skills? Furthermore, is it valid to compare educational qualifications across countries as if they indicated equivalent competencies and skills?

The newly available PIAAC data allows us for the first time to examine these questions empirically across 24 countries and adult cohorts. Educational level is measured using the International Standard Classification of Education (ISCED) 1997 and competencies are measured using psychometric tests. This presentation compares the skills and competencies of adults who achieved a specific level of educational attainment across countries and highlights conceptual and empirical differences between educational attainment (qualifications) and educational achievement (competencies and skills). We expect that 1) countries differ substantially with respect to the absolute levels of competency achieved by adults from comparable cohorts with equivalent educational qualifications but 2) that the order of educational qualifications in a country remains a significant predictor for skill levels in each country. In order to compare skill production by educational qualifications across countries, we perform both bivariate analyses as well as multivariate analyses using multilevel modeling. Due to possible changes in educational systems over time, especially different timing of educational expansion across countries, we also assess our research questions in different age groups.

OECD (2010), PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I).

## Mathematische Eingangsvoraussetzungen von Studierenden der Ingenieurwissenschaften

### Chair(s): Irene Neumann (IPN Kiel)

Der Mathematik kommt in den Ingenieurwissenschaften eine zentrale Rolle zu. So müssen beispielsweise Drehmomente berechnet werden, um erfolgreich Motoren konzipieren zu können, oder Kräfteverhältnisse bestimmt werden, um tragfähige Brücken bauen zu können. Normativ betrachtet sollten Studienanfänger über einen Grundstock an mathematischen Fähigkeiten und Fertigkeiten verfügen, die in den Lehrplänen der Länder für das (Fach-)Abitur oder den Bildungsstandards für die allgemeine Hochschulreife (KMK, 2012) beschrieben sind. Um jedoch studiengangspezifische mathematische Lehrveranstaltungen angepasst an das Vorwissen der Studierenden konzipieren und anbieten zu können, ist es unabdingbar zu wissen, inwieweit diese normativ beschriebenen Kompetenzen bei Studierenden auch tatsächlich vorhanden sind. Dieser Frage wird im vorliegenden Beitrag nachgegangen.

Mathematische Kompetenzen werden meist durch eine inhaltsbezogene und denkprozessbezogene Komponente beschrieben. Beispielsweise unterscheiden die Bildungsstandards inhaltliche „Leitideen“ und kognitive „allgemeine mathematische Kompetenzen“ auf drei hierarchisch gestuften Anforderungsniveaus (KMK, 2012). Ähnliche Dimensionierungen finden sich bei den internationalen Vergleichsstudien PISA (OECD, 2013) oder TIMSS (Baumert et al., 2000). Eine Konzeptualisierung mathematischer Kompetenzen von Gymnasiasten am Ende der Schullaufbahn durch eine inhalts- und eine denkprozessbezogene Komponente findet sich auch bei Hauck (2012). Zur Erstellung eines Itempools zur Erfassung dieser Kompetenzen unterschied Hauck neben anderen Aufgabenmerkmalen vier mathematische Themenbereiche – Algebra und Grundlagen, Analysis, Geometrie und Stochastik – sowie deklaratives und prozedurales Wissen. Die vier Themenbereiche ergaben sich bei Hauck aus einer curricularen Analyse von Standards und Lehrplänen und gelten als typische Inhaltsfelder der Mathematikurse in der Oberstufe in Deutschland. Insgesamt wurden auf dieser Basis 64 Aufgaben entwickelt.

Um die mathematischen Eingangsvoraussetzungen von Studierenden der Ingenieurwissenschaften zu untersuchen, wurden aus dem Aufgabenpool von Hauck (2012) – mit geringfügigen Überarbeitungen – 36 Items ausgewählt. Aufgaben zur Stochastik wurden dabei nicht berücksichtigt, da eine Analyse der Inhalte von Skripten und Modulhandbüchern der entsprechenden Veranstaltungen aus dem Ingenieursstudium keine bzw. nur geringfügige Thematisierung der Stochastik ergab (vgl. auch Ștefăniță, 2011). Aus Pilotierungsdaten des Itempools (Hauck, 2012) lag eine Raschanalyse vor, sodass bei der Aufgabenauswahl auf eine gleichmäßige Verteilung von Aufgabenschwierigkeiten, angemessene Passung zum Raschmodell sowie angemessene Trennschärfen geachtet werden konnte. Darüber hinaus wurden die Aufgaben so ausgewählt, dass die drei verbleibenden Inhaltsbereiche gleichmäßig berücksichtigt waren. Es wurden zwei Testhefte erstellt, die sich lediglich in der Reihenfolge der Aufgaben unterschieden, um Positionseffekte auszugleichen. Die Aufgaben wurden in einer Stichprobe von 1069 Studierenden eingesetzt (792 Universität, 237 Fachhochschule, 40 keine Angabe). Darunter waren 754 Studierende des Maschinenbaus, 119 des Bauingenieurwesens, und 86 der Fahrzeugtechnik (Rest sonstige oder keine Angabe).

Bei der Auswertung der Daten wurde ein Item aufgrund schlechter psychometrischer Eigenschaften (niedrige Trennschärfe) ausgeschlossen. Ein ein- und ein dreidimensionales Raschmodell wurden an die Daten angepasst, wobei das dreidimensionale Modell (pro Inhaltsbereich eine eigene Fähigkeit) einen signifikant besseren Fit zeigte ( $\Delta$ deviance=210,  $df=5$ ,  $p<.001$ ; BIC1-dim=34420,33, BIC3-dim=34245,04). Bei einer auf Null festgelegten mittleren Personenfähigkeit zeichnete sich der Trend ab, dass die Itemschwierigkeiten von Algebra und Grundlagen (AG;  $M=-0.85$ ,  $SD=1.54$ ) über Analysis (An;  $M=0.16$ ,  $SD=1.67$ ) hin zur Geometrie (Ge;  $M=1.76$ ,  $SD=1.26$ ) ansteigen. Die EAP/PV-Reliabilität war dabei in allen drei Dimensionen zufriedenstellend (AG: .67, An: .76, Ge: .70). Ein multivariater Vergleich der Personenparameter ergab in allen drei Inhaltsbereichen ein besseres Abschneiden der Universitäts- gegenüber den Fachhochschulstudierenden ( $F(3,1023)=6.93$ ,  $p<.001$ ). Gleichermäßen zeigte sich ein Effekt bzgl. der Studienfächer ( $F(6,1906)=3.05$ ,  $p<.01$ ), wobei nachfolgende univariate Analysen Effekte für Algebra und Grundlagen ( $F(2,954)=5.40$ ,  $p<.01$ ) sowie Geometrie ( $F(2,954)=5.68$ ,  $p<.01$ ) aufwies, jedoch nicht für Analysis ( $F(2,954)=2.46$ ,  $p=.09$ ). Unter Kontrolle des Hochschultyps lassen sich jedoch keine studiengangbezogenen Unterschiede feststellen.

Insgesamt stellte sich das eingesetzte Instrument als geeignet heraus, die Eingangsvoraussetzungen von Ingenieursstudierenden zu erfassen, und hinsichtlich der untersuchten Inhaltsbereiche Vergleiche von Studierendenfähigkeiten zulässt. Um weitere Aussagen über diese Fähigkeiten zu erhalten sind jedoch Detailanalysen nötig, wie sie beispielsweise in der Think-Aloud-Studie von Lehmann und Rösken-Winter durchgeführt wurden.

Baumert, J., Bos, W., Brockmann, J., Gruehn, S., Klieme, E., Köller, O., Lehmann, R., Lehrke, M., Neubrand, J., Schnabel, K. U., Schwippert, K., & Watermann, R. (2000). TIMSS/III-Deutschland: Der Abschlussbericht. [http://www.timss.mpg.de/TIMSS\\_im\\_Ueberblick/TIMSSIII-Broschuere.pdf](http://www.timss.mpg.de/TIMSS_im_Ueberblick/TIMSSIII-Broschuere.pdf)

Hauck, C. (2012). Vergleich der mathematischen Kompetenzen von Gymnasiasten an allgemein bildenden und technischen Gymnasien in Baden-Württemberg. Unveröffentlichte Masterarbeit, Universität Stuttgart.

Ștefăniță, F. (2011). Qualitative Analyse der thematischen Dimensionen und Anforderungen im Fach (Höhere / Angewandte) Mathematik im Rahmen des Maschinenbaustudiums an ausgewählten Hochschulen Baden-Württembergs. Unveröffentlichte Masterarbeit, Universität Stuttgart.

KMK (2012). Bildungsstandards für die Allgemeine Hochschulreife im Fach Mathematik. [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2012/2012\\_10\\_18-Bildungsstandards-Mathe-Abi.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf)

OECD (2013). PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy. OECD Publishing.

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G102



## **Mathematische und physikalische Problemlösekompetenzen von Studierenden der Ingenieurwissenschaften: Kategorisierung von Interdependenzen und Barrieren**

Die Item Response Theory (IRT) ermöglicht es, das Itemantwortverhalten von Personen verschiedener Leistungsstufen zu modellieren. Die Item Response Curve beschreibt dabei die nicht-lineare Beziehung zwischen der Wahrscheinlichkeit eines manifesten Antwortverhaltens in Abhängigkeit von der Ausprägung des zugrundeliegenden latenten Merkmals (Trait) (vgl. Embretson & Reise, 2000). Im Rahmen des Projektes KoM@ING aus der BMBF-Förderlinie KoKoHs wurden IRT-basierte Eingangstests zur Mathematik und Technischen Mechanik für Studienanfänger im Studiengang Maschinenbau entwickelt, die an umfangreichen Stichproben pilotiert sind (vgl. Beitrag von Neumann et al.). So können einzelne Kompetenzen reliabel und valide gemessen werden, jedoch bleiben die Prozesse, die Lernende beim Lösen von mathemathhaltigen Aufgaben in Technischer Mechanik durchlaufen, noch unkategorisiert. Über zusätzliche qualitative Analysen konnte jedoch ein Kategoriensystem entwickelt werden, anhand dessen Barrieren der Studierenden beim Problemlöseprozess beschrieben und erfasst werden können.

Als theoretischer Rahmen für die Entwicklung des Kategoriensystems wird zunächst auf die Arbeiten Polyas (1957) zurückgegriffen, welcher den Ablauf des Problemlöseprozesses fokussiert und diesen in vier Teilschritte unterteilt: Verstehen des Problems, Entwickeln eines Plans, Ausführen des Plans und Rückschau. Bruder und Collet (2011) dagegen rücken die Methoden in den Mittelpunkt und erklären den Problemlöseprozess mit Hilfe der heuristischen Hilfsmittel (Informative Figuren, Tabellen u.a.) und Strategien (z.B. Analogieschluss, Systematisches Probieren), die Lernende einsetzen, wenn sie eine Aufgabe lösen. Die Verwendung von Mathematik beim Lösen physikalischer Aufgaben wird durch Einbeziehung des Modells der Epistemic Games von Tuminaro und Redish (2007) berücksichtigt. Epistemic Games beschreiben die Art und Weise, wie Studierende ihr Vorwissen und ihre Ideen organisieren und Physikaufgaben entsprechend lösen, beispielsweise mittels Mapping Meaning to Mathematics. Diese drei Ansätze beschreiben denselben Problemlöseprozess aus unterschiedlichen Blickwinkeln und ermöglichen so, nicht nur Interdependenzen der jeweiligen Fachkulturen einzufangen, sondern videografierte Daten auch differenziert zu kategorisieren.

Strategien zur Aufgabenbearbeitung sind bisher nur implizit Thema in der Mathematik- und Physikausbildung für Ingenieurstudierende. Sie können jedoch der Schlüssel zu hoher Problemlösekompetenz sein. Für die Kompetenzmodellierung und -messung wird daher untersucht, über welche Basisstrategien Studierende verfügen, wie sie diese in Abhängigkeit vom Schwierigkeitsgrad der Aufgaben einsetzen und welche Parameter für eine erfolgreiche Bearbeitung entscheidend sind. Diese Fragen werden in einer qualitativen Längsschnitt-Videostudie unter Einsatz des Kategoriensystems beantwortet, mit welchem auch Entwicklungen über den Verlauf des Studiums abgebildet werden können. Die Forschungsfragen lauten: Wie können mit Hilfe eines auf Basis der o.g. Theorien fundierten Kategoriensystems Problemlösestrategien hinsichtlich ihrer Popularität, ihrer Universalität und ihres Potenzials eingestuft werden? Welche Schwierigkeitsparameter können durch das Kategoriensystem beschrieben werden? Welche Interdependenzen zwischen mathematischen und physikalischen Problemlösekompetenzen treten auf?

An der Videostudie nahmen insgesamt 37 Ingenieursstudierende aus dem ersten Semester teil. Die Personenfähigkeitsschätzer für Mathematik und Physik wurden mittels der oben erwähnten IRT-basierten Tests ermittelt. Aus diesen Testheften bearbeiteten die Probanden dann in Gruppen die fünf leichtesten und die fünf anspruchsvollsten Aufgaben. Eine solche Einteilung der Aufgabenschwierigkeit ist auf Grund der IRT-Basierung der Tests möglich. Unterschiede und Barrieren bei der Bearbeitung verschieden anspruchsvoller Aufgaben können somit erfasst werden. Die Proband(inn)en lösen die Aufgaben in Kleingruppen und verbalisieren ihre Ideen und Gedanken (Think-Aloud-Methode). Die Pilotierung des Kategoriensystems erfolgte an drei Videos, deren Transkriptionen in einem Interraterverfahren mit vier Ratern ausgewertet wurden. Mit Hilfe des Kategoriensystems kann gezeigt werden, dass die eingesetzten Problemlösestrategien geeignet klassifiziert werden können, um insbesondere fachspezifische Interdependenzen und Barrieren zu beschreiben. Neben den Personenfähigkeitsschätzern wurden auch Einstellungen zur Schulmathematik und -physik und Schulabschlussnoten als Prädiktoren für entsprechende Zusammenhangshypothesen erhoben. Im Vortrag wird aufgezeigt, inwiefern Einsatz, Häufigkeit und Intensität von Problemlösestrategien variiert und neben der objektiven Aufgabenschwierigkeit als subjektive Komponente die Performanz beeinflusst. Die Ergebnisse liefern Aufschluss darüber, wie spezifische Problemlösestrategien in die Kompetenzmodellierung mathematisch-physikalischer Kompetenzen einfließen müssen.

Bruder, R., & Collet, C. (2011). Problemlösen lernen im Mathematikunterricht. Berlin: Cornelsen Verlag Scriptor.

Embretson, S.E., & Reise, S. (2000). Item response theory for psychologists. Mahwah: Erlbaum Publisher.

Polya, G. (1957). How To Solve It – A New Aspect of Mathematical Method. Garden City: Doubleday

Tuminaro, J., Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. Physics Education Research, 3 (2), 020101-1 – 020101-22. doi: 10.1103/PhysRevSTPER.3.020101

Mittwoch,  
05.03.2014,  
09:00 - 10:45,  
PEG-Gebäude  
Raum 1.G102



## **Zyklische Verfeinerung von Interventionsmaßnahmen nach dem Prinzip des Design Research am Beispiel von Mathematik für Ingenieurstudierende**

Als Framework zur Verknüpfung von Theorie und Praxis bietet Design Research eine gleichberechtigte Wechselbeziehung der beiden Bereiche in Zyklen aus Entwicklung und Evaluation hinsichtlich der ausgewählten Interventionen und der initiierten Prozesse (vgl. Gravemeijer & Cobb, 2006). Dabei verfolgt der Verfeinerungsprozess das Doppelziel, die Theorie zu justieren und die Lehrpraxis zu optimieren (Burkhardt & Schoenfeld, 2003) - eine ideale Basis für langfristig angelegte Unterstützungsprojekte, die beforscht werden.

Das vom Stifterverband geförderte Projekt MP<sup>2</sup>-Mathe/Plus der Ruhr-Universität Bochum, in dem Studierende der Ingenieurwissenschaften in der Studieneingangsphase Unterstützung in Mathematik finden, ist eines von sechs ausgewählten Vorhaben des Kollegs Lehre<sup>an</sup> 2013. Die jährlichen Projektphasen im Wintersemester ermöglichen die Modifikation der Maßnahmen auf der Grundlage der empirischen Evaluationen. MP<sup>2</sup>-Mathe/Plus befindet sich aktuell im vierten Projektzyklus. Anfänglich basierte es auf der Hypothese, dass es Studierenden zu Beginn ihres Studiums vorrangig an Lernstrategien mangelt. Dies führte zu der Forschungsfrage, inwiefern die Instruktion von Lernstrategien in Kombination mit einer Förderung der Metakognition (Flavell, 1979) und der Selbstregulation (Bandura, 1977) den Lernerfolg begünstigt, was als Projektinterventionen diverse Hilfestellungen (u.a. wöchentliche Treffen, e-Learning-Kurs, ausführliches tägliches Lerntagebuch) implizierte.

Die Evaluation der Interventionen in den verschiedenen Projektzyklen stützt sich auf mehrere empirische Erhebungen, u.a. Prä- und Postumfragen zu Lernstrategien mit Hilfe einer Adaption des LIST-Fragebogens (Lernstrategien Im Studium, vgl. Wild & Schiefele, 1994). Hierfür wurden zu Beginn und am Ende des ersten Semesters die verwendeten Lernstrategien aller Studierenden der betroffenen Lehrveranstaltung abgefragt (2010/11: n(prä)=259, n(post)=145; 2011/12: n(prä)=542, n(post)=102; 2012/13: n(prä)=459, n(post)=270) und auf verschiedenen Skalen (z.B. Elaboration, Metakognition, Anstrengung oder Aufmerksamkeit) abgebildet. Eine anonyme Kodierung ermöglichte den Vergleich zwischen ProjektteilnehmerInnen (2010/11: n(TN)=75, 2011/12: n(TN)=82, 2012/13: n(TN)=84) und Nicht-TeilnehmerInnen. Zudem wurden Leistungstests nach diversen Gesichtspunkten (z.B. Regelmäßigkeit der Projektteilnahme, Leistungsstand zu Studienbeginn, Gender) ausgewertet. Ergänzend wurden Interviews durchgeführt, um die sozialen und affektiven Dimensionen qualitativ analysieren zu können.

Die Erkenntnisse des ersten Zyklus führten zu der modifizierten Hypothese, dass Motivation (Heckhausen, 1989) und Selbstwirksamkeit sowie eine stärkere Betonung affektiver Variablen (McLeod, 1992), insbesondere von Emotionen (Goldin, 2002; Lazarus, 1991) und ihrer Interaktion mit Kognitionen (Op 't Eynde, de Corte & Verschaffel, 2006; Vygotsky, 1978) für die Umsetzung der erlernten Lernstrategien bedeutsam sind. Dies führte zur Stärkung von Peer-Unterstützung und mehr gemeinschaftlicher Verantwortlichkeit der Studierenden (auch durch Einbeziehung sozialer Netzwerke) sowie zu einer engeren Verzahnung der Interventionen. Hingegen wurde das aktive Eingreifen der ProjektutorInnen in Lernprozesse in drei Phasen schrittweise reduziert, um die Autonomie der Studierenden nach Ablauf von MP<sup>2</sup>-Mathe/Plus zu unterstützen. Dieses Prozedere bewährte sich und wurde daher im nachfolgenden dritten Projektzyklus nur wenig modifiziert, lediglich die Zielsetzung des Lerntagebuchs wurde verschoben, weg von Dokumentation des Lernverhaltens hin zu unmittelbarem Feedback und motivierender digitaler Benutzeroberfläche im LearningLog online. Dass personelle Umstellungen die Qualität der Projektmaßnahmen nicht beeinflussten, unterstützt die beschriebenen Entscheidungen.

Die Auswertungen lieferten auch das Ergebnis, dass männliche Studierende weniger von den Interventionen profitierten als weibliche. Diese ausgeprägten Gender-Unterschiede können durch Einbeziehung spezifischer Attributionstheorien (u.a. Weiner, 1994) fundiert werden, jedoch ist deren Einflusspotenzial gering. Im vierten Projektjahr wird daher neben dem Ausbau des LearningLog online und der Vernetzung über soziale Netzwerke eine Abschwächung der geschlechtsspezifischen Ausprägungen durch positive Rollenmodelle angestrebt. Die Ausweitung auf weitere Ingenieursstudiengänge an der Ruhr-Universität Bochum und an anderen Hochschulen gestattet zudem eine Überprüfung des Maßnahmenpakets auf Übertragbarkeit.

Alles in allem hat sich nach Ablauf von drei und Planung von vier Projektzyklen ein tragfähiges Gesamtkonzept entwickelt, das neben einem Bündel von verzahnten Praxisinterventionen, die einander ergänzen und verstärken, auch eine Bereicherung der Theoriebasis umfasst.

Bandura, A. (1977). Social learning theory. Englewood Cliffs, N.J.: Prentice Hall.

Burkhardt, H., & Schoenfeld, A. (2003). Improving Educational Research: Toward a More Useful, More Influential, and Better-Funded Enterprise. *Educational Researcher*, 32(9), 3–14.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.

Goldin, G. A. (2002). Affect, meta-affect, and mathematical belief structures. In G. C. Leder, E. Pehkonen, & G. Törner (Hrsg.), *Mathematics education library: Vol. 31. Beliefs. A hidden variable in mathematics education?* (S. 59–72). Dordrecht, Boston: Kluwer Academic Publishers.

Gravemeijer, K., & Cobb, P. (2006). Design Research from a Learning Design Perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Hrsg.), *Educational design research* (S. 45–85). London, New York: Routledge.

Heckhausen, H. (1989). *Motivation und Handeln* (2. Auflage). Berlin, Heidelberg, New York: Springer.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Hrsg.), *Handbook of research on mathematics teaching and learning* (S. 575–596). New York: Macmillan.

Op 't Eynde, P., Corte, E. D., & Verschaffel, L. (2006). Accepting Emotional Complexity: A Socio-Constructivist Perspective on the Role of Emotions in the Mathematics Classroom. *Educational Studies in Mathematics*, 63(2), 193–207.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Weiner, B. (1994). *Motivationspsychologie* (3. Auflage). Weinheim: Beltz, Psychologie-Verlag-Union.

Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185–200.



## **Die Bedeutung repräsentationaler Flexibilität für die erfolgreiche Bearbeitung mathematischer Problemaufgaben**

Flexibles Denken und Handeln wird in vielen Lebensbereichen als vorteilhaft erachtet. Dies trifft auch für den Bereich mathematischer Bildung zu. Betrachtet man beispielsweise das Lösungsvorgehen von Schülerinnen und Schülern bei der Bearbeitung mathematischer (Problem)aufgaben, so wird deutlich, dass die flexible Nutzung verschiedener Repräsentationsformen oft diskutiert und gefordert wird (z. B. Heinze, Star & Verschaffel, 2009), eine empirische Überprüfung des Nutzes von repräsentationaler Flexibilität jedoch weiterhin aussteht. Die vorliegende Forschungsarbeit soll diesbezüglich etwas Licht ins Dunkel bringen.

Repräsentationale Flexibilität wird als Teilkomponente allgemeiner kognitiver Flexibilität aufgefasst, die folgendermaßen definiert werden kann: „the selection and/or modification of available problem solving techniques, methods, or strategies as a function of changes of the task or situation“ (Krems, 1995, p. 202). Die flexible Nutzung verschiedener Repräsentationen stellt einen grundlegenden Aspekt kognitiver Flexibilität dar, weil Repräsentationen Formate oder Strukturen darstellen, auf denen Strategien operieren (vgl. Tabachneck, Koedinger & Nathan, 1994; Lovett & Schunn, 1999). Dabei können unterschiedliche Formen von Repräsentationen unterschieden werden (vgl. Cox, 1999; Schnotz & Bannert, 2003): ikonische (z. B. Skizzen, Diagramme) und symbolische (z. B. Gleichungen, Formeln) Repräsentationen sowie externe (z. B. aufgeschriebene Rechnungen) und interne (z. B. Kopfrechnung) Repräsentationen.

Zwei Forschungsfragen sollen nachfolgend beantwortet werden: 1) Wie flexibel setzen Schülerinnen und Schüler verschiedener Klassenstufen Repräsentationen bei der Lösung mathematischer Problemaufgaben ein? und 2) In welchem Zusammenhang steht diese Flexibilität mit dem Lösungserfolg? Da mehrschrittige Aufgaben eingesetzt wurden, können zwei Formen repräsentationaler Flexibilität unterschieden werden: die flexible Nutzung von Repräsentationen innerhalb einer Aufgabe und die flexible Nutzung zwischen den Aufgaben.

Um die repräsentationale Flexibilität von Schülerinnen und Schülern bei der Bearbeitung mathematischer Problemaufgaben und ihren Zusammenhang mit dem Lösungserfolg zu untersuchen, wurden 268 Schülerinnen und Schüler aus Grundschulen (3., 4. Klassen) und Gymnasien (6., 9. Klassen) gebeten, vier problemhaltige Textaufgaben selbstständig zu bearbeiten. Das Lösungsvorgehen wurde videographiert und anhand eines Kodiersystems quantifiziert. Basierend auf den gewonnenen Daten wurden zwei Kennwerte gebildet, um repräsentationale Flexibilität innerhalb und zwischen Aufgaben betrachten zu können. Beide Kennwerte haben einen Range von null (keine Flexibilität) bis eins (maximale Flexibilität).

Die Datenauswertung zeigt, dass innerhalb der Aufgaben die flexible Nutzung der Repräsentationen mit der Klassenstufe zunimmt (3. Klasse:  $M = 0.24$ ; 4. Klasse:  $M = 0.33$ ; 6. Klasse:  $M = 0.40$ ; 9. Klasse:  $M = 0.48$ ). Zwischen den Aufgaben zeigt sich nicht ein so klares Bild. Hier nimmt die repräsentationale Flexibilität bis zur 6. Klasse zu (3. Klasse:  $M = 0.16$ ; 4. Klasse:  $M = 0.26$ ; 6. Klasse:  $M = 0.33$ ) und in der neunten Klasse wieder leicht ab ( $M = 0.25$ ).

Hinsichtlich des Zusammenhangs repräsentationaler Flexibilität mit dem Lösungserfolg wurde die Klassenstufe als konfundierende Variable berücksichtigt, da mit höherer Klassenstufe mehr Aufgaben richtig gelöst wurden. Partialkorrelationen zeigen, dass lediglich die flexible Nutzung von Repräsentationen zwischen den Aufgaben signifikant mit dem Lösungserfolg assoziiert ist ( $pr = .17$ ,  $p = .02$ ), die flexible Nutzung innerhalb der Aufgaben jedoch nicht ( $pr = .01$ ,  $p = .90$ ). Diese Ergebnisse entsprechen Befunden aus der mathematischen Strategieforschung (Elia, van den Heuvel-Panhuizen & Kolovou, 2009), wo gezeigt wurde, dass sich lediglich der flexible Strategiewechsel zwischen verschiedenen Aufgaben positiv auf den Lösungserfolg auswirkt.

Die flexible Nutzung von Repräsentationen scheint für den Lösungserfolg bei der Bearbeitung mathematischer Problemaufgaben bedeutsam zu sein. Hier ist jedoch lediglich der flexible Einsatz verschiedener Repräsentationsformen zwischen den Aufgaben relevant, die flexible Nutzung innerhalb von Aufgaben nicht. Es bedarf weiterer Untersuchungen um diese Befundlage abzusichern. Darüber hinaus werfen die Ergebnisse Fragen für zukünftige Forschung auf: Wie wird der Einfluss repräsentationaler Flexibilität auf den Lösungserfolg durch Aufgaben- oder Personencharakteristika beeinflusst? Wie kann man die flexible Nutzung von Repräsentationen im Schulkontext vermitteln? Ist repräsentationale Flexibilität immer nützlich bzw. unter welchen Bedingungen ist sie kontraproduktiv?

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9, 343-363.

Elia, I., Van den Heuvel-Panhuizen, M. & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM – The International Journal on Mathematics Education*, 41, 605-618.

Heinze, A., Star, J. R. & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM – The International Journal on Mathematics Education*, 41, 535-540.

Krems, J. (1995). Cognitive flexibility and complex problem solving. In P. Frensch & J. Funke (Hrsg.), *Complex Problem Solving: The European Perspective* (S. 201-218). Hillsdale: Erlbaum.

Lovett, M. C. & Schunn, C. D. (1999). Task representations, strategy variability and base-rate neglect. *Journal of Experimental Psychology: General*, 128(2), 107-130.

Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141-156.

Tabachneck, H. J. M., Koedinger, K. R. & Nathan, M. J. (1994). Towards a theoretical account of strategy use and sense making in mathematical problem solving. In A. Ram, & K. Eiselt, *Proceedings of the 16th annual conference of the cognitive science society* (S. 836-841). Hillsdale, NJ: Erlbaum.



Lars Jenßen (Humboldt-Universität zu Berlin), Simone Dunekacke (Humboldt-Universität zu Berlin),  
Sigrid Blömeke (Humboldt-Universität zu Berlin)

## **Das Zusammenspiel von mathematischer Kompetenz und Mathe-Angst bei angehenden Erzieherinnen - eine Latent-State-Trait-Analyse**

Die Präsentation ist Teil des Projekts „KomMa“, welches die professionelle Kompetenz angehender Erzieherinnen und Erzieher im Bereich Mathematik untersucht (Dunekacke et al., 2013).

Erziehungswissenschaftliche Kompetenzkonzepte gehen davon aus, dass Kompetenz ein Trait ist. Weinert (2001) definiert Kompetenz z.B. als die Fähigkeit, Probleme in variablen Situationen zu bewältigen. Diese Vorstellung entspricht dabei der Definition eines Traits (Liebert & Liebert, 1998). Nach Klieme, Hartig und Rauch (2008, S. 5) hat die Gleichsetzung von „Kompetenz“ und „Trait“ den Vorteil, größere Gruppen hinsichtlich ihrer interindividuellen Leistungsunterschiede zu untersuchen, da die Annahme getroffen werden kann, dass derartige Unterschiede nur auf Personenmerkmale zurückzuführen sind. Evidenzen deuten darauf hin, dass mathematische Kompetenz als stabil angesehen werden kann, wobei davon ausgegangen werden kann, dass sie sich über die Lebensspanne hinweg entwickelt (Aunola et al., 2004).

Dieser als „Dispositionismus“ vertretene Ansatz steht konträr zu dem sog. „Situationismus“-Ansatz in der Psychologie, der davon ausgeht, dass Messungen immer Zustände von Personen widerspiegeln, die auch situative Einflussfaktoren mit einbeziehen (z.B. Mischel, 1968). Dieser Kontroverse innerhalb der Psychologie begegnet die Latent State-Trait Theorie (LST), indem sie es ermöglicht, situative Einflüsse auf das Zustandekommen einer Leistung von dispositionellen Faktoren zu trennen (Steyer, Schmitt & Eid, 1999). Ausgehend von der klassischen Testtheorie besteht die LST formal in einer Dekomposition der Zustandsvariablen („latent state variable“) in einen unsystematischen Fehlerwert, in einen „trait“-Teil (Disposition) und in ein latentes State-Residuum (Charakteristika der Situation und Interaktion von Person und Situation, „occasion specifics“). Die LST wurde mittlerweile in zahlreichen Bereichen der Psychologie angewandt (Geiser & Lockhart, 2012), jedoch noch nicht im bildungswissenschaftlichen Kontext.

Im Bereich mathematischer Kompetenz wird Mathe-Angst als Prädiktor für schlechtere Leistungen in Mathematik diskutiert (Evans, 2000; Ma, 1999). Theoretische Vorstellungen und empirische Evidenzen deuten darauf hin, dass Angst sowohl aus Trait-Anteilen besteht, als auch auf situationspezifische Einflüsse zurückgeht. Studien, die Mathe-Angst untersuchten, konnten diese Annahme auch speziell für dieses Konstrukt belegen (Goetz et al., 2013).

Ziel dieser Studie ist zu untersuchen, ob auch bei mathematischer Kompetenz bei angehenden Erzieher/-innen davon ausgegangen werden kann, dass situationspezifische Einflüsse im Sinne der LST einen bedeutsamen Einfluss haben. Darüber hinaus soll das Zusammenspiel von mathematischer Kompetenz und Mathe-Angst bei angehenden Erzieher/-innen auf latenter State-Trait-Ebene untersucht werden.

Um die Forschungsfrage zu beantworten, werden 350 angehende Erzieherinnen und Erzieher innerhalb von drei Wochen zweimal mit einem im Projekt KomMa entwickelten Mathe-Test getestet. Die Mathe-Angst soll mithilfe der Mathematics Anxiety Scale (Bai et al., 2009) erfragt werden. Im Anschluss erfolgt sukzessiv die LST-Modellierung: In einem ersten Schritt wird ein State-Trait Modell der mathematischen Kompetenz geschätzt, um Trait-Varianz und situationspezifische Varianz zu trennen. Im Anschluss wird ein State-Trait Modell der Mathe-Angst geschätzt. Zuletzt wird das komplexe Zusammenspiel mathematischer Kompetenz mit Mathe-Angst in einem Multistate-Multitrait-Multiconstruct-Modell (Eid & Diener, 2004) untersucht.

Anhand theoretischer Überlegungen zum Kompetenzkonstrukt und Studien zur mathematischen Kompetenz wird davon ausgegangen, dass die Trait-Varianz mathematischer Kompetenz höher ist als die Varianz der occasion-specifics, dass also Personcharakteristika bedeutsamer sind als situationspezifische Einflüsse. Für die Mathe-Angst wird, ausgehend von Theorien und aktuellen Studien, angenommen, dass sowohl Personcharakteristika als auch situationspezifische Faktoren gleichermaßen bedeutsam sind. Außerdem wird davon ausgegangen, dass Mathe-Angst einen negativen Einfluss auf die Trait-Varianz der mathematischen Kompetenz hat, aber ebenso einen negativen Einfluss auf situationspezifischen Anteile in dem Sinne, dass Mathe-Angst diese Anteile mathematischer Kompetenz bedeutsam ausformt.

In der Diskussion soll auf den Nutzen der Methode für zukünftige Untersuchungen eingegangen werden (im Sinne der Konstruktvalidierung und Testkonstruktion). Die LST-Modellierung soll zu einem besseren Verständnis mathematischer Kompetenz beitragen. Ob Mathe-Angst eher einen negativen Einfluss auf den Trait-Anteil oder den situationspezifischen Anteil hat, ist bisher nicht erforscht. Mögliche praktische Implikationen für die Ausbildung der Erzieher/-innen sollen vor dem Hintergrund der Ergebnisse diskutiert werden.

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology*, 96(4), 699-713.

Bai, H., Wang, L.S., Pan, W., & Frey, M. (2009). Measuring Mathematics Anxiety: Psychometric Analysis of a Bidimensional Affective Scale. *Journal of Instructional Psychology*, 36(3), 185-193.

Dunekacke, S., Jenßen, L., Baack, W., Schmude, C., Wedekind, H., Grassmann, M., & Blömeke, S. (2013). KomMa – Structure, Level and Development of Professional Competencies of Pre-School (Kindergarten) Teachers in the field of Mathematics. In: S. Blömeke & O. Zlatkin-Troitschanskaia (Eds.) (2013). The German funding initiative “Modeling and Measuring Competencies in Higher Education”: 23 research projects on engineering, economics and social sciences, education and generic skills of higher education students. (KoKoHs Working Papers, 3). Berlin & Mainz: Humboldt University & Johannes Gutenberg University.

Epstein, S. (1984). The stability of behavior across time and situations. In R. Zucker, J. Aronoff, & A.I. Rabin (Eds), *Personality and the Prediction of Behavior* (pp. 209-268). San Diego: Academic Press.

Eid, M., & Diener, E. (2004). Global judgments of subjective well-being: Situational variability and long-term stability. *Social Indicators Research*, 65, 245-277.

Geiser, C., & Lockhart, G. (2012). A Comparison of Four Approaches to Account for Method Effects in Latent State-Trait Analyses. *Psychological Methods*, 17(2), 255-283.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R. & Hall, N.C. (2013). Do Girls Really Experience More Anxiety in Mathematics? *Psychological Science*, 1-9.

Klieme, E., Hartig, J., & Rauch, D. (2008). The Concept of Competence in Educational Contexts. In J. Hartig, E. Klieme, & D. Leutner, *Assessment of Competencies in Educational Contexts* (p. 3-22). Göttingen: Hogrefe.

Ma, X. (1999). A Meta-Analysis of the Relationship between Anxiety toward Mathematics and Achievement in Mathematics. *Journal for Research in Mathematics Education*, 30(5), 520-540.

Marteau, T. M. and Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31, 301-306.

Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent State-Trait Theory and Research in Personality and Individual Differences. *European Journal of Personality*, 13, 389-408.

Weinert, F.E. (2001). Concept of Competence: A Conceptual Classification. In D.S. Rychen, & L. Hersh Salganik (Eds.). *Defining and Selecting Key Competencies*. Göttingen: Hogrefe.

Ziegler, M., Ehrlenspiel, F., & Brand, R. (2009). Latent state-trait theory: An application in sport psychology. *Psychology of Sport and Exercise*, 10(3), 344-349.



Chair(s): Tobias Rausch (Otto-Friedrich-Universität Bamberg)

Diskutant/in: Tina Hascher (Universität Bern)

Einschätzungen von Schülermerkmalen durch Lehrkräfte sind im Schulalltag allgegenwärtig. Dies bezieht sich jedoch nicht nur auf Lernergebnisse, Lernvoraussetzungen oder emotional-motivationale Merkmale, sondern auch auf die Beurteilung von Lernprozessen sowie auf die prognostische Einschätzung von Schülerinnen und Schülern im Hinblick auf die weitere Gestaltung des Unterrichts oder auf Schullaufbahneempfehlungen.

Dieser Vielfalt beim professionellen Urteilen von Lehrkräften wird zunehmend auch in der Forschung Rechnung getragen. Dadurch hat sich die Forschung zur professionellen Urteilskompetenz von Lehrkräften stark ausdifferenziert. Die zunächst eher deskriptiven Befunde zur Güte der Einschätzung von Schülermerkmalen wurden durch die Betrachtung potentieller Hintergrundvariablen ergänzt. Ebenso wurde der Beurteilungsprozess mehr in den Fokus gerückt. Eine solche Ausdifferenzierung, ebenso wie die Betrachtung professioneller Urteile aus unterschiedlichen disziplinären Blickwinkeln, hatte auch den Einsatz vielfältiger methodischer Herangehensweisen zur Folge. Ziel des Symposiums ist es, durch die Betonung der Methoden- und Perspektivenvielfalt in den Beiträgen zu einem breiter gefassten Verständnis der professionellen Urteilskompetenz von Lehrkräften näher zu kommen.

Im ersten Beitrag wird der Einfluss von Wissen über schwierigkeitsgenerierende Merkmale von Texten und Aufgaben, sowie von Wissen über adäquate Lesestrategien auf die in einer Felduntersuchung erhobene Urteilsgüte der Lehrkräfte betrachtet. Es wurde vermutet, dass es Lehrkräften mit größerem Wissen leichter fällt, korrekte aufgabenspezifische Urteile über die Schüler ihrer Klasse zu fällen. Dies konnte in der Untersuchung nicht grundlegend bestätigt werden.

Im Gegensatz zu den in der Literatur vorherrschenden Felduntersuchungen bietet der experimentelle Ansatz, wie er im Simulierten Klassenraum realisiert wird, den Vorteil, die interessierenden Variablen isoliert betrachten zu können. Im zweiten Beitrag wird eine Untersuchung in einem solchen experimentellen Setting beschrieben. Es finden sich hier keine direkten Effekte des manipulierten sozioökonomischen Status der simulierten Schüler auf die Leistungseinschätzung durch (angehende) Lehrkräfte.

Weniger mit der Beurteilung von Lernergebnissen sondern mit der Diagnose von Lernprozessen, Wissenslücken und naiven Konzepten beschäftigt sich der dritte Beitrag. Hier wird anhand von Videovignetten untersucht, inwiefern Lehramtsstudierende und erfahrene Lehrkräfte Strategien zur Lernprozessdiagnose im Fach Biologie identifizieren können. Die gefundenen Unterschiede zwischen den beiden Gruppen in der Bewertung von hinderlichen und förderlichen Strategien werden mit Interviewdaten weiter untermauert.

Ebenfalls bezogen auf den Urteilsprozess wird im vierten Beitrag untersucht, ob die Art und Weise wie Schülerinformationen verarbeitet und gespeichert werden einen Einfluss auf das Lehrerurteil hat. Es wird angenommen, dass die Information, die zuerst wahrgenommen wird auch am besten erinnert wird und somit das Urteil über einen Schüler maßgeblich mitbestimmt. Mithilfe eines Eyetracking-Verfahrens konnten in der Studie Blickbewegungen bei der Informationsaufnahme durch Schülervignetten nachvollzogen werden.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
Casino-Gebäude  
Raum 823

Tobias Rausch (Otto-Friedrich-Universität Bamberg), Jacqueline Matthäi (Otto-Friedrich-Universität Bamberg),  
Cordula Artelt (Otto-Friedrich-Universität Bamberg)

### ***Mit Wissen zu akkurateren Urteilen?***

#### ***Zum Zusammenhang von Wissensgrundlagen und Urteilsgüte im Bereich des Textverstehens***

Die Einschätzung und Beurteilung von Schülerinnen und Schülern spielt im Schulalltag eine wichtige Rolle und ist eine der zentralen Aufgaben von Lehrkräften. Jedoch zeigen bisherige Studien, dass Lehrereinschätzungen häufig nicht die Schülerleistung akkurat vorhersagen können, und dass es große interindividuelle Unterschiede in der Urteilsgüte gibt. Studien belegen, dass die Expertise von Lehrkräften mit der Genauigkeit von Schülerbeurteilungen positiv zusammenhängt (z.B. Herppich et al., 2013). Dabei wurde die Expertise von Lehrkräften bisher meist über distale Merkmale (z.B. über die Berufserfahrung, Abschlussnoten im Studium oder die Anzahl besuchter Universitätskurse) operationalisiert. Das domänenspezifische Wissen als zentrale Bedingung von Expertise (Gruber & Mandl, 1996) hingegen wird im Hinblick auf die Prädiktionskraft für die Urteilsgüte bisher selten betrachtet. Es wird davon ausgegangen, dass ein detailliertes fachliches, fachdidaktisches und diagnostisches Wissen die Urteilsgüte der Einschätzung von Schülerleistungen vorhersagen kann (Artelt & Rausch, in Druck). Eine empirische Untersuchung dieser Vermutung steht für den Bereich des Textverstehens jedoch bisher noch aus. Im Rahmen des BiKS-Projekts (Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter) wurde der Frage zum Zusammenhang von fachlichen und fachdidaktischen Wissensgrundlagen einerseits und der Urteilsgüte im Bereich Textverstehen andererseits nachgegangen.

Basierend auf psychologischen, linguistischen und didaktischen Theorien bezüglich Text- und Aufgabenschwierigkeit, sowie zu kognitiven Prozessen beim Lösen von Aufgaben zum Textverstehen wurde ein Test entwickelt, der darauf abzielt, die diesbezüglichen Wissensgrundlagen von Deutschlehrkräften der Sekundarstufe I zu quantifizieren. Im Mittelpunkt steht dabei die Erfassung des Wissens über schwierigkeitsgenerierende Merkmale von Texten und Aufgaben, sowie über adäquate Strategien im Umgang mit Texten. Auf der Textebene sollen fünf kurze Texte hinsichtlich ihrer Schwierigkeit in eine Rangreihe gebracht werden, anspruchsvolle Textpassagen umgeschrieben werden und schwierige Textstellen markiert werden. Bei diesen Aufgaben gaben die Lehrkräfte im Mittel 57% korrekte Antworten. Auf der Aufgabenebene wurden die von den Lehrkräften vorgenommenen Schwierigkeitseinschätzungen von sieben Fragen zu einem Text mit der in der BiKS-Studie gefundenen empirischen Aufgabenschwierigkeit der gleichen Fragen verglichen. Dabei verschätzten die Lehrkräfte sich im Mittel um 16,5 Prozentpunkte. Beim Vergleich von Aufgabenpaaren hinsichtlich ihrer Schwierigkeit wurden im Mittel nur 38% der Aufgaben korrekt eingeschätzt. Das Wissen über den angemessenen Einsatz von Lesestrategien wurde in Anlehnung an den WLST 7-12 über paarweise Vergleiche von Strategien zu vier verschiedenen Szenarien erhoben. Die Lehrkräfte konnten hier je nach Szenario zwischen 50 und 87% der Strategien korrekt einschätzen. Um nun den Zusammenhang von Wissensgrundlagen und der Urteilsgüte untersuchen zu können, wurde der Test auch bei N = 26 Lehrkräften administriert, die im Rahmen einer Feldstudie zur Urteilsgenauigkeit im Bereich des Textverstehens teilnahmen. Die Urteilsgüte von globalen und aufgabenspezifischen Einschätzungen bewegt sich im Rahmen der Ergebnisse vorheriger Untersuchungen (z.B. Karing et al., 2011). Während bei globalen Urteilen die Einschätzung der Schülerfähigkeit mit dem allgemeinen Konstrukt Textverstehen in Bezug gesetzt werden muss, besteht die Anforderung bei aufgabenspezifischen Urteilen darin, die Einschätzung der Fähigkeit des Schülers mit der Einschätzung der Schwierigkeit der entsprechenden Aufgabe in Beziehung zu setzen (Karing et al., 2011).

Aufgrund der größeren Nähe der aufgabenspezifischen Urteile zum Wissenstest wird vermutet, dass größeres Wissen über schwierigkeitsgenerierende Merkmale von Texten und Aufgaben, sowie Wissen über adäquate Lesestrategien mit akkurateren aufgabenspezifischen Urteilen über Schüler einhergeht. Erste Analysen deuten jedoch an, dass es keinen Zusammenhang zwischen dem mit dem beschriebenen Test erfassten Wissen und der Urteilsgüte im Textverstehen gibt. Dies gilt sowohl für aufgabenspezifische als auch für globale Urteile zum Textverstehen. Derzeit werden unter Einbeziehung zusätzlicher Variablen weitere detailliertere Analysen durchgeführt, die ein differenzierteres Bild der Zusammenhänge zwischen den Wissensgrundlagen und der Güte diagnostischer Urteile erlauben.



## ***Die Auswirkung des sozioökonomischen Status von SchülerInnen auf die Urteilsgenauigkeit von Lehrkräften im Simulierten Klassenraum***

Die Urteilsgenauigkeit von Lehrkräften spielt in Lehr- und Lernkontexten eine wesentliche Rolle. Neben der Auswahl von geeigneten Unterrichtsmaterialien stellt vor allem die korrekte Beurteilung der Leistungen von SchülerInnen einen gewichtigen Punkt dar. Diese Beurteilungen sind Grundlage für das Erreichen von Schulabschlüssen sowie für die Versetzung am Jahresende und es ist erstrebenswert, Bewertungen z.B. in Form von Noten ausschließlich nach dem Leistungsprinzip zu vergeben und eventuell auftretende Urteilsverzerrungen zu minimieren. Die Forschung zeigt jedoch, dass die Notenvergabe die Gütekriterien von Messungen nicht immer ausreichend erfüllt. Insbesondere konnte dokumentiert werden, dass leistungsfremde Merkmale wie die Handschrift aber auch der sozioökonomische Status (SES) die Benotung beeinflussen können (Ingenkamp & Lissmann, 2008). So wurde in zahlreichen Studien gezeigt, dass Kinder mit niedrigem SES selbst bei Kontrolle kognitiver Grundfähigkeiten sowie der Lesekompetenz der SchülerInnen eine geringere Chance haben, eine Empfehlung für das Gymnasium zu erhalten, als Kinder mit hohem SES (z.B. Bos et al., 2004; Ditton, Krüsken, & Schauenberg, 2005). Jedoch bietet die Forschung bislang wenige Studien, die sich mithilfe eines experimentellen Settings der Frage nach dem Zusammenhang zwischen Lehrerurteil und SES der SchülerInnen angenommen haben. Dieser Beitrag möchte diese Lücke verkleinern und mithilfe des Simulierten Klassenraumes (SKR) untersuchen, ob (angehende) Lehrkräfte SchülerInnen mit geringem SES schlechter bewerten als SchülerInnen mit mittlerem oder hohem SES. Vermutet wurde, dass mit höherem SES eine höhere Leistungsbeurteilung einhergeht. Beim SKR handelt es sich um die Computersimulation einer Unterrichtssituation, in der die Versuchspersonen die Rolle einer Lehrkraft übernehmen, mit virtuellen SchülerInnen interagieren, Zusatzinformationen über die SchülerInnen erhalten und deren experimentell gesteuerte Leistung beurteilen sollen (Südkamp & Möller, 2009). Die Stichprobe bestand aus  $n = 29$  Grundschullehrkräften (90% weiblich; Alter:  $M = 42.9$ ,  $SD = 12.1$  Jahre; Berufserfahrung:  $M = 14.7$ ,  $SD = 12.8$  Jahre) und  $n = 30$  Lehramtsstudierenden (57% weiblich; Alter:  $M = 24.4$ ,  $SD = 1.9$  Jahre; Semester:  $M = 8.4$ ,  $SD = 2.7$ ). Unabhängige Variable war der SES der simulierten SchülerInnen, dreistufig operationalisiert über die Berufsangaben der Eltern (niedriger, mittlerer und hoher SES: ISEI entsprechend zwischen 16 und 40, 41 und 64, 65 und 90), die mit weiteren (unwichtigen) Informationen über die SchülerInnen dargeboten wurden. Als weitere UV wurde die Leistungsfähigkeit der Schüler variiert (schwach, durchschnittlich, stark), operationalisiert über die Wahrscheinlichkeit, eine richtige Antwort auf im SKR gestellte Fragen zu geben. Die abhängige Variable stellte die Leistungsbeurteilung in Prozent dar („Wie viel Prozent der an ihn/sie gestellten Aufgaben hat jede/r Schüler/in richtig beantwortet?“). Als Manipulationscheck wurden die Versuchspersonen gebeten, im Anschluss an die Leistungsbeurteilung die Berufsangaben der Eltern zu den entsprechenden SchülerInnen zuzuordnen. Die Auswertung erfolgte unter Berücksichtigung der hierarchischen Datenstruktur mittels Regressionsanalyse in MPlus getrennt für die Lehrer- und Studentenstichprobe. Es wurde geprüft, ob der SES der SchülerInnen unter Berücksichtigung der tatsächlich erbrachten Leistung der SchülerInnen einen Einfluss auf die Leistungsbeurteilung der (angehenden) Lehrkräfte hatte. Entgegen unserer Erwartungen ergab sich für beide Stichproben kein statistisch signifikanter Effekt des SES der SchülerInnen auf die Leistungsbeurteilung. Allerdings zeigte sich in der Studierendenstichprobe ein Effekt, wenn anstelle des manipulierten SES der SchülerInnen der aus der Erinnerung zugeordnete (und als Manipulationscheck intendierte) SES in der Regressionsanalyse berücksichtigt wurde. SchülerInnen, denen aus der Erinnerung Elternberufe zugeordnet wurden, die einen niedrigen SES bedeuten, wurden unter Kontrolle der tatsächlichen Leistung signifikant schlechter beurteilt als SchülerInnen mit hohem SES. Die Studie wirft weitere Fragen auf. Diskussionsbedürftig ist sicherlich auch die Manipulation des SES. Wird die Anfälligkeit von Urteilsverzerrungen mit Eintritt in das Berufsleben geringer? Über welche Variablen wird der im Feld gefundene Einfluss des SES vermittelt und lässt er sich experimentell abbilden? Die vorliegenden Ergebnisse können als Ausgangspunkt für weitere Studien genutzt werden.

## **Lernprozessdiagnose im Dialog – Erkennen und Bewerten diagnostisch relevanter Interaktionsstrategien im Video**

Lernprozesse von Schülerinnen und Schülern (SuS) akkurat zu diagnostizieren, ist eine wichtige Anforderung an Lehrkräfte. Lehrkräfte sollten zum Beispiel einschätzen können, welche Wissenslücken oder naive Konzepte bei SuS bestehen. Studien zeigen jedoch, dass es angehenden und erfahrenen Lehrkräften sogar im Einzelgespräch mit SuS schwer fällt, solche Lernprozesse zu diagnostizieren. Diese Schwierigkeit hängt möglicherweise mit den Zielen und korrespondierenden Interaktionsstrategien der Lehrkräfte zusammen. So fördern bestimmte Interaktionsstrategien das Erreichen eines Zieles, während andere Strategien die Erreichung eher behindern. Bezogen auf die Diagnose von Lernprozessen sind jene Interaktionsstrategien förderlich, die SuS anregen, aktiv ihr Wissen zu äußern (z.B. offene Fragen oder Nachhaken). Strategien hingegen, bei denen vor allem die Lehrkraft aktiv ist (z.B. Erklären oder geschlossene Fragen stellen), erlauben es den SuS weniger, ihr Wissen zu explizieren. Der Lehrkraft wird es folglich erschwert, das Wissen der SuS akkurat einzuschätzen. Passend hierzu konnte gezeigt werden, dass Lehrkräfte einen interaktiveren Lehrstil als Lehramtsstudierende verwendeten, der mit einer akkurateren Beurteilung der SuS einherging. Es stellt sich die Frage, welche impliziten Vorstellungen angehende und berufserfahrene Lehrkräfte bezüglich der Eignung von Strategien zur Lernprozessdiagnose haben und wie die adäquate Verwendung dieser Strategien gefördert werden kann. Vor diesem Hintergrund haben wir untersucht, inwiefern sich Lehrkräfte und Lehramtsstudierende im Erkennen und Interpretieren diagnostisch relevanter Interaktionsstrategien unterscheiden, um auf diese Weise ihre impliziten Vorstellungen über Lernprozessdiagnose im Dialog zu erfassen. Sind erfahrene Lehrkräfte eher in der Lage, bei der Beobachtung und Analyse eines Lehrer-Schüler-Dialogs, in diagnostischer Hinsicht förderliche sowie hinderliche Strategien zu identifizieren als Lehramtsstudierende? Wie unterscheiden sich angehende und erfahrene Lehrkräfte in der Interpretation und Bewertung beobachteter Strategien? An unserer Studie nahmen 22 Biologie-Lehramtsstudierende der Universität Freiburg und 22 Biologielehrkräfte (Gymnasialzweig) teil. Die Versuchsteilnehmenden sahen ein fünf-minütiges Video zum Thema Blutkreislauf eines nach theoretischen Gesichtspunkten konstruierten Dialogs zwischen einem Lehrer und einem Schüler. In diesem Video wurden unterschiedliche Interaktionsstrategien seitens des Lehrers realisiert, welche auf Grundlage empirischer Forschungsbefunde a priori als förderlich bzw. hinderlich für die Diagnose kategorisiert worden waren. Die Teilnehmenden beobachteten das Video und drückten eine grüne bzw. rote Taste, wenn sie eine Interaktionsstrategie des Lehrers als förderlich bzw. hinderlich für die Lernprozessdiagnose bewerteten. Nach der Videobeobachtung wurde jede/r Teilnehmer/in dazu befragt, was ihm/ihr an den jeweiligen durch Tastendruck markierten Stellen aufgefallen war, in wie weit der Lehrer durch die Strategie Informationen über den Lernprozess des Schülers sammeln konnte und was der Lehrer anders oder besser hätte machen können. Aufgrund unserer theoriegestützten a priori Kategorisierung gab es bei der Auswertung pro Interaktionsstrategie sechs Optionen: Förderliche Strategie nicht bemerkt, hinderliche Strategie nicht bemerkt, förderliche Strategie als förderlich bewertet, hinderliche Strategie als hinderlich bewertet, förderliche Strategie als hinderlich bewertet (falsch negativ) und hinderliche Strategie als förderlich bewertet (falsch positiv). Die Ergebnisse zeigen, dass erfahrene Lehrkräfte signifikant mehr diagnoserelevante Interaktionsstrategien bemerkten als Lehramtsstudierende. Bei detaillierterer Betrachtung der fördernden und hinderlichen Interaktionsstrategien wurde deutlich, dass die erfahrenen Lehrkräfte signifikant häufiger die von ihnen identifizierten Strategien als hinderlich bewerteten, auch wenn diese nach unserer Kategorisierung förderlich für die Lernprozessdiagnose waren („Negativity-Bias“). Lehramtsstudierende hingegen bewerteten Interaktionsstrategien signifikant häufiger als förderlich, obwohl wir sie a priori als hinderlich klassifiziert hatten („Positivity-Bias“). Offenbar nahmen Lehramtsstudierende und erfahrene Lehrkräfte Diagnose-relevante Interaktionsstrategien unterschiedlich wahr. Zudem scheinen sich die Bewertungen von der auf wissenschaftlicher Basis von uns vorgenommenen normativen Klassifizierung der Strategien systematisch und gruppenspezifisch zu unterscheiden. Derzeit führen wir eine inhaltsanalytische Auswertung der Interviewdaten durch, um diese Schlussfolgerungen näher zu beleuchten und weiter zu untermauern. Unsere Ergebnisse sollen dazu beitragen, die Aus- und Weiterbildung von Lehrkräften im Bereich der Lernprozessdiagnose zu verbessern.

## **Noten oder sozialer Hintergrund?**

### **Der erste Blick beeinflusst das Gedächtnis für Schülerinformationen und die Genauigkeit des Urteils**

Wenn Lehrkräfte ihre Schüler beurteilen, fließt als wichtigster Faktor die Schulleistung in das Urteil ein (Nölle, Hörstermann, Krolak-Schwerdt, & Gräsel, 2009). Allerdings sind Lehrerurteile nicht frei von leistungsfernen Informationen, wie der soziale Hintergrund eines Schülers (Glock, Krolak-Schwerdt, Klapproth, & Böhmer, 2013). Ein möglicher Mechanismus, der diesem Einfluss zugrunde liegen könnte, ist die Art und Weise wie Schülerinformationen verarbeitet und gespeichert werden. Aus der sozialen Kognitionsforschung ist bekannt, dass verschiedene Arten von Informationen unterschiedlichen Einfluss auf das Gedächtnis und das Urteil über Personen haben. Die Information, die zuerst wahrgenommen wird, wird am besten erinnert (Hastie & Kumar, 1979; Srull & Wyer, 1989) und bestimmt maßgeblich das Urteil (Hastie & Park, 1986). Dies kann über die Theorie der Assoziativen Netzwerke (Smith, 1998) erklärt werden. Jede Information stellt einen Knotenpunkt dar, der über Links mit anderen Knoten und der Person verknüpft wird. Die erste wahrgenommene Information stellt den ersten Knoten dar, der mit der Person verbunden wird. Jede weitere Information wird in Verbindung mit dieser ersten Information wahrgenommen und interpretiert. Somit erhält die erste Information die stärksten und auch die größte Anzahl an Links, was es wahrscheinlicher macht, dass diese Information erinnert wird. Dadurch, dass weitere Informationen im Lichte der ersten Information wahrgenommen und interpretiert werden, bestimmt diese auch das Urteil. Forschung zur Personenwahrnehmung hat bisher immer determiniert, welche Information zuerst dargeboten wird. In der vorliegenden Studie wurde die gesamte Schülerinformation gleichzeitig auf einer Bildschirmseite dargeboten und mithilfe eines Eyetracking-Verfahrens aufgezeichnet, welche Art von Schülerinformation die Probanden zuerst beachteten. Da sowohl Noten als auch soziale Hintergrundinformation eine wesentliche Rolle in den Lehrerurteilen spielen, wurde deren Position variiert. Einer Gruppe von Probanden wurde die Noteninformation auf der oberen linken Hälfte des Bildschirms dargeboten, einer anderen Gruppe wurde die soziale Hintergrundinformation an dieser Position dargeboten. Die Aufgabe der Probanden war das Treffen einer Schullaufbahnpfehlung für jeden der vier Schüler. Anschließend erfolgte eine kurze Distraktoraufgabe und abschließend eine Free Recall-Aufgabe, in welcher die Probanden die Informationen notierten, die sie noch von den einzelnen Schülerprofilen erinnern konnten. 40 Studierende der Universität Luxemburg nahmen an der Studie teil. Die Schülerprofile beruhten auf tatsächlichen Schülerfällen aus Luxemburg. Für diese Fälle war die korrekte Schullaufbahnpfehlung bekannt. Die Ergebnisse der Eye-tracking Daten zeigten, dass die Probanden sich tatsächlich zuerst der Information zuwandten, die auf der oberen linken Hälfte des Bildschirms dargeboten wurde. Die Gruppe, die die Noten an dieser Position hatte, betrachtete die Noteninformation insgesamt länger als alle andere Information, während dieser Effekt für die soziale Hintergrundinformation in der anderen Gruppe nicht gefunden werden konnte. Die Gruppe, die die Noten an erster Position hatte, erinnerte die Noteninformation besser, sowohl im Vergleich zu der anderen enthaltenen Information auf den Schülerprofilen als auch im Vergleich zu der anderen Gruppe, die zuerst die soziale Hintergrundinformation erhielt. Diese Gruppe erinnerte die soziale Hintergrundinformation besser als die Gruppe, die die Noten zuerst erhielt. Die Akkuratheit der Schullaufbahnpfehlungen war bei der Gruppe, die die sozialen Hintergrundinformationen an erster Stelle erhielt, höher als bei der Gruppe, die die Noten an erster Position erhielt. Im Vergleich zu erfahrenen Lehrkräften, die die gleichen Schülerprofile mit sozialer Hintergrundinformation an erster Stelle erhielten, aber selbst jedoch angaben, hauptsächlich auf die Noteninformation zu achten, erwiesen sich die Urteile der Gruppe, die die soziale Hintergrundinformationen zuerst erhielten, ebenfalls als genauer. Dieser Befund kann über die Widersprüche in der Noteninformation erklärt werden, die eher ins Gewicht fallen, wenn die Noten die erste wahrgenommene Information darstellen.

Obwohl noch mehr Forschungsbedarf besteht, liefert diese Studie erste Hinweise darauf, wie wichtig die Noteninformation für die Beurteilung von Schülern ist und zeigt gleichzeitig auf, dass die soziale Hintergrundinformationen diagnostisch einen geringeren Stellenwert hat.

## ***Die Rolle der Lehrkraft für die Unterrichtsqualität und den Lernerfolg von Schülerinnen und Schülern***

Chair(s): Dirk Richter (Institut zur Qualitätsentwicklung im Bildungswesen), Uta Klusmann (IPN Kiel)

Diskutant/in: Olaf Köller (IPN Kiel)

Die professionelle Kompetenz von Lehrkräften gilt als eine zentrale Voraussetzung für gelingende Lehr-Lernprozesse im Unterricht und damit für den schulischen Lernerfolg von Schülerinnen und Schülern. In Deutschland beschäftigten sich in den vergangenen 10 Jahren mehrere Studien mit der Erfassung professioneller Kompetenz von Lehrkräften (z.B. COACTIV; Kunter et al., 2011; TEDS-M: Blömeke, Kaiser, & Lehmann, 2010). Trotz der vergleichsweise umfangreichen Forschung zur Modellierung professioneller Kompetenz von Lehrkräften gibt es im deutschsprachigen Raum nur wenige Arbeiten, die sich mit dem Zusammenhang zwischen Merkmalen von Lehrkräften und den erreichten Schülerleistungen beschäftigen. Dieses Symposium untersucht deshalb in vier Beiträgen die Bedeutung psychologischer Merkmale und des Fortbildungsverhalten für die Qualität des Unterrichts und den Lernerfolg von Schülerinnen und Schülern.

Der Beitrag von Voss et al. untersucht an einer Stichprobe junger Mathematiklehrkräfte den Zusammenhang zwischen pädagogisch-psychologischem Wissen und der Unterrichtsqualität aus Sicht von Schülerinnen und Schülern. Der Beitrag von Klusmann et al. nimmt das berufliche Wohlbefinden von Grundschullehrkräften in den Blick und untersucht dessen Relevanz für den Lernerfolg von Schülerinnen und Schülern im Bereich Lesen und Mathematik. Der Beitrag von Richter et al. analysiert Überzeugungen von Lehrkräften zu den wahrgenommenen Funktionen von den bundesweit durchgeführten Vergleichsarbeiten und untersucht deren Bedeutung für die Gestaltung von Unterricht und die von Schülerinnen und Schülern erreichten Kompetenzen im Lesen und Mathematik. Im Aufsatz von Dalehefte et al. steht die Evaluation des mathematisch-naturwissenschaftlichen Fortbildungsprogramms SINUS an Grundschulen im Zentrum. Die Autoren gehen der Frage nach, ob sich Lehrkräfte des SINUS-Programms von Lehrkräften anderer Grundschulen in ihrem Kooperations- und Weiterbildungsverhalten sowie in den von ihren Schülerinnen und Schülern erbrachten Leistungen in Mathematik unterscheiden.

Dieses Symposium vereint psychologisch ausgerichtete Beiträge (Voss et al. und Klusmann et al.) und Beiträge mit einem stärkeren erziehungswissenschaftlichen Schwerpunkt (Richter et al und Dalehefte et al.). Damit repräsentiert dieses Symposium einen Teil der Disziplinenvielfalt, die sich auch in der Forschung zur Professionalität von Lehrkräften wiederfindet.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
IG-Farben-Haus  
Raum 311

Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010). TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich. Münster: Waxmann.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV. Waxmann: Münster.

Thamar Voss (Eberhard Karls Universität Tübingen), Verena Hoehne (Eberhard Karls Universität Tübingen), Johanna Seiz (Goethe-Universität Frankfurt), Mareike Kunter (Goethe-Universität Frankfurt)

## ***Pädagogisch-psychologisches Wissen, Überzeugungen und Motivation von Lehrkräften: Bedeutung für Unterrichtsqualität und Lehrerwohlbefinden***

Theoretischer Hintergrund: Die professionelle Kompetenz von Lehrkräften wird in neueren Forschungsansätzen zunehmend in den Vordergrund gestellt und als notwendige Voraussetzung für erfolgreiches Handeln im Unterricht betrachtet (Baumert & Kunter, 2006; Bromme, 1997). Professionelle Kompetenz wird dabei als mehrdimensionales Konstrukt aufgefasst, das aus kognitiven sowie nicht-kognitiven Aspekten besteht. Im Rahmen der COACTIV-Studie werden beispielsweise motivationale Orientierungen und die Fähigkeit zur Selbstregulation, Überzeugungen sowie professionelles Wissen von Lehrkräften unterschieden (Kunter et al., 2011).

Fragestellung: Im folgenden Beitrag wird der Frage nachgegangen, welche Bedeutung kognitive und nicht-kognitive Aspekte der professionellen Kompetenz von Referendarinnen und -referendaren für deren späteren Berufserfolg haben. Es wurden das pädagogisch-psychologische Wissen (PPK) und die lerntheoretischen Überzeugungen als kognitive Aspekte sowie der Enthusiasmus der Lehrkräfte als motivationaler Aspekt fokussiert.

Berufserfolg wurde multikriterial betrachtet und zwei Jahre später einerseits die Unterrichtsqualität aus Schülersicht, andererseits das Wohlbefinden der Lehrkräfte erfasst.

Methode: Die Datengrundlage stellt die COACTIV-R-Studie dar, eine Studie zum Erwerb professioneller Kompetenz von Mathematiklehramtskandidatinnen und -kandidaten während des Referendariats (Kunter et al., 2011). Es wurde auf Daten von zwei Messzeitpunkten zurückgegriffen: Zu MZP1 wurden 181 Referendarinnen und Referendare während des Vorbereitungsdienst untersucht. Zu MZP2 wurden ca. zwei Jahre später – als die vormaligen Referendare bereits im Schuldienst standen – deren 7968 Schülerinnen und Schüler zur Unterrichtsqualität befragt sowie eine Lehrerbefragung u.a. zu deren beruflichem Wohlbefinden durchgeführt.

Als Aspekte der professionellen Kompetenz von Lehrkräften wurde zu MZP1 PPK anhand eines Messinstruments mit 39 Items (Voss et al., 2011) erfasst, die sich auf Wissen über eine effiziente Klassenführung, Unterrichtsmethoden, Leistungsbeurteilung und Schülerheterogenität bezogen ( $\alpha=.81$ ). Weiterhin wurde die Zustimmung zu traditionellen lerntheoretischen Überzeugungen, wie z.B. der Auffassung, dass Lernen ein Prozess der Wissenstransmission ist ( $\alpha=.78$ , vgl. Dubberke et al., 2008) sowie der Enthusiasmus für das Unterrichten ( $\alpha=.85$ , Kunter et al., 2008) erhoben.

Zu MZP2 schätzten die Schülerinnen und Schüler die Effizienz der Klassenführung (Unterrichtsstörungen und Monitoring), das Potenzial des Unterrichts zur kognitiven Aktivierung sowie die konstruktive Unterstützung ein (vgl. Kunter et al., 2011,  $.55 < \alpha < .85$ ,  $.11 < ICC < .39$ ). Die Lehrkräfte bearbeiteten zu MZP2 u.a. Fragen zu deren emotionaler Erschöpfung ( $\alpha=.75$ , vgl. Maslach & Jackson, 1996), beruflichem Commitment ( $\alpha=.69$ , Meyer et al., 1993) und Lebenszufriedenheit ( $\alpha=.89$ , Pavot, Diener & Suh, 1998).

Emotionale Erschöpfung wurde zusätzlich aus Schülersicht erfasst ( $\alpha=.83$ ,  $ICC=.25$ ).

Statistische Verfahren

Mithilfe von Zwei-Ebenen-Strukturgleichungsmodellen wurde analysiert, ob die Aspekte professioneller Kompetenz der Referendare (MZP1) prädiktiv für deren späteren Berufserfolg (MZP2) sind. In einem Modell wurde die Unterrichtsqualität aus Schülersicht als Kriterium verwendet, in einem zweiten Modell das Wohlbefinden der Lehrkräfte vorhergesagt.

Ergebnisse: Die Ergebnisse zeigen, dass alle untersuchten Aspekte der professionellen Kompetenz signifikant zur Vorhersage des späteren Berufserfolgs beitragen, jedoch mit unterschiedlichen Befundmustern: Bei der Vorhersage der Unterrichtsqualität gaben Schülerinnen und Schüler von Lehrkräften, die während des Referendariats ein höheres PPK und einen höheren Enthusiasmus für das Unterrichten aufwiesen, zu MZP2 eine höhere konstruktive Unterstützung im Unterricht an (Enthusiasmus:  $\beta=.20$ , PPK:  $\beta=.34$ ). PPK sagte weiterhin das spätere Monitoring ( $\beta=.21$ ), der Enthusiasmus ein geringes Ausmaß an Unterrichtsstörungen ( $\beta=-.19$ ) vorher. Zudem berichteten Lernende von Lehrkräften mit weniger transmissiven Überzeugungen von einem signifikant höheren Potenzial zur kognitiven Aktivierung ( $\beta=-.32$ ).

Für das Wohlbefinden der Lehrkräfte zu MZP2, erwies sich insbesondere der Enthusiasmus zu MZP1 als bedeutsam: Lehrkräfte, die während des Referendariats enthusiastischer waren, gaben an sich zu MZP2 ihrem Beruf stärker verbunden zu fühlen ( $\beta=.40$ ), zufriedener ( $\beta=.25$ ) und weniger erschöpft ( $\beta=-.21$ ) zu sein. Berufliches Commitment war zudem höher für Lehrkräfte mit weniger transmissiven Überzeugungen zu MZP1.

Die Ergebnisse liefern wichtige Hinweise auf die Bedeutung von kognitiven sowie nicht-kognitiven Aspekten der professionellen Kompetenz für den späteren beruflichen Erfolg angehender Lehrkräfte und bieten Ansatzpunkte zur Steigerung der Unterrichtsqualität und des Wohlbefindens der Lehrkräfte.

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469–520. doi: 10.1007/s11618-006-0165-2

Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Ed.), Enzyklopädie der Psychologie: Psychologie des Unterrichts und der Schule (pp. 177–212). Göttingen: Hogrefe.

Dubberke, T., Kunter, M., McElvany, N., Brunner, M., & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften: Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. Zeitschrift für Pädagogische Psychologie, 22(3–4), 193–206.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2011). Professionelle Kompetenz von Lehrkräften – Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann.

Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. Learning and Instruction, 18(5), 468–482.

Maslach, C., & Jackson, S.E. (1996). Maslach Burnout Inventory Manual (3rd ed.). Palo Alto, CA: Consulting Psychological Press.

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to Organizations and Occupations: Extension and Test of the three Component Conceptualization. Journal of Vocational Behavior, 78(4), 538–551.

Pavot, W., Diener, E., & Suh, E. (1998). The Temporal Satisfaction With Life Scale. Journal of Personality Assessment, 70(2), 304–354.

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. Journal of Educational Psychology, 103(4), 952–969.

GEBF TAGUNG  
2014

F02 a

Symposiums  
beitrag

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
IG-Farben-Haus  
Raum 311



Die Rolle der Lehrkraft für die Unterrichtsqualität und den Lernerfolg



## **Beanspruchungserleben von Lehrkräften und Schülerleistung: Eine Analyse des IQB-Ländervergleichs in der Primarstufe**

Theoretischer Hintergrund: Die Bedeutung der Lehrkräfte für die schulische Entwicklung von Schülerinnen und Schülern wurde in den letzten Jahren vielfach betont (Hattie, 2009; Nye, Konstantopoulos, & Hedges, 2004). Mittlerweile konnte gezeigt werden, dass neben dem Wissen und Können von Lehrkräften auch motivationale und emotionale Merkmale, wie Enthusiasmus, Selbstwirksamkeitserwartungen und berufliche Selbstregulation einen Effekt auf die Unterrichtsgestaltung und das Erleben und Verhalten der Schülerinnen und Schüler haben (Baumert et al., 2010; Caprara, Barbaranelli, Steca, & Malone, 2006; Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009; Kunter et al., 2013). Inwiefern neben positiven emotionalen und motivationalen Merkmalen von Lehrkräften auch negative Erlebensqualitäten wie Erschöpfung, Burnout und Stress Konsequenzen für die Schülerinnen und Schüler haben, wurde bislang kaum empirisch untersucht. Dabei wird in der Forschung zu Burnout bei Lehrkräften schon lange vermutet, dass die von Lehrkräften erlebte Beanspruchung negative Konsequenzen für ihr Unterrichtsverhalten und somit auch für die Motivation und Leistung der Schülerinnen und Schüler hat (Maslach & Leiter, 1999). Fragestellung: Die aktuelle Studie untersucht die Frage, inwieweit die berufliche Beanspruchung von Lehrkräften in Zusammenhang mit der Leistung ihrer Schülerinnen und Schüler steht. Den Ausgangspunkt bildet die Annahme, dass sich die motivationalen, kognitiven und expressiven Konsequenzen eines erhöhten Beanspruchungserlebens in der Unterrichtsvorbereitung, der Unterrichtsdurchführung sowie im Interaktionsverhalten manifestieren, was sich wiederum auf das Erleben und Verhalten der Schülerinnen und Schüler auswirkt (Maslach & Leiter, 1999). Bisherige empirische Studien legen nahe, dass das Beanspruchungserleben von Lehrkräften für die Schülerinnen und Schüler erkennbar wird (Evers et al., 2004) und sich in einer ungünstigeren Unterrichtsgestaltung, wie z.B. einer geringeren individuellen Unterstützung der Schülerinnen und Schüler, manifestiert (Klusmann et al., 2006). Für die Unterrichtsqualität wiederum konnte mehrfach gezeigt werden, dass sie – neben individuellen kognitiven und motivationalen Voraussetzungen der Schülerinnen und Schüler – maßgeblich zur Leistungsentwicklung beiträgt (Helmke, 2012).

Methode: Die Datenbasis bilden eine repräsentative Befragung von Schülerinnen und Schüler und ihrer Lehrkräfte im Rahmen des IQB Ländervergleichs 2011 in der Primarstufe (Stanat, Pant, Böhme, & Richter, 2012). Es stehen im Fach Deutsch für insgesamt 22389 Schüler(innen) und 1120 Lehrkräfte und im Fach Mathematik für 22002 Schüler(innen) und 1104 Lehrkräfte Angaben zur Verfügung.

Operationalisiert wird die berufliche Beanspruchung über das Leitsymptom von Burnout, die emotionale Erschöpfung (Maslach, Schaufeli, & Leiter, 2001). Die Leistung der Schülerinnen und Schüler wurde mit Leistungstests im Bereich Lesen und Mathematik erfasst (Weirich et al., 2012). Auf der Seite der Schülerinnen und Schüler kontrollieren wir für kognitive, motivationale und soziodemographische Merkmale sowohl auf individueller als auch auf Klassenebene. Auf der Seite der Lehrkräfte kontrollieren wir die Berufserfahrung, die quantitative Arbeitsbelastung und die Lehrbefähigung im entsprechenden Fach.

Zur Vorhersage der Kompetenzen der Schülerinnen und Schüler werden Mehrebenenmodelle verwendet (Raudenbush & Bryk, 2002).

Ergebnisse: Es zeigte sich, dass die mittlere Mathematik- und Leseleistung in Klassen, die von Lehrkräften mit einer höheren emotionalen Erschöpfung unterrichtet wurden, statistisch signifikant geringere ist als in Klassen, die von weniger erschöpften Lehrkräften unterrichtet werden. Diese Effekte sind auch unter Kontrolle von Berufserfahrung, Arbeitsumfang und Qualifikation der Lehrkräfte im Fach stabil. Wurden die Kompositionseffekte in soziodemographischen, motivationalen und kognitiven Voraussetzungen auf Klassenebene kontrolliert, konnte für die Leseleistung weder ein signifikanter Effekt der emotionalen Erschöpfung noch der übrigen Lehrkräfte mehr gefunden werden, für die Mathematikleistung hingegen blieb der Effekt der emotionalen Erschöpfung bestehen.

Die Studie liefert insgesamt anhand einer für Deutschland repräsentativen Schüler- Stichprobe erste Hinweise dafür, dass die Leistung von Schülerinnen und Schülern nicht nur mit kognitiven Merkmalen ihrer Lehrkraft assoziiert ist, sondern auch mit negativen emotionalen Erlebensqualitäten zusammenhängt. Demnach scheint eine erhöhte berufliche Beanspruchung nicht nur Kosten für die Lehrkraft selbst, sondern auch für die Performanz im schulischen Kontext haben kann.

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A. et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133-180.
- Caprara, G. V., Barbaranelli, C., Steca, P. & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44 (6), 473-490.
- Evers, W. J. G., Tomic, W. & Brouwers, A. (2004). Burnout among teachers: Students' and teachers' perceptions compared. *School Psychology International*, 25 (2), 131-148.
- Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R. & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101(3), 705-716.
- Hattie, J. A. C. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Helmke, A. (2012). Unterrichtsqualität und Lehrprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts (4. überarbeitete Aufl., Schule weiterentwickeln – Unterricht verbessern. Orientierungsband). Seelze: Klett-Kallmeyer.
- Klusmann, U., Kunter, M., Trautwein, U. & Baumert, J. (2006). Lehrerbelastung und Unterrichtsqualität aus der Perspektive von Lehrenden und Lernenden. *Zeitschrift für Pädagogische Psychologie*, 20 (3), 161-173.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105 (3), 805-820.
- Maslach, C. & Leiter, M. P. (1999). Teacher burnout: A research agenda. In R. Vandenburghe & A. M. Huberman (Eds.), *Understanding and preventing teacher stress: A sourcebook of international research and practice* (pp. 295-303). Cambridge, UK: Cambridge University Press.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52, 397-422.
- Nye, B., Konstantopoulos, S. & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237-257.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage Publications.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (2012). Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011. Münster: Waxmann.
- Weirich, S., Haag, N. & Roppelt, A. (2012). Testdesign und Auswertung des Ländervergleichs: Technische Grundlagen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 277-290). Münster: Waxmann.



Dirk Richter (IQB, Humboldt-Universität zu Berlin), Katrin Böhme (IQB, Humboldt-Universität zu Berlin), Michael Becker (DIPF), Hans Anand Pant (IQB, Humboldt-Universität zu Berlin), Petra Stanat (IQB, Humboldt-Universität zu Berlin)

## **Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten: Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen von Schülerinnen und Schülern**

Theoretischer Hintergrund: Die Vergleichsarbeiten (VERA) bilden seit mehreren Jahren ein wichtiges Instrument der Kompetenzdiagnostik, das auf den Bildungsstandards der Kultusministerkonferenz basiert. Sie dienen in erster Linie der Unterrichts- und Schulentwicklung, werden teilweise aber auch zur flächendeckenden Information der Schulaufsicht über den Leistungsstand der Einzelschulen genutzt (KMK, 2012). Welche dieser beiden Funktionen Lehrkräfte VERA zuschreiben, ist bislang in der empirischen Forschung kaum betrachtet worden. Es gibt jedoch Hinweise darauf, dass die von Lehrkräften wahrgenommenen Funktionen von VERA nicht immer deckungsgleich mit den intendierten Funktionen dieses Instruments sind (z.B. Kühle & Peek, 2007). Da die meisten empirischen Arbeiten zu diesem Thema bereits in den Anfangsjahren der Vergleichsarbeiten entstanden sind, ist unklar, ob Lehrkräfte gut 10 Jahre nach der Einführung von VERA nach wie vor ähnliche Überzeugungen aufweisen.

Die mit den Leistungstests verbundenen Funktionen sind bedeutsam für die Nutzung der Testergebnisse durch die Lehrkraft und damit auch für die Gestaltung des Unterrichts und den Lernerfolg der Schülerinnen und Schüler (Maag Merki, 2010). Tests, die primär der Unterrichtsentwicklung dienen, bieten Lehrkräften durch die Rückmeldung der Leistungsergebnisse die Möglichkeit, die Stärken und Schwächen der eigenen Klasse anhand eines standardisierten Instruments zu diagnostizieren und Maßnahmen zur Förderung der Kompetenzen zu ergreifen (z.B. Helmke, 2004). Tests, die zur Rechenschaftslegung schulischer Leistungen auf der Ebene der Klasse oder der Schule verwendet werden, können dazu beitragen, dass der Unterricht primär auf die Verbesserung der Testergebnisse ausgerichtet wird und somit Tests auch negative Folgen, wie etwa eine Verengung des Curriculums, mit sich bringen (Maag Merki, 2010).

Für die deutschen Vergleichsarbeiten gilt bisher als belegt, dass die Rückmeldungen der Testergebnisse nur von einem geringen Teil der Lehrkräfte zur Ableitung von Entwicklungsmaßnahmen im Unterricht herangezogen werden (Dedering, 2011). Gleichzeitig gibt es bislang allerdings auch keine Hinweise darauf, dass Vergleichsarbeiten zu unerwünschten Effekten geführt hätten (Wacker & Kramer, 2012). In den bisherigen Arbeiten wurde jedoch nicht berücksichtigt, dass die Vergleichsarbeiten neben der Entwicklungsfunktion teilweise auch zur Rechenschaftslegung der Leistungen von Lehrkräften und Schulen herangezogen werden. Es ist daher offen, inwiefern die von Lehrkräften wahrgenommenen Funktionen von VERA bedeutsam für das unterrichtliche Handeln und den Lernerfolg von Schülerinnen und Schülern sind.

Fragestellung: Der vorliegende Beitrag untersucht, inwieweit VERA als Instrument der Unterrichtsentwicklung bzw. der Rechenschaftslegung und Kontrolle von Lehrkräften wahrgenommen wird. Nach einer deskriptiven Betrachtung der Konstrukte werden Zusammenhänge zwischen den von Lehrkräften wahrgenommenen Funktionen einerseits und Merkmalen von Unterricht und den Leistungen von Schülerinnen und Schülern im Lesen und in Mathematik andererseits untersucht.

Methode: Die vorliegende Studie basiert auf den Daten des IQB-Ländervergleichs 2011, an dem über 27000 Schülerinnen und Schüler der vierten Jahrgangsstufe teilnahmen und in dem die Leistungen in den Fächern Deutsch und Mathematik getestet wurden (Stanat, Pant, Böhme, & Richter, 2012). Die Lehrkräfte der getesteten Schülerinnen und Schüler beantworteten Fragen zu den wahrgenommenen Funktionen von VERA und zu Veränderungen im Unterricht durch die Einführung von Schulleistungstests. Die Analysen basieren auf Mehrebenenmodellen, in denen die selbst berichteten Veränderungen im Unterricht sowie die Kompetenzen der Schülerinnen und Schüler durch die wahrgenommenen Funktionen von VERA unter Kontrolle von schüler- und lehrerseitigen Hintergrundmerkmalen vorhergesagt wird.

Ergebnisse: Aus Sicht der befragten Grundschullehrkräfte dienen Vergleichsarbeiten weder in besonderem Maße zur Unterrichtsentwicklung noch zur Kontrolle von Schulen. Mehrebenenanalysen zur Vorhersage des Unterrichts zeigen, dass Lehrkräfte, die die Vergleichsarbeiten als Mittel der Unterrichtsentwicklung ansehen, stärker dazu tendieren, ihren Unterricht auf den Erwerb von Kompetenzen auszurichten und auf die Leistungsunterschiede in der Klasse mit Differenzierungsmaßnahmen einzugehen. Die vorliegende Studie konnte ebenfalls Belege dafür finden, dass die Überzeugungen zu den Vergleichsarbeiten mit den Leistungen der Schülerinnen und Schüler in einem statistisch bedeutsamen Zusammenhang stehen, der auch unter Berücksichtigung zentraler individueller und klassenbezogener Hintergrundmerkmale signifikant bleibt.

Dedering, K. (2011). Hat Feedback eine positive Wirkung? Zur Verarbeitung extern erhobener Leistungsdaten in Schulen. *Unterrichtswissenschaft*, 39(1), 63–83.

Helmke, A. (2004). Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. *Das Seminar*(2), 90-112.

KMK. (2012). Vereinbarung zur Weiterentwicklung von VERA. Berlin: KMK.

Kühle, B. & Peek, R. (2007). Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnissrückmeldungen in Schulen. *Empirische Pädagogik*, 21(4), 428–447.

Maag-Merki, K. (2010). Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 145–169). Wiesbaden: Verlag für Sozialwissenschaften.

Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

Wacker, A. & Kramer, J. (2012). Vergleichsarbeiten in Baden-Württemberg. *Zeitschrift für Erziehungswissenschaft*, 15(4), 683–706.



## Neun Jahre SINUS in deutschen Grundschulen: Evaluation im Rahmen der TIMSS 2011-Erhebung

Theoretischer Hintergrund: SINUS an Grundschulen (SGS) ist ein Lehrerprofessionalisierungs- und Unterrichtsentwicklungsprogramm mit dem Ziel, die Effizienz des mathematisch- naturwissenschaftlichen Unterrichts zu steigern. Der Erfolg des Programms wird in der Begleitforschung auf der Ebene der Schulen, der Lehrkräfte, des Unterrichts und der Leistungen der Schülerinnen und Schüler geprüft (Fischer et al., 2012). In diesem Beitrag wird SGS in Bezug zu einem programmunabhängigen Außenkriterium (Kontrollgruppe) mit Hilfe von Daten aus „Trends in International Mathematics and Science Study“ (TIMSS) 2011 (Bos et al., 2012) gesetzt. Ziel eines Vergleichs ist es, mehr über die Wirksamkeit des Programms zu erfahren. An dieser Stelle soll auf das Fortbildungsverhalten und die Professionalisierungsmöglichkeiten der Lehrkräfte im Bereich der Mathematik fokussiert werden. Anschließend wird auf die Mathematikkompetenzen der Schülerinnen und Schüler eingegangen.

Soll eine Professionalisierungsmaßnahme gelingen, muss es Strukturen geben, die Veränderungen des Wissens, der Fähigkeiten und der Einstellungen der Lehrkräfte erleichtern (Hochberg & Desimone, 2010). Schulen und Lehrkräfte brauchen Unterstützung und Anregung durch inner- und außerschulische Expertise (Timperley, 2008). Ein solches Gerüst an Strukturen wird vom Programmträger (IPN, Kiel) durch die bundesweite Koordination des Programms und das Angebot an Fortbildungsveranstaltungen gewährleistet. Zehn Module und fünf Schwerpunkte, die theoretisch und empirisch fundiert sind, bilden die Grundlage des Programms. Sie basieren auf Problembereichen des deutschen Bildungssystems, die aus Befunden internationaler Schulleistungsstudien abgeleitet und in alltagstauglichen, handlungs-orientierten und beispielbezogenen Beschreibungen erhältlich sind (Demuth, Walther, & Prenzel, 2011). Lehrkräfte und Schulen haben in SGS viele Freiräume. Sie setzen selbst Schwerpunkte und steuern eigenverantwortlich die Umsetzung des Programms an ihrer Schule.

Fragestellung: Aufgrund der vielen Freiräume des SGS-Programms ist es von Interesse zu erfahren, inwieweit die Mathematiklehrkräfte Fortbildungsangebote und –möglichkeiten verstärkt wahrnehmen. Auf Seiten der SGS-Lehrkräfte wird vermutet, dass sie verstärkt Gelegenheiten zur Professionalisierung wahrnehmen. Auf Seiten der Schülerinnen und Schüler wäre zu erwarten, dass die Teilnahme der SGS-Lehrkräfte am Professionalisierungskonzept in höheren mathematischen Kompetenzen resultiert.

Methode: Es wurden für die Erhebung 78 SGS-Schulen einbezogen, die mindestens seit fünf Jahren am SINUS beteiligt waren. Die SGS-Gesamtstichprobe bestand aus 1.581 Schülerinnen und Schülern, 78 Schulleitungen, sowie 105 Lehrpersonen, von denen 68 Mathematiklehrpersonen in die vorliegende Studie eingehen. Die als Vergleichsmaßstab herangezogene Stichprobe aus TIMSS 2011 umfasst Daten für 3.995 Grundschulkindern, 197 Schulleitungen und 206 Mathematiklehrkräfte (Wendt, Tarelli, Bos, Frey, & Vennemann, 2012). Die Befunde zur Professionalisierung der Lehrkräfte beziehen sich auf Häufigkeitsanalysen. Die Leistungsdaten (Plausible Values) der Schülerinnen und Schüler wurden unter Berücksichtigung bestimmter Kontrollvariablen mit Tests auf Mittelwertsunterschiede analysiert.

Ergebnisse: Insgesamt geben die Befunde wichtige Hinweise über das Fortbildungsverhalten der Lehrkräfte und die dazugehörigen Rahmenbedingungen im Programm. In SGS werden im Vergleich zu TIMSS mehr Schülerinnen und Schüler von Mathematiklehrkräften unterrichtet, die angeben, in den letzten zwei Jahren Fortbildungen zu mathematischen und mathematikdidaktischen Inhalten, zum Lehrplan, zur Integration von Kommunikationstechniken, zur Leistungsfeststellung sowie zum Eingehen auf individuelle Bedürfnisse der Lernenden besucht zu haben. Die SGS-Werte sind am höchsten in Bereichen, die mit dem SINUS-Ansatz einhergehen. Mathematiklehrkräfte in SGS zeigen generell eine hohe Fortbildungsbereitschaft, die sich auch auf Fortbildungsbereiche über die SINUS-Inhalte hinaus erstreckt (z.B. Kommunikationstechniken). Bei der wahrgenommenen Möglichkeit zum Erwerb der Medienkompetenz gibt es ebenfalls signifikante Unterschiede zugunsten von SGS. Die SGS-Lehrkräfte erhalten jedoch nicht signifikant mehr Zeit für Fortbildungen oder mehr Anreize seitens der Schule als die Kolleginnen und Kollegen in TIMSS. Im Hinblick auf die Kompetenzwerte der Lernenden erzielen SGS-Schülerinnen und -Schüler höhere Testwerte bei den Mathematikaufgaben als die Lernenden in TIMSS. Weitere Analysen sind jedoch notwendig, um festzustellen, ob die Kompetenzunterschiede der Schülerinnen und Schüler zu Gunsten von SGS auf die Wirksamkeit des Programms zurückzuführen sind.

Bos, W., Wendt, H., Köller, O. & Selter, C. (2012). TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.

Demuth, R., Walther, G. & Prenzel, M. (2011). Unterricht entwickeln mit SINUS. 10 Module für den Mathematik- und Sachunterricht in der Grundschule. Seelze: Friedrich Verlag.

Fischer, C., Kobarg, M., Dalehefte, I. M. & Trepke, F. (2012). Wirkungen von Maßnahmen zur Lehrerprofessionalisierung feststellen - Unterrichtsentwicklung im Programm "SINUS an Grundschulen" mit verschiedenen Instrumenten und Methoden erheben. In: Gläser-Zikuda, M., Seidel, T., Rohlf, C., Gröschner, A. & Ziegelbauer, S. (Hrsg.). Mixed Methods in der empirischen Bildungsforschung. Münster: Waxmann

Hochberg, E.D. & Desimone, L.M. (2010). Professional Development in the Accountability Context: Building Capacity to Achieve Standards. Educational Psychologist, 45(2), 89-106.

Timperley, H. (2008). Teacher Professional Learning and Development. Netherlands: International Academy of Education / International Bureau of Education.

Wendt, H., Tarelli, I., Bos, W., Frey, K. & Vennemann, M. (2012). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2011). In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.) TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 27-68). Münster: Waxmann.



## ***Akademische und psychosoziale Entwicklung in nichtgymnasialen Schulformen: Ergebnisse und methodische Herausforderungen der TRAIN-Studie***

Chair(s): Kathrin Jonkmann (Universität Tübingen), Michael Becker (DIPF)

Diskutant/in: Martin Brunner (FU Berlin, Institut für Schulqualität Berlin und Brandenburg ISQ)

Wie stark hängt die Entwicklung von Kindern und Jugendlichen davon ab, in welchem Bundesland sie zur Schule gehen, auf welche Schulform sie nach der Grundschule wechseln, wer ihre Klassenkameraden sind und welche Lehrkräfte sie unterrichten? Diesen Fragen, so oder ähnlich von Eltern und Politikern hunderttausendfach formuliert, gehen Bildungsforscher seit mehreren Jahrzehnten nach. Dennoch ist der gesicherte Kenntnisstand zu diesen Kernthemen der empirischen Bildungsforschung noch immer teils fragmentarisch, und es ist nach wie vor nicht hinreichend geklärt, wie sich die Mehrgliedrigkeit des Schulsystems auf die akademische und psychosoziale Entwicklung der Schülerinnen und Schüler auswirkt. Die TRAIN-Studie „Tradition und Innovation im Sekundarschulwesen“ versucht einen Beitrag zur Schließung dieser Forschungslücken zu leisten. Ein wesentlicher Anstoß für die TRAIN-Studie war die seit einigen Jahren beklagte Situation an Hauptschulen als eine Art „Restschule“ mit einer möglicherweise kritischen leistungsbezogenen und sozialen Komposition, was in einigen Bundesländern zu ihrer Umstrukturierung oder sogar Abschaffung geführt hat.

In der Konzeption, Durchführung und Auswertung vereint die TRAIN-Studie psychologische, erziehungswissenschaftliche und soziologische Perspektiven. Darüber hinaus profitiert sie von der Unterstützung durch Fachdidaktiker und Lehrplanexperten der teilnehmenden Länder. TRAIN begleitete über vier Messzeitpunkte von der Klasse 5 bis zur Klasse 8 die Entwicklung von über 3.800 Schülerinnen und Schülern und ihren Lehrkräften in Haupt- und Realschulen in Baden-Württemberg und Mittelschulen als einzige nichtgymnasiale Schulform in Sachsen. Ziel des sowohl inhaltlich als auch methodisch orientierten Symposiums ist es, die Studie, ihre zentralen Befunde sowie die Vielfalt der Fragestellungen, mit denen sich Bildungsforscher unterschiedlicher disziplinärer Schwerpunkte in ihrem Rahmen beschäftigen, vor- und zur Diskussion zu stellen.

Nach einem kurzen Überblick über das Design von TRAIN untersucht Beitrag 1 Schereneffekte der Leistungsentwicklung: Welche Lernverläufe lassen sich in den Domänen Mathematik und Deutsch im Vergleich der Bundesländer und Schulformen feststellen? Beitrag 2 widmet sich der Analyse von Kompositionseffekten: Erzielen Klassen, die eine günstigere Zusammensetzung hinsichtlich der Leistungsfähigkeit, des sozialen und des kulturellen Hintergrundes der Kinder aufweisen, höhere Lernzuwächse? In Beitrag 3 wird die Bedeutung der Bildungsgänge für die Entwicklung des Fachinteresses und der Anstrengungsbereitschaft der Schülerinnen und Schüler untersucht: Wirken auch hier Schulformen als differentielle Entwicklungsmilieus? In Beitrag 4 wird schließlich eine der methodischen Herausforderung großangelegter Schulleistungsstudien thematisiert: Wie wirkt sich die Reihenfolge der Itemdarbietung auf die Schätzung der Leistungen aus? Das Symposium schließt mit einer Diskussion aus sowohl bildungswissenschaftlicher als auch bildungspraktischer Perspektive durch Martin Brunner, Professor für Evaluation und Qualitätssicherung im Bildungswesen und Geschäftsführer des Instituts für Schulqualität der Länder Berlin und Brandenburg (ISQ).

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G168



Michael Becker (DIPF), Marko Neumann (DIPF), Kathrin Jonkmann (Eberhard Karls Universität Tübingen), Norman Rose (Eberhard Karls Universität Tübingen), Hanna Dumont (DIPF), Karin Berendes (Eberhard Karls Universität Tübingen), Gabriel Nagy (IPN Kiel), Ulrich Trautwein (Eberhard Karls Universität Tübingen)

## ***Die Entwicklung von Mathematik- und Deutschleistungen an den nichtgymnasialen Bildungsgängen Baden-Württembergs und Sachsens***

In der Literatur werden die Schulformen des gegliederten deutschen Sekundarschulsystems häufig als „differenzielle Lern- und Entwicklungsmilieus“ (vgl. Baumert, Stanat, & Waterman, 2006; Maaz, Baumert, & Trautwein, 2009) bezeichnet, die Kindern bei gleichen individuellen Lernvoraussetzungen unterschiedliche Entwicklungschancen bieten. Obwohl dies gerade in der jüngeren Forschung vielfach diskutiert wird, gibt es bislang immer noch wenige Arbeiten, die diese Vermutung direkt adressieren (können), weshalb die Befundlage bislang heterogen und mitunter widersprüchlich erscheint (vgl. Becker, 2009). Dies bezieht sich sowohl auf die deskriptive Ebene (z.B. Beschreibung der Entwicklungsverläufe zwischen den Schulformen) als auch auf die Frage, ob von den potentiell unterschiedlichen Entwicklungsmilieus eine direkte Wirkung ausgeht.

Hier setzt der vorliegende Beitrag an. Gegenstand ist die vergleichende Untersuchung der Leistungsentwicklung in Mathematik und im Leseverständnis in den nichtgymnasialen Schulformen in der Sekundarstufe I (Klassenstufe 5 bis 8) in Baden-Württemberg und Sachsen. In Sachsen besteht neben dem Gymnasium nur die Mittelschule als weitere Schulform, während sich der nichtgymnasiale Bereich in Baden-Württemberg im Zeitraum der Untersuchung im Wesentlichen in Haupt- und Realschule gliederte. Entsprechend wird neben der Systemperspektive (Bundeslandvergleich) auch eine schulformbezogene Perspektive eingenommen, in der die Leistungsentwicklungen aller drei Schulformen näher betrachtet werden.

Grundlage der Analysen bildet die TRAIN-Studie, in der Schülerinnen und Schüler über vier Messzeitpunkte vom Anfang der Sekundarstufe I in Klasse 5 bis zur Klasse 8 untersucht wurden. Die nachfolgenden Analysen basieren auf der Schülerstichprobe von N = 3880 Schülerinnen und Schülern aus 22 Mittelschulen, 24 Realschulen und 59 Hauptschulen. Die Leistungstests bestanden im Wesentlichen aus einem auf dem Literacy-Konzept basierenden Aufgabenpool der Bildungsstandards und wurden mit einem probabilistischen Testmodell über die vier Erhebungen auf einer einheitlichen Metrik skaliert.

Im ersten Analyseschritt wurden deskriptiv die Entwicklungsverläufe der Schülerleistungen betrachtet. Im zweiten Schritt wurde geprüft, inwiefern die Unterschiede in den deskriptiven Entwicklungsverläufen zwischen den Bundesländern und Schulformen statistisch signifikant waren und bestehen blieben, wenn in den Analysen zusätzlich die individuellen Eingangsvoraussetzungen der Schülerinnen und Schüler berücksichtigt wurden (z.B. Geschlecht, kognitive Grundfähigkeiten, familiärer Hintergrund).

In beiden Leistungsdomänen fanden sich in Sachsen und Baden-Württemberg ähnliche Ausgangswerte zu Beginn der Sekundarstufe und günstigere Leistungswerte zum letzten Messzeitpunkt (Beginn 8. Klasse). In Mathematik ließ sich der Unterschied in der Leistungsveränderung zwischen den Bundesländern signifikanzstatistisch absichern, sowohl auf Grundlage querschnittlicher Mittelwertvergleiche als auch im wellenübergreifenden Wachstumskurvenmodell. Im bundeslandübergreifenden Schulformvergleich war der Unterschied im Lernzuwachs zwischen der Mittelschule und der Realschule statistisch signifikant. Unterschiede im Zuwachs zwischen Mittel- und Hauptschule fielen hingegen nicht statistisch signifikant aus. Die spezifischen Effekte blieben erhalten, wenn Unterschiede in den individuellen Eingangsvoraussetzungen der Schülerinnen und Schüler berücksichtigt wurden. Im Leseverstehen ließen sich Mittelwertunterschiede zwischen den Bundesländern lediglich querschnittlich in der 8. Klasse belegen. Die Entwicklung der Veränderung, unter Berücksichtigung der Gesamtentwicklung über die drei Schuljahre, unterschied sich jedoch nicht statistisch bedeutsam zwischen den Bundesländern. Für die bundeslandübergreifenden Schulformunterschiede zeigten sich statistisch bedeutsame Unterschiede im intercept, aber keine Unterschiede in den linearen slopes. Das Ergebnismuster blieb auch unter Kontrolle individueller und sozialer Lernausgangsvoraussetzungen und Hintergrundmerkmalen erhalten. Zusammenfassend weisen die Auswertungen darauf hin, dass sich günstigere Entwicklungen für Sachsen andeuteten, und dies in tendenziell günstigeren Entwicklungen in der Mittelschule begründet liegt. Jedoch ist dieses Bild nicht über beide Domänen einheitlich zu replizieren, sondern konsistenter für die Mathematik.

Baumert, J., Stanat, P., & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat, & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differentielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, M. (2009). Kognitive Leistungsentwicklung in differenziellen Lernumwelten: Effekte des gegliederten Sekundarschulsystems in Deutschland. Berlin: Max-Planck-Institut für Bildungsforschung.

Maaz, K., Baumert, J., & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In J. Baumert, K. Maaz, & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft: Sonderheft No. 12) (S. 11-46). Wiesbaden: VS Verlag für Sozialwissenschaften.



Kathrin Jonkmann (Eberhard Karls Universität Tübingen), Hanna Dumont (DIPF),  
Norman Rose (Eberhard Karls Universität Tübingen)

## **Effekte der Klassenkomposition auf die Leistungsentwicklung in Hauptschulen, Realschulen und Mittelschulen**

Wenn Schülerinnen und Schüler bei gleichen individuellen Lernvoraussetzungen in Abhängigkeit von der leistungsmäßigen, sozioökonomischen, kulturellen oder lernbiographischen Zusammensetzung ihrer Klasse oder Schule unterschiedlich hohe Lernzuwächse aufweisen, spricht man von Kompositionseffekten der Schülerschaft. Kompositionseffekte gilt es sowohl zu analysieren aber auch zu kontrollieren, um Schulen und Schulformen als potentielle differentielle Lern- und Entwicklungsmilieus (Baumert, Stanat, & Watermann, 2006) zu identifizieren (siehe Dumont, Neumann, Maaz, & Trautwein, 2013, für eine ausführliche Darstellung des Forschungsstandes). Betrachtet man Kompositionseffekte getrennt für unterschiedliche Schulformen, dann geben diese Analysen Auskunft auf die Frage, ob es einen Unterschied macht, welche Einzelschule bzw. Klasse ein Kind besucht. Wird beispielsweise in manchen Hauptschulklassen mehr gelernt als in anderen? Welche Merkmale der Klassenzusammensetzung wirken sich besonders günstig oder ungünstig auf die Leistungsentwicklung aus? Und welche Schulformen sind besonders sensitiv für Effekte ungünstiger Klassenkomposition? Auf diesen schulformspezifischen Effekten liegt der Schwerpunkt des vorliegenden Beitrags. Darüber hinaus ist die Kontrolle von Kompositionseffekten bedeutsam, um potentielle Schulformeffekte zu isolieren, die beispielsweise auf systematisch variierende Curricula, Ausbildungswege der Lehrkräfte und pädagogische Traditionen zurückgehen könnten (z. B. Kunter et al., 2005, Terhart, 2003). Bisherige Studien zu Kompositionseffekten, die beispielsweise auf Basis der querschnittlichen PISA-Daten (Baumert & Schümer, 2000) oder der Längsschnittstudie KESS (Gröhlich, Guill, Scharenberg, & Bos, 2010; Stanat, Schwippert, & Gröhlich, 2010) durchgeführt wurden, zeigen, dass in Klassen, in denen ein höheres mittleres Leistungsniveau vorliegt und die im Mittel einen günstigeren sozioökonomischen Status aufweisen, höhere Leistungen erzielt werden. Der Migrantenanteil scheint hingegen keinen eigenständigen, über das Fähigkeitsniveau und den mittleren sozioökonomischen Status hinausgehenden Effekt zu haben. Desweiteren deutet sich an, dass Hauptschulen besonders sensitiv auf ungünstige Zusammensetzungen der Schülerschaft reagieren (Baumert et al., 2006). Ziel dieses Beitrags ist es, den empirischen Kenntnisstand zu schulformspezifischen Kompositionseffekten zu erweitern, indem die Leistungsentwicklung über drei Schuljahre in drei Fächern in Hauptschulen und Realschulen in Baden-Württemberg und sächsischen Mittelschulen analysiert wird. Die Forschungsfrage lautet: Wie wirkt sich unter Kontrolle zentraler Indikatoren der individuellen Lernvoraussetzungen die mittlere Ausgangsleistung, der mittlere sozioökonomische Status und der Migrantenanteil der Klasse zu Beginn der Klassenstufe 5 auf das Leistungsniveau in Deutsch, Mathematik und Englisch in der Klassenstufe 8 innerhalb der einzelnen Schulformen aus? Die Analysen wurden auf Basis des TRAIN-Datensatzes in 61 Hauptschulklassen, 27 Realschulklassen und 48 Mittelschulklassen mit einer Gesamtstichprobengröße von N = 3880 Schülerinnen und Schülern durchgeführt. In Mehrgruppen-Mehrebenenanalysen wurden auf der Individualebene Geschlecht und Migrationshintergrund sowie Unterschiede in der Vorleistung, in den kognitiven Grundfähigkeiten, im sozioökonomischen Status (höchster ISEI in der Familie), im Bücherbesitz und im höchsten Bildungsabschluss der Eltern kontrolliert. Auf der Klassenebene wurden die Kompositionsmerkmale zur Vorhersage des Klassenmittelwertes aufgenommen. Die Befunde lassen sich wie folgt zusammenfassen. Die mittlere Leistung in Klassenstufe 5 war in den Mittelschulen für Leseverstehen, Mathematik und Englisch ein statistisch signifikanter Prädiktor der Lernzuwächse. In den Hauptschulen war sie nur für Mathematik bedeutsam, in den Realschulen war der Effekt für keine der Domänen statistisch signifikant. Der mittlere sozioökonomische Status war hingegen nur in den Realschulen ein signifikanter positiver Prädiktor der Leistungszuwächse in Mathematik. In allen Schulformen zeigten sich bei Einzelbetrachtung der Kompositionsmerkmale signifikante negative Effekte des Migrantenanteils auf die Leistungszuwächse in Mathematik. Darüber hinaus zeigten sich diese für Lesen in den Haupt- und Mittelschulen und für Englisch in den Realschulen. Diese Effekte waren jedoch bei gleichzeitiger Berücksichtigung der Vorleistung nicht länger statistisch bedeutsam. In weiteren Analysen sollen mögliche Schulformeffekte unter Kontrolle der Kompositionseffekte beleuchtet werden.

Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.

Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183.

Gröhlich, C., Guill, K., Scharenberg, K. & Bos, W. (2010). Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 - Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (S. 100-106). Münster: Waxmann.

Kunter, M., Brunner, M., Baumert, J., Klusmann, U., Krauss, S. & Blum, W. (2005). Der Mathematikunterricht der PISA-Schülerinnen und -Schüler. *Schulformunterschiede in der Unterrichtsqualität*. *Zeitschrift für Erziehungswissenschaft*, 8, 502-520.

Terhart, E. (2003). Die Lehrerbildung. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 787-810). Reinbek: Rowohlt.

Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb: Längsschnittliche Überprüfung eines umstrittenen Effekts. *Zeitschrift für Pädagogik*, 55. Beiheft, 147-164.

## **Interessiert und fleißig? Entwicklung der Motivation in nichtgymnasialen Schulformen**

Schulleistungsstudien weisen darauf hin, dass der Leistungsstand und die Leistungsentwicklung in Hauptschulen nicht zufriedenstellend sind (Becker, Lüdtke, Trautwein, & Baumert, 2006; Prenzel, et al., 2008). Zudem wird oftmals gemutmaßt, dass an vielen Hauptschulen ein breites Desinteresse und wenig Bereitschaft zur Anstrengung vorherrsche, was wiederum den Lernerfolg gefährde (vgl. Knigge, 2009; Schümer, 2004). Betrachtet man jedoch die wissenschaftliche Befundlage lassen sich diese Aussagen aufgrund von fehlenden empirischen Belegen weder belegen noch widerlegen. Insgesamt gibt es nur wenige Studien, die sich mit der Entwicklung von Interesse und Anstrengungsbereitschaft in Bezug auf unterschiedliche Lernmilieus beschäftigt haben. Während diese Studien den generellen Abfall in den Interessen und der Anstrengungsbereitschaft im Laufe der Schulzeit bestätigen (Marks, 2000; Trautwein, Lüdtke, Kastens, & Köller, 2006; Watt, 2004), zeigten sich hierbei Hinweise auf leicht unterschiedliche Entwicklungsverläufe in unterschiedlichen Lernmilieus. So scheint das Interesse einen leicht positiveren Entwicklungsverlauf in leistungsschwachen Lerngruppen als in leistungsstarken Gruppen zu nehmen, wohingegen die Anstrengungsbereitschaft in leistungsstärkeren Lerngruppen tendenziell höher ausgeprägt ist als in leistungsschwachen Lerngruppen (Carbonaro, 2005; Frenzel, Goetz, Pekrun, & Watt, 2010). Es fehlt jedoch an Studien, die zwischen schulischem Interesse und schulischer Anstrengungsbereitschaft unterscheiden sowie diese beiden Konstrukte in mehreren Fächern und mit einem längsschnittlichen Design untersuchen und zudem neben der Hauptschule noch andere Schulformen als Vergleichsgruppen berücksichtigen. Im vorliegenden Beitrag wurde versucht diesem Defizit entgegenzuwirken.

In der repräsentativen TRAIN-Stichprobe von Schülerinnen und Schülern der Haupt- und Realschule in Baden-Württemberg sowie der Mittelschule in Sachsen wurden Schulformunterschiede in den Interessen und der Anstrengungsbereitschaft in drei Schulfächern (Mathematik, Deutsch und Englisch) sowie deren Entwicklungsverläufe zwischen der fünften und der achten Jahrgangsstufe untersucht. Hierbei wurde zum einen beleuchtet, inwieweit sich die Schülerinnen und Schüler der untersuchten Schulformen in ihrem fachspezifischen Interesse und der fachspezifischen Anstrengungsbereitschaft in der Jahrgangsstufe 5 sowie in der Jahrgangsstufe 8 voneinander unterscheiden. Zum anderen wurden mit Hilfe von Wachstumskurvenmodellen die Veränderungen im fachspezifischen Interesse und der fachspezifischen Anstrengungsbereitschaft über die drei Schuljahre untersucht.

Es zeigte sich, dass zu Beginn der Sekundarstufe die Schülerinnen und Schüler der Hauptschule das höchste Interesse berichteten, gefolgt von Schülerinnen und Schüler der Realschule und der Mittelschule. Die Schülerinnen und Schüler in der Mittelschule berichteten zudem auch die am niedrigsten ausgeprägte Anstrengungsbereitschaft. Die Anstrengungsbereitschaft fiel am höchsten für die Realschule aus. Bei Betrachtung der Veränderung über die vier Messzeitpunkte zeigte sich in allen untersuchten Schulformen eine bedeutsame Abnahme von Interesse und Anstrengungsbereitschaft. Es traten zudem sowohl für die fachspezifischen Interessen als auch die fachspezifische Anstrengungsbereitschaft geringere jährliche Abnahmen in der Hauptschule als in der Realschule und Mittelschule auf. Die vorliegenden Befunde zeigen somit auf, dass die Berücksichtigung der unterschiedlichen Lernmilieus bei der Betrachtung der Motivationslage von Schülerinnen und Schülern wichtig ist. Es bedarf nun weiterer Forschung, die die für die differentielle Entwicklung relevanten Wirkmechanismen näher beleuchtet.

- Becker, M., Lüdtke, O., Trautwein, U., & Baumert, J. (2006). Leistungszuwachs in Mathematik: Evidenz für einen Schereneffekt im mehrgliedrigem Schulsystem? *Zeitschrift für Pädagogische Psychologie/ German Journal of Educational Psychology*, 20, 233-242.
- Carbonaro, W. (2005). Tracking, Students' Effort, and Academic Achievement. *Sociology of Education*, 78, 27-49. doi: 10.1177/003804070507800102
- Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, 20, 507-537. doi: 10.1111/j.1532-7795.2010.00645.x
- Knigge, M. (2009). *Hauptschüler als Bildungsverlierer*. Münster: Waxmann.
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37, 153-184. doi: 10.3102/00028312037001153
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., et al. (Eds.). (2008). *PISA 2006 in Deutschland - Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Münster: Waxmann.
- Schümer, G. (2004). Zur doppelten Benachteiligung von Schülern aus unterprivilegierten Gesellschaftsschichten im deutschen Schulwesen. In G. Schümer, K. J. Tillmann & M. Weiss (Eds.), *Die Institution Schule und die Lebenswelt der Schüler: Vertiefende Analysen der PISA-2000-Daten zum Kontext von Schülerleistungen* (pp. 73-114). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Trautwein, U., Lüdtke, O., Kastens, C., & Köller, O. (2006). Effort on homework in grades 5-9: Development, motivational antecedents, and the association with effort on classwork. *Child Development*, 77, 1094-1111. doi: 10.1111/j.1467-8624.2006.00921.x
- Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development*, 75, 1556-1574. doi: 10.1111/j.1467-8624.2004.00757.x



## **Modellierung von Itempositionseffekten mit mehrdimensionalen IRT Modellen**

In bildungswissenschaftlichen Large Scale-Assessments wird zumeist eine große Zahl von Testaufgaben zur Erfassung von Schülerleistungen verwendet, um die Leistungsdomäne möglichst umfassend im Test zu repräsentieren. Aus Gründen der Effizienz und Zumutbarkeit wird den einzelnen Teilnehmerinnen und Teilnehmern einer Studie jeweils nur eine Auswahl der Testitems, z. B. in Form von booklets, zur Bearbeitung vorgelegt. Das Testdesign beschreibt die Zuweisung der Items und Tests zu den Schülerinnen und Schülern. In der TRAIN Studie wurde ein rotiertes Matrixdesign verwendet, das Erinnerungseffekte aufgrund der längsschnittlichen Untersuchungsanlage ausschließt. Der Nachteil von Bookletdesigns liegt darin, dass differential item functioning (DIF) über die verschiedenen Testbooklets auftreten kann. So deuten empirische Befunde darauf hin, dass Items geringere Lösungshäufigkeiten aufweisen, wenn sie am Ende statt zu Beginn eines Tests präsentiert werden. Dieser häufig replizierte Positionseffekt wird oft als Ermüdungseffekt interpretiert. Alternativ sind auch Übungseffekte denkbar, die sich in einer positiven Korrelation von Lösungshäufigkeit und Itemposition im Test zeigen würde. Man kann erwarten, dass Ermüdung oder Übung interindividuelle Varianz aufweisen, und somit unterschiedliche Effekte hinsichtlich der Testleistung haben. Je nach Test und Fragestellung stellen solche Itempositionseffekte somit eine konstruktirrelevante Varianzquelle und folglich eine mögliche Validitätsgefährdung dar. Geeignete IRT-basierte Verfahren erlauben die Analyse und Kontrolle von Positionseffekten in quer- und längsschnittlichen Assessments. Vor diesem Hintergrund adressiert die vorgestellte Studie drei Fragestellungen: Erstens, wie können Positionseffekte analysiert, getestet und modelliert werden? Zweitens, gibt es Korrelate für interindividuelle Unterschiede in Itempositionseffekten? Und drittens, welchen Einfluss hat die Kontrolle von Itempositionseffekten auf Zusammenhänge zwischen den resultierenden Personenparameterschätzern und Drittvariablen? Es wurden die TRAIN-Daten des Leseverständnistests zum ersten Messzeitpunkt in Klassenstufe 5 verwendet. Die Stichprobengröße beläuft sich auf N = 2774 Schülerinnen und Schüler aus Haupt-, Real- und Mittelschulen. Der Test bestand aus 51 Items, die in neun verschiedenen Testheften administriert wurden. Zur Analyse der Itempositionseffekte wurden vier verschiedene IRT Modelle berechnet: (a) das herkömmliche 2PL-Modell als Referenzmodell, (b) das 2PL-Modell mit festen Itempositionseffekten hinsichtlich des Schwierigkeitsparameters, (c) das 2PL-Modell mit zufälligen Positionseffekten für die Itemschwierigkeit und (d) das 2PL-Modell mit zufälligen Itempositionseffekten für die Itemschwierigkeiten und festen Effekten für die Itemdiskrimination. Die Modelle stellen eine Erweiterung mehrdimensionaler IRT-Modelle dar, in denen die zufälligen Effekte als latente Variablen unter Verwendung virtueller Items definiert wurden. Im Weiteren wurden Kovariaten wie allgemeine kognitive Fähigkeiten, basale Lesefertigkeiten, Aufmerksamkeit und Konzentration, verbales Selbstkonzept und habituelle Lesemotivation in die verschiedenen Modelle mit einbezogen, um deren Zusammenhang mit der latenten Variable (Leseverständnis) vergleichend zu analysieren. Die Ergebnisse der Analysen zeigen statistisch bedeutsame Itempositionseffekte hinsichtlich der Itemschwierigkeit und der Itemdiskrimination. Der Vergleich der verschiedenen IRT Modelle anhand des Bayesian information criterion (BIC) ergab den besten Modellfit für das vierte Modell mit interindividuell variierenden Positionseffekten hinsichtlich der Itemschwierigkeiten und einem negativen Positionseffekt für die Itemdiskriminationen. Das Befundmuster spricht für eine Interpretation als Ermüdungseffekte. Die Zusammenhänge zwischen Kovariaten und latenter Fähigkeitsvariable variieren über die verschiedenen Modelle. Die Ergebnisse zeigen somit, dass Itempositionseffekte auch im Sinne der Testvalidität relevant sind.

## ***Determinanten motivationaler Dynamik von Schülerinnen und Schülern im Kontext schulischer Bildung***

Chair(s): Miriam Marleen Gebauer (TU Dortmund), Nele McElvany (TU Dortmund)

Diskutant/in: Jens Möller (Christian-Albrechts-Universität zu Kiel)

Fokus des Symposiums sind Effekte und Determinanten der motivationalen Dynamik. Sowohl selbstbezogene Kognitionen als auch motivationale Orientierungen sind für das Erbringen schulischer Leistung und Erfolg von entscheidender Bedeutung (Krapp & Ryan, 2002). Dabei werden motivationale und selbstbezogene Aspekte von individuellen und sozialen Merkmalen der Schülerinnen und Schüler mitbestimmt. Wichtige Bedingungen der motivationalen Orientierung und selbstbezogener Kognitionen von Schülerinnen und Schüler sind unter anderem die Lernzielorientierung, Geschlecht, Migrationshintergrund und familiäre Merkmale der Jugendlichen. Ziel des Symposiums ist es mit Vorträgen aus verschiedenen Perspektiven und einer anschließenden Diskussion einen Beitrag dazu zu liefern, genauere Erkenntnisse über die Zusammenhänge zwischen den genannten Aspekten zu erhalten. Mit vier Beiträgen aus vier unterschiedlichen Arbeitsgruppen werden zentrale Aspekte der motivationalen Dynamik betrachtet und diskutiert. Im Folgenden werden die Referenten und Titel der Beiträge genannt: Der erste Beitrag von Weidinger, Spinath und Steinmayr mit dem Titel: Moderatoren des Zusammenhangs zwischen Fähigkeitsselbstkonzept und Intrinsischer Motivation: Welche Rolle spielen Zielorientierungen? Der zweite Beitrag des Symposiums von Häfner, Flunger, Dicke, Gaspard, Schreier, Nagengast und Trautwein trägt den Titel: Der familiäre Hintergrund von Schülerinnen und Schülern: Zusammenhänge mit akademischer Motivation und Leistung. Der nächste Beitrag von Schöber, Köller, McElvany, Ferdinand, Gebauer, Bos & Huelmann mit dem Titel: Reziproke Zusammenhänge zwischen Selbstwirksamkeitsüberzeugungen und Leistungen im Lesen und in Mathematik in Gruppen mit und ohne Migrationshintergrund. Den letzte Vortrag werden Heyder und Kessels halten und das Symposium um den Blickwinkel der geschlechtsspezifischen Bedeutsamkeit erweitern. Der Titel des Beitrag lautet „Stören oder mitarbeiten? Zur differenziellen Funktionalität schulischen (Des-) Engagements bei männlichen und weiblichen Jugendlichen“. Die anschließende Diskussion von Herrn Prof. Dr. Möller wird die Vorträge rahmen, konstruktive Kritiken geben und mögliche Ergänzungen bieten. Somit umfasst das Symposium mehrere Perspektiven auf die motivationale Dynamik von Schülerinnen und Schüler. Es werden individuelle aber auch familiäre und soziale Aspekte berücksichtigt und den Effekt für die Aspekte der motivationalen Dynamik untersucht.

## **Moderatoren des Zusammenhangs zwischen Fähigkeitsselbstkonzept und Intrinsischer Motivation: Welche Rolle spielen Zielorientierungen?**

In bekannten Motivationsmodellen wird explizit oder implizit davon ausgegangen, dass die Wahrnehmung, die eine Person von ihren eigenen Fähigkeiten hat (Fähigkeitsselbstkonzept, FSK), ihre Intrinsische Motivation (IM) beeinflusst (Harter, 1978; Ryan & Deci, 2000; White, 1959; Wigfield & Eccles, 2000). In mehreren längsschnittlichen Studien wurde diese Annahme für Schüler/innen der ersten bis zwölften Klasse bezogen auf unterschiedliche Schulfächer empirisch untersucht, wobei allenfalls schwache Hinweise dafür gefunden wurden, dass sich das FSK zu einem früheren Messzeitpunkt auf die IM zu einem späteren Zeitpunkt auswirkt (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Nurmi & Aunola, 2005; Skaalvik & Valas, 1999; Spinath & Spinath, 2005; Spinath & Steinmayr, 2008, 2012). Uns sind darüber hinaus drei Studien bekannt, die den postulierten Kausalzusammenhang im Rahmen der Feedbackforschung experimentell untersuchten (Sansone, 1986; Vallerand & Reid, 1984, 1988). Sie stellten fest, dass der Effekt von Leistungsfeedback auf die IM durch die Kompetenzwahrnehmung der Personen mediiert wurde. In diesen Experimenten wurde allerdings nicht für den Einfluss des FSKs und der IM vor Erhalt des Feedbacks sowie für den Einfluss der tatsächlichen Aufgabenleistung der Personen auf deren IM kontrolliert, was zu einer Verzerrung der Ergebnisse geführt haben könnte. Hinsichtlich interindividueller Unterschiede im Abfall der IM scheint es darüber hinaus angezeigt, den Einfluss des FSK auf die IM unter Berücksichtigung weiterer motivationaler Variablen wie der Zielorientierung experimentell zu erforschen (vgl. Dweck, 1986).

In einem computerbasierten Experiment wurde deshalb untersucht, ob sich die experimentelle Manipulation des FSKs für numerische Intelligenztestaufgaben von Studierenden einer deutschen Universität auf deren IM für ebendiese Aufgaben auswirkt (N = 90, M = 22.74 Jahre, SD = 2.82). Darüber hinaus wurde der theoretischen Annahme nachgegangen, ob der Zusammenhang von FSK und IM in Abhängigkeit der Zielorientierung der Person variiert. Unabhängig von ihrer tatsächlichen Leistung in 30 numerischen Intelligenztestaufgaben erhielten die Versuchspersonen zufällig entweder einer negative oder eine positive normativ-basierte Leistungsrückmeldung. Vor und nach dem Erhalt der Rückmeldung wurden sowohl das aufgabenspezifische FSK als auch die aufgabenspezifische IM über Selbstberichtsfragebögen erfasst. Eine Varianzanalyse mit Messwiederholung zeigte, dass – auch bei Kontrolle der tatsächlichen Aufgabenleistung – sowohl das FSK als auch die IM der Personen nur nach negativer Leistungsrückmeldung abgesunken sind. Eine Mediationsanalyse mittels Pfadmodell zeigte, dass das FSK den Einfluss der Rückmeldung auf die Veränderung der IM vollständig mediierte. Um Alternativerklärungen für den gefundenen Kausaleffekt ausschließen zu können, wurde zusätzlich für Veränderungen im momentanen Gefühlszustand kontrolliert sowie die IM für neutrale Aufgaben (i.e. Aufgaben, zu denen die Personen keine Rückmeldung erhielten) erfasst, welche sich in Abhängigkeit der Rückmeldung nicht signifikant veränderte. Betrachtete man den Zusammenhang des FSKs und der IM für numerische Aufgaben getrennt für hoch und niedrig lernzielorientierte Personen, so zeigte sich, dass die IM sich nur bei niedrig lernzielorientierten veränderte, während sie bei hoch lernzielorientierten Personen nahezu unverändert blieb. Eine Moderation durch die Annäherungsleistungszielorientierung konnte nicht beobachtet werden, während sich ein Moderationseinfluss der Vermeidungsleistungsziele zeigte: Die IM sank nach negativer Rückmeldung bei hoch Vermeidungsleistungszielorientierten stärker ab als bei niedrig Vermeidungsleistungszielorientierten. Die Ergebnisse bekräftigen die Annahme, dass sich das FSK einer Person auf deren IM auswirkt und dass dieser Einfluss von der Zielorientierung der Person abhängig ist. Die Ergebnisse werden in Bezug auf ihre theoretischen und praktischen Implikationen diskutiert.



## ***Der familiäre Hintergrund von Schülerinnen und Schülern: Zusammenhänge mit akademischer Motivation und Leistung***

Zusammenhänge zwischen dem familiären Hintergrund von Schülerinnen und Schülern und der schulischen Motivation und Leistung konnten in mehreren Studien der Bildungsforschung nachgewiesen werden (Sirin, 2005; Steinmayr, Dinger, & Spinath, 2012). Studien, die sich in diesem Kontext mit dem familiären Hintergrund befassten, verwendeten allerdings meist unterschiedliche Indikatoren für den familiären Hintergrund. In einigen Studien wurde der sozioökonomische Status als Indikator des familiären Hintergrundes herangezogen (Sirin, 2005). In anderen Studien wurden hingegen verschiedene Familiencharakteristika berücksichtigt, wie z.B. die Eltern-Kind Beziehung und elterliche Werte (Simpkins, Fredricks, & Eccles, 2012; Yan & Lin, 2005). Auch gibt es Studien, die das elterliche Verhalten als einen weiteren zentralen Aspekt des familiären Hintergrunds betrachten (Cheung & Pomerantz, 2012; Jeynes, 2007).

Bei der Erforschung der Assoziationen zwischen familiärem Hintergrund und der schulischen Motivation und Leistung von Schülerinnen und Schülern wurden jedoch selten die verschiedenen Indikatoren simultan berücksichtigt. Somit kann der differentielle Zusammenhang zwischen den unterschiedlichen Indikatoren des familiären Hintergrunds und Schülermerkmalen nicht adäquat beurteilt werden.

Die vorliegende Studie beschäftigt sich mit der Frage, wie verschiedene Indikatoren des familiären Hintergrundes und des elterlichen Verhaltens mit verschiedenen Indikatoren der Schülermotivation und –leistung zusammenhängen. Insbesondere war die adäquate Schätzung der Zusammenhänge unter Auspartialisierung der anderen Indikatoren des familiären Hintergrundes von Interesse: Welche Zusammenhänge zwischen dem sozioökonomischen Status, den verschiedenen Familiencharakteristika und dem elterlichen Verhalten mit der Schülermotivation und –leistung lassen sich unter gleichzeitiger Berücksichtigung der anderen Indikatoren finden? Aufbauend auf Arbeiten von Baumert, Watermann und Schümer (2003) und Hoover-Dempsey und Sandler (1995) konzeptualisierten wir vier Dimensionen des familiären Hintergrundes. Neben dem sozioökonomischen Status wurden soziale (Eltern-Kind Beziehung) und motivationale Familiencharakteristika (elterliches Interesse, Werte und Selbstkonzept) berücksichtigt, sowie der Kontext der Familie (subjektive Zeit und Energie der Eltern). Darüber hinaus wurde das elterliche Verhalten multidimensional erfasst (Eltern-Kind Kommunikation, akademische Einmischung und unterstützendes Verhalten). Außerdem erfassten wir verschiedene Indikatoren der Schülermotivation (Selbstkonzept, Anstrengungsbereitschaft und Interesse in Mathematik), sowie der Mathematikleistung (Mathematiknote und Leistung in einem Mathematiktest).

Daten von 1573 Schülerinnen und Schülern, sowie ihren Eltern wurden mithilfe von Eltern- und Schülerfragebögen erhoben. Die Schülerinnen und Schüler stammten aus 82 Klassen der 9. Jahrgangsstufen von 25 Gymnasien. Zur Beantwortung der Fragestellung wurden verschiedene Strukturgleichungsmodelle in Mplus berechnet, in denen die Zusammenhänge zwischen den verschiedenen Indikatoren des familiären Hintergrundes und den verschiedenen Indikatoren der Schülermotivation und –leistung schrittweise untersucht wurden. Als Kontrollvariablen wurden das Geschlecht, die Intelligenz, sowie die Mathematiknote am Ende der 8. Klasse berücksichtigt. In einem ersten Schritt wurden die Zusammenhänge der Variablen im Querschnitt analysiert. In einem nächsten Schritt wurden die Zusammenhänge im Längsschnitt analysiert, in dem die Indikatoren der Schülermotivation und –leistung vier Monate später vorhergesagt wurden, während das Niveau zum ersten Testzeitpunkt kontrolliert wurde.

Mithilfe der verschiedenen Strukturgleichungsmodelle konnte gezeigt werden, dass die sozialen sowie die motivationalen Familiencharakteristika, unter Auspartialisierung der anderen Indikatoren des familiären Hintergrundes, positiv mit den verschiedenen Indikatoren Schülermotivation und –leistung zusammenhängen. Während das familiäre Interesse positiv mit dem Selbstkonzept und dem Interesse der Schülerinnen und Schüler in Mathematik sowie der Leistung in einem Mathematiktest zusammenhing, zeigten sich positive Zusammenhänge der Eltern-Kind-Beziehung mit der Anstrengungsbereitschaft der Schülerinnen und Schüler. Des Weiteren fanden wir negative Zusammenhänge zwischen der akademischen Einmischung der Eltern und dem Selbstkonzept der Schülerinnen und Schüler in Mathematik, sowie der Leistung im Mathematiktest, während sich keine Zusammenhänge der Eltern-Kind Kommunikation, sowie des elterlichen unterstützenden Verhaltens mit der Schülermotivation und Leistung finden ließen. Längsschnittlich betrachtet fanden sich kaum Zusammenhänge zwischen den Indikatoren des elterlichen Hintergrundes und der Schülermotivation und –leistung. Das komplexe Ergebnismuster zeigt, dass verschiedene Familiencharakteristika und elterliches Verhalten neben dem sozioökonomischen Status wichtig für die Schülermotivation sind. Zusätzlich fanden wir differentielle Zusammenhänge abhängig von dem gewählten Indikator der Schülermotivation und –leistung.



Christian Schöber (IPN Kiel), Olaf Köller (IPN Kiel), Nele McElvany (TU Dortmund), Hanna D. Ferdinand (TU Dortmund), Miriam Marleen Gebauer (TU Dortmund), Wilfried Bos (TU Dortmund), Thorben Huelmann (TU Dortmund)

### ***Reziproke Zusammenhänge zwischen Selbstwirksamkeitsüberzeugungen und Leistungen im Lesen und in Mathematik in Gruppen mit und ohne Migrationshintergrund***

Im Bereich der Forschung zu schulischen Selbstkonzepten hat sich theoretisch und empirisch die Überzeugung durchgesetzt, dass reziproke Zusammenhänge zwischen fachspezifischen Selbstkonzepten und Fachleistungen bestehen (Nebeneinander von Self-Enhancement Model und Skill-Development Model). Inwieweit sich ein reziprokes Kausalmodell auf fachspezifische Selbstwirksamkeitsüberzeugungen übertragen lässt, ist dagegen kaum untersucht, obwohl theoretische Arbeiten dies nahelegen (im Überblick Köller & Möller, 2010). Weiterhin ist bislang vergleichsweise wenig untersucht worden, ob reziproke Zusammenhänge zwischen Selbstwirksamkeitsüberzeugungen (SWÜ) und Leistungen subgruppeninvariant sind.

Um hier weitere Klärung herbeizuführen, wurde in einer Messwiederholungsstudie der Zusammenhang von fachlichen Leistungen und fachspezifischen SWÜ von Schulkindern (N = 1598; 49.4% weiblich; N = 1404 im Längsschnitt) im 7. Jahrgang an unterschiedlichen Schulen in Deutschland analysiert.

Für beide Leistungsdomänen wurde erwartet, dass SWÜ durch frühere Leistungen und Leistungen durch frühere SWÜ vorhergesagt werden, und zwar nach Kontrolle der Stabilitäten. Zur Erfassung der Lese- und Mathematikleistungen kamen standardisierte Leistungstests zum Einsatz. Die Erhebungen fanden zum Beginn und zum Ende des Schuljahres statt. Die Items stammten u.a. aus der Lernausgangslagenuntersuchung (LAU; vgl. u.a. Lehmann et al., 2001) und wurden ergänzt durch veröffentlichte Aufgaben aus den Vergleichsarbeiten für die 8. Jahrgangsstufe des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB). Dabei wurde ein Ankeritem-Design gewählt, das die längsschnittliche Modellierung der Leistungen auf einer gemeinsamen Skala ermöglicht. Die Leistungsdaten wurden dazu mit Hilfe von IRT-Modellen (1-Parameter-Analysen in ConQuest 3.0, vgl. Adams, Wu, & Wilson, 2012, Generierung von Plausible Values (PV)) ausgewertet.

Die Reliabilitäten für die ermittelten PV lagen zu beiden Messzeitpunkten über .90. Zur Erfassung der SWÜ wurden jeweils 6 Items eingesetzt (Beispielitem Lesen: „Ich bin sicher, dass ich auch den Stoff in Unterrichtstexten, die sehr schwierig zu lesen sind, verstehen kann.“ Mathematik: „Ich bin überzeugt, dass ich in Mathematik-Hausaufgaben und Mathematik-Klassenarbeiten gute Leistungen erzielen kann.“ Vierstufiges Antwortformat von 1 „trifft nicht zu“ bis „trifft zu“). Die Reliabilitäten (Cronbachs alpha) lagen zu beiden Zeitpunkten über .85. Zur Überprüfung der Hypothesen wurden Pfadanalysen in Mplus 6.12 (Muthén & Muthén, 2010) durchgeführt. Zuvor wurden fehlende Werte mittels multipler Imputation geschätzt (25 Datensätze) und die so gewonnenen Datensätze in Mplus ausgewertet.

Insgesamt zeigen sich die erwarteten reziproken Effekte, ohne dass sich eine kausale Prädominanz von einer der beiden Variablen zeigt. Zusätzlich durchgeführte Multi-Gruppenanalysen, in denen die Modelle getrennt für Schülerinnen und Schüler aus Deutschland, der Türkei und Osteuropa (Polen und ehemalige Sowjetunion) geschätzt wurden, weisen darauf hin, dass die Zusammenhänge vom Herkunftsland moderiert werden. Die Befunde werden mit Hinweis auf Unterschiede in der kulturellen Herkunft der Jugendlichen diskutiert.

## **Stören oder mitarbeiten? Zur differenziellen Funktionalität schulischen (Des-) Engagements bei männlichen und weiblichen Jugendlichen**

Jungen erreichen bei gleichen Kompetenzen schlechtere Noten als Mädchen, was zu einem großen Teil auf Geschlechtsunterschiede in überfachlichen Kompetenzen, schulerfolgsrelevantem Verhalten und Persönlichkeitsmerkmalen zu Gunsten der Mädchen zurückgeführt werden kann. So können die bei gleichen Kompetenzen besseren Noten der Mädchen zum Teil durch ihr positiveres (Sozial-)Verhalten (DiPrete & Jennings, 2012), ihre größere Selbstdisziplin (Duckworth & Seligman, 2006) und ihre größere Verträglichkeit (Steinmayr & Spinath, 2008) erklärt werden. Auch zeigen Mädchen größere Lernbereitschaft und positiveres Lernverhalten (Lam et al., 2012; Neugebauer, 2011) sowie geringere Arbeitsvermeidung (Steinmayr & Spinath, 2008) als Jungen. Darüber hinaus stören Jungen häufiger den Unterricht als Mädchen (z.B. McClowry et al., 2013; Trautwein, Köller, & Baumert, 2004).

Die Gründe für das geringere schulische Engagement und das weniger schulangepasste Verhalten von Jungen sind empirisch nicht hinreichend untersucht. Trautwein und Kollegen (2004) konnten zwar Geschlechtsunterschiede in der Prävalenz von Problemverhalten bestätigen, fanden jedoch keine Geschlechtsunterschiede in seiner Funktionalität: Jungen und Mädchen profitierten gleichermaßen von Störverhalten hinsichtlich ihres Selbstkonzepts der sozialen Anerkennung und ihres Selbstkonzeptes der eigenen Durchsetzungsfähigkeit. Qualitative Forschung zum „laddish behaviour“ von Jungen (Jackson, 2002; 2003) lieferte interessante Erklärungsansätze für die Geschlechtsunterschiede im schulerfolgsrelevanten Verhalten: Aus Äußerungen in Interviews und Gruppendiskussionen wurde abgeleitet, dass (britische) Jungen glauben, Störverhalten und demonstrativ wenig Mitarbeit im Unterricht würden ihr Ansehen unter Gleichaltrigen steigern und insgesamt besonders männlich wirken (Jackson, 2002; 2003). Da diese Annahmen ausschließlich auf Interviews mit britischen Jungen basieren, ist bisher weder die vermutete Kausalität noch die Geschlechtsspezifität des zugrundeliegenden Mechanismus noch die Gültigkeit der Befunde für deutsche Jugendliche hinreichend nachgewiesen. Aus deutschen Vignettenexperimenten ist bislang lediglich bekannt, dass hochleistende Schülerinnen und Schüler, die sich sehr für die Schule anstrengen und als Streber tituliert werden, von ihren MitschülerInnen weniger gern gemocht werden als hochleistende Schülerinnen und Schüler, die sich weniger offensichtlich für die Schule anstrengen, wobei diese Effekte bei Jungen stärker waren als bei Mädchen (Rentzsch, Schütz, & Schröder-Abé, 2011).

Fragestellung: Unsere Studie untersucht, ob das Ansehen durchschnittlich leistungsstarker Schülerinnen und Schüler unter deutlich wahrnehmbarem schulischem Engagement leidet und von Störverhalten profitiert. Wir erwarten, dass vor allem Jungen annehmen, dass Jungen durch geringes schulisches Engagement und vermehrtes Störverhalten ihre Beliebtheit bei Gleichaltrigen steigern können.

Methode: In einem 2 (Geschlecht Targetperson) x 2 (Geschlecht Versuchsperson) –faktoriellen Vignettenexperiment wurde die Funktionalität schulischen Des-Engagement und Störverhaltens auf das Ansehen in der Peergroup untersucht. Jugendliche der 9. Jahrgangsstufe einer Integrierten Sekundarschule bearbeiteten eine Vignette, die ein Mädchen und einen Jungen beschreibt, die neu in die Klasse kommen und überlegen, wie sie sich am besten verhalten sollen, um in der neuen Klasse „gut anzukommen“. Anschließend beurteilten die Jugendlichen verschiedene Verhaltensweisen danach, ob das Mädchen/ der Junge damit ihre/seine Beliebtheit in der Klasse steigern könnte. Die zur Auswahl stehenden Verhaltensweisen waren an zwei studentischen Stichproben vorgetestet, dass sie eindeutig entweder schulisches Engagement oder Störverhalten im Unterricht signalisieren.

Ergebnisse: Die Jugendlichen glaubten, dass Jungen und Mädchen ihre Beliebtheit auf unterschiedliche Weise würden steigern können: Das Mädchen könne seine Beliebtheit dadurch steigern, dass es im Unterricht gut mitarbeitet und möglichst wenig stört, während sich für den Jungen die angenommenen Auswirkungen schulischen Engagements einerseits und Störverhalten andererseits in ihrer Wirkung auf seine Beliebtheit nicht unterschieden. Entgegen der Erwartungen zeigten sich keine Unterschiede zwischen männlichen und weiblichen Probanden.

Die Ergebnisse werden vor dem Hintergrund diskutiert, inwiefern schulisches Desengagement als besonders funktional für Jungen gelten kann.

## Untersuchung des Professionswissens von Lehramtsstudierenden und Lehrkräften in den naturwissenschaftlichen Fächern

Chair(s): Jörg Großschedl (IPN Kiel), Ute Harms (IPN Kiel), Daniela Mahler (IPN Kiel)

Diskutant/in: Josef Riese (Universität Paderborn)

Ziel des Symposiums ist die Zusammenführung und Diskussion von Studien, die sich aktuell mit dem fachbezogenen Professionswissen von Lehramtsstudierenden und im Dienst befindlichen Lehrkräften der naturwissenschaftlichen Fächer auseinandersetzen. Im Mittelpunkt des Symposiums stehen (1) die jeweilige Testkonstruktion, (2) die Rolle von Lerngelegenheiten und personenbezogenen Leistungsmerkmalen für den Erwerb fachbezogenen Professionswissens und (3) dessen Bedeutung für die Unterrichtsqualität. Die Arbeiten sind verortet in drei Großprojekten („Professionswissen in den Naturwissenschaften“ [ProwiN], „Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen“ [KiL] und „Kompetenzmodellierung und -erfassung zum Wissenschaftsverständnis über naturwissenschaftliche Arbeits- und Denkweisen bei Studierenden [Lehramt]“ [Ko-WADIS]). In den Projekten ProwiN und KiL wird das Professionswissen breit, in den Facetten Fachwissen, fachdidaktisches und pädagogisches Wissen, erfasst. Innerhalb dieser Facetten wird wiederum zwischen Inhaltsbereichen und kognitiven Prozessen bzw. Wissenstypen unterschieden. Das Projekt Ko-WADIS nimmt im Detail Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung in den Blick.

Der Beitrag aus dem Projekt ProwiN betrachtet das Professionswissen als Prädiktor der Unterrichtsqualität. Diese wird mittels Videografie erfasst und über einen Mehrebenen-analytischen Ansatz mit dem Lernerfolg von Schülerinnen und Schülern in Beziehung gesetzt. Die Beiträge aus dem KiL-Projekt widmen sich der Validierung von Tests zur Erfassung des biologie- und physikbezogenen Professionswissens. Als Validitätshinweise werden Korrelationsanalysen zu fachunabhängigen und fachbezogenen Leistungsmerkmalen der Studierenden berichtet. Zudem werden Analysen zur Überprüfung der diskriminanten Validität durchgeführt. Im Bereich des physikbezogenen Professionswissens werden diese Analysen durch Think-Aloud-Protokolle ergänzt, die zeigen, ob die entwickelten Aufgaben die intendierten Inhaltsbereiche und Wissenstypen des fachbezogenen Professionswissens tatsächlich erfassen. Die Studienergebnisse stützen die Validität der biologie- und physikbezogenen Tests. Der Beitrag aus Ko-WADIS widmet sich dem Erwerb von Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung in den Fächern Biologie, Chemie und Physik. Er fokussiert dabei auf die Bedeutung des Studienfortschritts für den Erwerb von Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung. Außerdem widmet er sich der Frage, ob die Berücksichtigung von Methodenkompetenzen im Studium bzw., ob das zweite Unterrichtsfach für den Erwerb von Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung relevant sind. Die Ergebnisse der Studie weisen darauf hin, dass die Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung deutliche Schnittmengen zwischen den Fächern aufweisen, und dass ein Kompetenzzuwachs im Studienverlauf stattfindet. Eine explizite Vermittlung von Methodenkompetenzen scheint einer impliziten Vermittlung nicht überlegen. Allerdings übertreffen Studierende mit zwei naturwissenschaftlichen Fächern erwartungsgemäß ihre Kommilitonen mit nur einem naturwissenschaftlichen Fach. Das Symposium liefert Erkenntnisse über die Bedeutung des Professionswissens für Unterrichtsqualität und Lernerfolg von Schülerinnen und Schülern und erlaubt Schlussfolgerungen dazu, wie der Erwerb des Professionswissens unterstützt werden kann.

- Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell (Ed.), *Handbook of research on science teacher education* (pp. 1105-1149). Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1988). *Kognitive Psychologie*. Heidelberg: Spektrum.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469-520.
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (pp. 29-53). Münster: Waxmann.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180.
- Borowski et al. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN) – Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 341-349.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Ed.), *Enzyklopädie der Psychologie: Pädagogische Psychologie*, Bd. 3. *Psychologie des Unterrichts und der Schule* (pp. 177-212). Göttingen: Hogrefe.
- Germ, M., Müller, A., & Harms, U. (in Druck). Naturwissenschaftsdidaktische Lernaufgaben, generatives Lernen und wahrgenommene Kohärenz im naturwissenschaftlichen Lehramtsstudium. *Zeitschrift für Didaktik der Naturwissenschaften*.
- Großschedl, J., Harms, U., & Glowinski, I. (2013). Universitäre Biologielehrerbildung als fachdidaktischer Forschungsgegenstand. *Biologie in unserer Zeit (BiuZ)*, 43, 147-149.
- Hattie, J. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement*. London: Routledge.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42, 371-406.
- Jüttner, M., Boone, W., Park, S., & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational Assessment, Evaluation and Accountability*, 25, 45-67. doi:10.1007/s11092-013-9157-y
- KMK (2008). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2008/2008\\_10\\_16-Fachprofile-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf) (22.01.2013).
- Kunter, M., Baumert, J., Voss, T., Klusmann, U., Richter, D., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instrumental quality and student Development. *Journal of Educational Psychology*, 105, 805-820.
- Magnusson, S., Krajcik, J. S., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *PCK and Science Education* (pp. 95-132). Netherland: Kluwer Academic Publisher.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Eds.), *Theorien in der biologiepädagogischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden* (pp. 177-186). Berlin: Springer.
- Messick, S. (1995). Validity of Psychological Assessment – Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning. *American Psychologist*, 50, 741-749.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus (Version 5.21) [ComputerSoftware]*. LosAngeles: Author.
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Studien zum Physik- und Chemielernen [Dissertation]. Berlin: Logos Verlag.
- Riese, J., & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167-187.
- Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35, 265-296.
- Seidel, T., & Shavelson, R. J. (2007). Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results. *Review of Educational Research*, 77, 454-499. doi:10.3102/0034654307310317
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14.
- Shulman, L. S. (1987). Knowledge and teaching of the new reform. *Harvard Educational Review*, 57, 1-22.
- Tepper, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., ... Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.
- Thillmann, H., & Wirth, J. (in Vorb.). Assessment of science teachers' pedagogical knowledge.
- Upmeyer zu Belzen, A., & Krüger, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 41-57.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G161



Sonja Werner (LMU München), Christian Förtsch (LMU München),  
Melanie Sczudlek (Friedrich-Alexander-Universität Erlangen-Nürnberg), Birgit Neuhaus (LMU München)

## **Das Projekt ProwiN – Ergebnisse aus der Perspektive der Biologie**

Die Forschung zur Unterrichtsqualität (engl.: Teaching Effectiveness) und Lehrerprofessionalisierung (engl. Teacher Professionalism) hat bereits eine lange Tradition. Neu sind aber die Ansätze beide Forschungslinien miteinander zu verbinden und hierbei besonders auf fachspezifische Aspekte zu fokussieren. Ziel der hier vorgestellten Studie ist es daher, Zusammenhänge zwischen verschiedenen Dimensionen des Professionswissens einer Biologielehrkraft, ihrer Unterrichtsgestaltung und dem Leistungszuwachs ihrer Schülerinnen und Schüler zu analysieren und hieraus Rückschlüsse für die Gestaltung der Biologielehrerausbildung zu ziehen.

### Theoretischer Hintergrund

Theoretische Grundlage der Studie im Bereich des Professionswissens ist ein im Rahmen des Projektes entwickeltes Modell (Tepner et al., 2012), in dem das Professionswissen, aufbauend auf Arbeiten von Shulman (1987) und COACTIV (Baumert & Kunter, 2011) in drei Dimensionen untergliedert wird: das pädagogische Wissen (pedagogical knowledge, PK), das fachdidaktische Wissen (pedagogical content knowledge, PCK) und fachliche Wissen (content knowledge, CK). Die drei Dimensionen werden in Anlehnung an Anderson (1988) und Paris et al. (1983) in drei Wissensarten unterteilt (deklaratives Wissen, prozedurales Wissen und konditionales Wissen). PK und PCK werden zudem in verschiedene Facetten untergliedert (vgl. Tepner et al., 2012). Für PCK sind dies Experimente, Modelle/Konzepte und Schülervorstellungen, für PK sind es Klassenführung, Unterrichtsmethoden, individuelle Lernprozess und Leistungsbeurteilung. Basierend auf diesem Modell wurde für die Biologie in der ersten Projektphase ein Messinstrument entwickelt, mit dem es möglich ist, die drei Dimensionen valide, reliable und objektiv zu erfassen (PCK und CK: Jüttner, Boone, Park, & Neuhaus, 2013, PK: Thillmann & Wirth, in Vorb.). Theoretische Grundlage der Studie im Bereich der Unterrichtsqualität sind die in den letzten Jahrzehnten durchgeführten Meta-Analysen zur Unterrichtsqualität (vgl. z. B. Hattie, 2009; Seidel & Shavelson, 2007). Die COACTIV-Studie konnte zeigen, dass das Fachwissen von Mathematiklehrern sehr hoch mit dem fachdidaktischen Wissen der Lehrkraft korreliert ( $r = 0,79$ ), das fachdidaktische Wissen Einfluss auf verschiedene Facetten der Unterrichtsgestaltung nimmt und diese wiederum in unterschiedlicher Ausprägung die Lernleistung der Schülerinnen und Schüler beeinflusst. Die Unterrichtsqualität wurde jedoch in der COACTIV-Studie über Schülerbefragungen erfasst, in der hier vorgestellten Studie werden Videos aufgenommen.

### Fragestellung:

Aus dem oben dargestellten Hintergrund lassen sich folgende Fragestellungen entwickeln:

- (1) Wie wirken sich verschiedene Facetten des Fachwissens, des fachdidaktischen Wissens und pädagogischen Wissens von Biologielehrkräften auf die Unterrichtsgestaltung aus?
- (2) Wie wirken die Gestaltung des Biologieunterrichts auf den Leistungszuwachs der Schülerinnen und Schüler aus?
- (3) Welche Konsequenzen können aus den gewonnenen Erkenntnissen für die Biologielehrerausbildung gezogen werden?

### Methode:

Zur Beantwortung der Fragestellungen werden 40 Lehrkräfte im Biologieunterricht der 9. Jahrgangsstufe videografiert, ihr Professionswissen wird mittels der in der 1. Phase entwickelten Fragebögen erhoben. Zudem wird der Wissenszuwachs der Schülerinnen und Schüler im Prä-Post-Test-Design erfasst.

### Erste Ergebnisse:

Im Rahmen des Vortrages werden das Design der Studie, Ergebnisse zur Qualität der Fragebögen und erste Auswertungen zu bayerischen Unterrichtsvideos dargestellt. Dabei wird auf die Facetten Experimente/Erkenntnisgewinnung mit den Subdimensionen ‚reflektierter Umgang mit Experimenten‘ und ‚Modellen‘ sowie ‚Einbettung von Originalen‘ fokussiert. Die Beurteilerübereinstimmungen für alle Kategorien wiesen akzeptable Werte auf (alle  $\kappa$ -Werte lagen zwischen 0,54 und 0,96). In mehr als der Hälfte der Stunden wurden Originale (73 %) und Experimente (89 %) verwendet. In 17 % der Stunden arbeiteten Lehrkräfte und Schüler mit Modellen. 12,41 % der Unterrichtszeit wurde auf die Arbeit mit Modellen verwendet, 68,91 % auf die Arbeit mit Experimenten und 36,55 % auf die Arbeit mit Originalen.

### Ausblick/Diskussion:

Ziel ist es nach Abschluss des Projektes differenzierte Aussagen dazu machen zu können, wie sich die verschiedenen Dimensionen der Lehrerprofessionalität mit ihren unterschiedlichen Facetten auf die Unterrichtsgestaltung auswirken. Das Testinstrument, das sehr gut zwischen den drei Dimensionen Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen differenzieren kann, bildet dafür eine solide Grundlage.



## **Entwicklung und Validierung eines Instruments zur Erfassung des biomedizinischen Wissens angehender Biologielehrkräfte**

### Theoretischer Hintergrund

Das Professionswissen von Lehrkräften wird als wichtiger Einflussfaktor auf die Instruktionsqualität im Unterricht angesehen und spielt damit eine zentrale Rolle für den Lernprozess von Schülerinnen und Schülern (Baumert et al., 2010; Hill, Rowan, & Ball, 2005). In der Forschung hat sich eine Dreiteilung des Professionswissens in (1) Fachwissen, (2) fachdidaktisches Wissen und (3) bildungswissenschaftliches Wissen etabliert (Baumert & Kunter, 2006; Shulman 1986, 1987). Dieser Beitrag legt einen Schwerpunkt auf das fachdidaktische Wissen angehender Biologielehrkräfte, das im KiL -Projekt („Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen“) in fünf Inhaltsbereichen (Wissen über [a] Schülerkognitionen, [b] Instruktionsstrategien, [c] Curriculum, [d] Bewertung, [e] fachdidaktische Forschung) konzeptualisiert wird (Großschedl, Harms, & Glowinski, 2013). Neben den Inhaltsbereichen werden auf einer zweiten Dimension kognitive Prozesse unterschieden, die die intendierten Fähigkeiten angehender Lehrkräfte bei der Nutzung dieses Wissens konkretisieren ([a] Erinnern & Abrufen, [b] Verstehen & Anwenden, [c] Generieren von Handlungsoptionen). Die Konzeptualisierung des biomedizinischen Wissens stellt dabei eine Synthese aus etablierten Modellen zur Beschreibung fachdidaktischen Wissens (z. B. Shulman, 1986; Magnusson, Krajcik, & Borko, 1999), den fachbezogenen Lehrerbildungsstandards (KMK, 2008) und 27 Modulkatalogen verschiedener Hochschulen dar (Großschedl et al., 2013). Da es bisher an Instrumenten zur Erfassung des biomedizinischen Wissens angehender Lehrkräfte mangelt, zielt diese Studie auf eine Validierung eines Tests ab, der auf Basis der oben beschriebenen Konzeptualisierung biomedizinischen Wissens entwickelt wurde.

### Fragestellungen

Lässt sich das biomedizinische Wissen angehender Biologielehrkräfte mithilfe des entwickelten Instruments reliabel und valide erfassen?

### Empirische Forschungsmethodik

An einer Fragebogenerhebung nahmen 274 Lehramtsstudierende (Alter: M = 22,8 Jahre [ SD = 2,3]; Geschlecht: 78,5% weiblich) mit dem Unterrichtsfach Biologie teil (Studiendauer: M = 4,7 Semester [ SD = 2,5]; gymnasiales Lehramt: 64,7%, nicht-gymnasiales Lehramt: 35,3%). Die Probanden bearbeiteten Tests zur Erfassung des biomedizinischen Wissens ( n =66 Items), biologischen Fachwissens ( n =77 Items, EAP/PV-Reliabilität=.85) und bildungswissenschaftlichen Wissens ( n =33 Items, EAP/PV-Reliabilität =.75). Der Itempool wurde auf vier Testhefte verteilt (Multimatrix-Design). Zur Überprüfung der diskriminanten Validität der Messinstrumente wurde eine konfirmatorische Faktorenanalyse mit Mplus (Muthén & Muthén, 2007) durchgeführt. Um weitere Validitätshinweise zu finden, wurden latente Regressionen zwischen dem biomedizinischen Wissen und den folgenden Kriteriumsvariablen durchgeführt: (a) Skalen des KFT (Figurenanalogien; Zahlenreihen; Wortklassifikationen), (b) Abiturnote, (c) angestrebter Lehramtszugang (gymnasial vs. nicht-gymnasial), (d) Zweitfach (naturwissenschaftlich vs. nicht-naturwissenschaftlich) und (e) Lehrerfahrung.

### Ergebnisse

Das Messinstrument zur Erfassung des biomedizinischen Wissens weist eine gute Reliabilität auf (EAP/PV-Reliabilität=.88). Die Faktorenanalyse zeigt, dass das biomedizinische Wissen eine vom biologischen Fachwissen und bildungswissenschaftlichen Wissen empirisch trennbare Fähigkeit darstellt. Das biomedizinische Wissen korreliert dabei hoch mit dem Fachwissen ( r =.68; p <.01) und moderat mit dem bildungswissenschaftlichen Wissen ( r =.35; p <.01). Zwischen dem Fachwissen und dem bildungswissenschaftlichen Wissen besteht kein signifikanter Zusammenhang ( r =.11; p >.05).

Die latenten Regressionsanalysen zeigen ein höheres biomedizinisches Wissen angehender Gymnasiallehrkräfte (im Vergleich zu Lehrkräften anderer Schularten; B =.72\*\*\*, SE =.16). Dies weist darauf hin, dass sich eine intensive biologische Fachausbildung auch auf das biomedizinische Wissen positiv auswirkt. Außerdem zeigen die Analysen, dass sich ein zweites naturwissenschaftliches Unterrichtsfach ( B =.31\*, SE =.16) sowie eine gute Abiturnote ( B =-.63\*\*\*, SE =.14) positiv auf das biomedizinische Wissen auswirken. Darüber hinaus besteht ein positiver Zusammenhang zwischen dem biomedizinischen Wissen und der Anzahl eigener Unterrichtsversuche ( B =.24\*\*, SE =.07). Diese Ergebnisse stützen die Validität des entwickelten Messinstruments.

### Relevanz der Forschungsergebnisse

Im Rahmen des KiL -Projekts wurde ein Messinstrument zur Erfassung des biomedizinischen Wissens angehender Biologielehrkräfte entwickelt, welches nun in längsschnittlichen Untersuchungen weiter genutzt werden soll. Unsere Ergebnisse weisen darauf hin, dass die drei Wissensbereiche häufig nur unzulänglich vernetzt werden. Diesem Problem könnte durch gemeinsame Lehrveranstaltungen der Fachdidaktiken, Fachwissenschaften und Bildungswissenschaften und durch diese Wissensbereiche vernetzende Lernaufgaben (Germ, Müller & Harms, in Druck) im Studium begegnet werden.

## Erfassung des Professionswissens angehender Physiklehrkräfte im Rahmen des Projekts KiL

### Theoretischer Hintergrund

Der professionellen Kompetenz von Lehrkräften wird eine zentrale Bedeutung für die Gestaltung guten Unterrichts zugeschrieben (Abell 2007; Kunter et al., 2013). Das Professionswissen stellt dabei einen wichtigen Teil der professionellen Kompetenz dar (Baumert & Kunter, 2006). In Anlehnung an Shulman (1986, 1987) hat sich eine Einteilung des Professionswissens in die Dimensionen Fachwissen (CK), fachdidaktisches Wissen (PCK) und pädagogisches Wissen (PK) etabliert (vgl. Abell, 2007; Borowski et al., 2010). Sie konnte im Rahmen der COACTIV-Studie für das Fach Mathematik empirisch bestätigt werden (Baumert & Kunter, 2006). Für die naturwissenschaftlichen Fächer wird im Projekt ProWiN auf Basis eines entsprechenden Modells das Professionswissen von Lehrkräften und dessen Einfluss auf deren Unterricht untersucht (Borowski et al., 2010). Es kann jedoch nicht ohne Weiteres angenommen werden, dass sich Befunde über das Professionswissen von Lehrkräften auf Studierende übertragen lassen. Zum Professionswissen Lehramtsstudierender liegen jedoch nur wenige Befunde vor. Für das Fach Physik sind hier die Arbeiten von Riese (2009) zu nennen (vgl. auch Riese & Reinhold, 2010). Die entsprechenden Ergebnisse beruhen jedoch auf einer Fokussierung auf einzelne Aspekte wie die Mechanik und das Experimentieren. Hier setzt das Projekt KiL an, in dem die Entwicklung des Professionswissens in voller Breite erfasst werden soll.

### Fragestellung

Folgende Forschungsfragen sollen für das Fach Physik untersucht werden:

1. Inwieweit lassen sich Befunde über die Struktur des Professionswissens von Lehrkräften auf angehende Lehrkräfte übertragen?
2. Wie verändert sich die Struktur des Professionswissens angehender Lehrkräfte gegebenenfalls im Laufe der universitären Ausbildung?

### Methode

Nach der Modellierung des zu erfassenden Konstrukts wurde die fachdidaktische Literatur auf veröffentlichte Aufgaben zur Erfassung des fachlichen und fachdidaktischen Wissens geprüft. Externe Experten wurden für die Aufgabenentwicklung hinzugezogen, um gezielt Lücken in der Abdeckung des Modells zu schließen. Die Aufgaben wurden im Rahmen einer bundesweiten Pilotstudie, aufgeteilt auf zwei Testhefte, an 12 Hochschulen einer Stichprobe von  $N = 166$  Studierenden vorgelegt. Anschließend wurden die Aufgaben von Experten intensiv diskutiert. Kritische Aufgaben wurden zudem Studierenden vorgelegt, um Erkenntnisse zur kognitiven Validität (nach Messick, 1995) zu gewinnen. Nach einer Überarbeitungsphase wurden die verbliebenen Aufgaben zu einem neuen Testheft zusammengestellt, mit dem im Rahmen der Hauptuntersuchung im Sommersemester 2013 bundesweit Daten erhoben werden. Dabei wird eine Stichprobengröße von  $N = 200$  angestrebt. Parallel dazu werden 24 zufällig gezogene Aufgaben in einer Think-Aloud-Studie auf ihre kognitive Validität hin überprüft.

### Ergebnisse

Als Maß für die Reliabilität wurde Cronbachs Alpha berechnet. Die Analyse ergab  $\alpha = .70$  für Testheft I bzw.  $\alpha = .77$  für Testheft II im Bereich CK und  $\alpha = .62$  für Testheft I und  $\alpha = .60$  für Testheft II im Bereich PCK. Zur Untersuchung der strukturellen Validität des Tests wurden Zusammenhänge zwischen CK und PCK getrennt für die Testhefte untersucht. Zur Bestimmung der externen Validität wurden Korrelationen mit Abiturnote und Semesterzahl der Probanden berechnet. Zusätzlich wurde der Zusammenhang mit verbalen, nonverbalen und quantitativen kognitiven Fähigkeiten bestimmt. Zwischen CK und PCK ergab sich für Heft I eine Korrelation von  $r = .54^{**}$  und für Heft II von  $r = .39^{**}$ . Bezüglich der externen Validität liegen die Ergebnisse im Wesentlichen im erwarteten Bereich. Dabei zeigt sich vor allem ein Einfluss der Semesterzahl und der verbalen und nonverbalen Subskalen des Kognitive-Fähigkeiten-Tests. Die Korrelation mit der quantitativen Subskala fällt nur für den CK des ersten Testhefts statistisch bedeutsam aus, was in einer Häufung von Aufgaben, die mathematische Berechnungen erfordern, begründet sein dürfte.

Bei der Think-Aloud-Studie konnte bei 23 von 24 geprüften Aufgaben ein überwiegender Anteil konstruktrelevanter Lösungsstrategien identifiziert werden. Damit kann davon ausgegangen werden, dass die Aufgaben überwiegend die intendierten kognitiven Prozesse evaluieren. Neben den obigen Pilotierungsergebnissen sollen im Vortrag auch neue Daten aus der im Sommersemester 2013 durchgeführten Hauptstudie vorgestellt werden.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G161

Stefan Hartmann (Humboldt-Universität zu Berlin), Annette Upmeier zu Belzen (Humboldt-Universität zu Berlin), Dirk Krüger (Freie Universität Berlin)

## **Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung als Teil des Professionswissens zukünftiger Lehrerinnen und Lehrer: Erste Ergebnisse einer Längsschnittstudie**

In Lehramtsstudiengängen der Fächer Biologie, Chemie und Physik sind Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung als Teil des Professionswissens (Bromme, 1997) von zentraler Bedeutung. Sowohl das Arbeiten im Laborpraktikum zum Erwerb fachlichen Wissens als auch die Anwendung von Methoden der Erkenntnisgewinnung beim naturwissenschaftlichen Unterrichten erfordern Methodenkompetenzen wie das Formulieren von Forschungsfragen und Hypothesen, das Planen von Untersuchungen, das Auswerten und Interpretieren von Ergebnissen sowie das Arbeiten mit naturwissenschaftlichen Modellen (Mayer, 2007; Upmeier zu Belzen und Krüger, 2010). Während Methoden der Erkenntnisgewinnung mittlerweile ein fester Bestandteil bei der Evaluation schulischer Bildungsprozesse sind, steht eine systematische Erfassung dieser Kompetenzen in der akademischen Ausbildungsphase von Lehrkräften der naturwissenschaftlichen Fächer noch aus. Dabei stellen sich drei zentrale Forschungsfragen: (1) Wie entwickeln sich Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung im Verlauf eines Lehramtsstudiums in den Fächern Biologie, Chemie und Physik? (2) Inwieweit unterscheiden sich die Leistungen Lehramtsstudierender, in deren Studium die Vermittlung von Methodenkompetenzen fest implementiert ist, von den Leistungen Studierender, in deren naturwissenschaftlichem Studium Kompetenzen der Erkenntnisgewinnung nicht explizit vermittelt werden? (3) Inwieweit unterscheiden sich die Leistungen Lehramtsstudierender mit einem naturwissenschaftlichen Fach von den Leistungen Lehramtsstudierender mit zwei naturwissenschaftlichen Fächern?

Für die Untersuchung dieser Fragestellungen wurde ein schriftlicher Test entwickelt, mit dem anhand verschiedener Kontexte aus den drei Fächern Biologie, Chemie und Physik Kompetenzen der Erkenntnisgewinnung evaluiert werden. Für die längsschnittliche Testung wurden 123 Aufgaben (42 Biologie, 33 Chemie, 48 Physik) im geschlossenen Format entwickelt. Die Multiple-Choice-Antwortoptionen wurden aus Aussagen Studierender generiert und damit existierende Vorstellungen berücksichtigt (Sadler, 1998). Die Auswahl geeigneter Items erfolgte auf Basis der Ergebnisse von zwei Pilotierungsstudien ( $N_1 = 259$ ,  $N_2 = 831$ ). Das Instrument wird als Multikohorten-Studie im Längsschnitt mit vier Messzeitpunkten im Verlauf des Lehramtsstudiums an drei deutschen Universitäten zum Einsatz kommen. Zu Vergleichszwecken werden auch Studierende der Monostudiengänge Biologie, Chemie und Physik sowie Studierende in Österreich getestet. Bei Letzteren ist eine explizite Vermittlung von Kompetenzen der Erkenntnisgewinnung im Studium nicht vorgesehen. Ein Multimatrixdesign ermöglicht eine zeitsparende und ökonomische Testung der Studierenden.

Mit den Daten aus der zweiten Pilotierungsstudie wurden erstmals Analysen zur Beantwortung der Forschungsfragen durchgeführt. Dabei wurden verschiedene ein- und mehrdimensionale Skalierungen vorgenommen, die zufriedenstellende Reliabilitäten der Personenschätzer aufweisen. Der Test weist eine angemessene Schwierigkeit auf und die Varianz der Personenfähigkeiten wird über die Itemschwierigkeiten hinreichend abgedeckt.

Die Ergebnisse zeigen, dass die Leistungen in den drei Fächern hoch korrelieren ( $.81 \leq r_{\text{latent}} \leq .94$ ). Dies wird als Hinweis darauf gewertet, dass die getesteten Kompetenzen in hohem Maß über die fachspezifischen Kontexte hinweg generalisierbar sind. Die Forschungsfragen wurden mit latenten Regressionsanalysen geprüft, bei denen verschiedene Eigenschaften der Versuchspersonen als Prädiktoren auf die eindimensional skalierte Testleistung in die Berechnungen eingingen.

Die Frage nach dem Verlauf der Kompetenzentwicklung lässt sich auf Basis der querschnittlichen Pilotierung nicht direkt beantworten. Die Ergebnisse der latenten Regressionsanalyse zeigen jedoch einen Effekt der Studienphase (Bachelor vs. Master) auf die Leistungen: Masterstudierende erzielen bessere Leistungen als Bachelorstudierende ( $B = .461$ ,  $SE = .09$ ). Dies weist darauf hin, dass mit fortschreitendem Studium ein Kompetenzzuwachs zu verzeichnen ist. Bezüglich der zweiten Forschungsfrage wurde kein Effekt gefunden: In der getesteten Stichprobe zeigt sich kein Einfluss der Studienart (kompetenzbasiert vs. nicht kompetenzbasiert) auf die Testleistung. Dies könnte ein Hinweis darauf sein, dass solche Kompetenzen im naturwissenschaftlichen Studium auch dann erworben werden, wenn sie in den Lehrveranstaltungen nicht explizit vermittelt werden. Bezüglich der dritten Forschungsfrage wurde wiederum ein signifikanter Effekt gefunden: Studierende mit zwei naturwissenschaftlichen Fächern erzielen bessere Leistungen als Studierende mit einem naturwissenschaftlichen Fach ( $B = .154$ ,  $SE = .05$ ).

Die Befunde der Pilotierungsstudie geben erstmals Auskunft über die Struktur von Kompetenzen der Erkenntnisgewinnung im naturwissenschaftlichen Lehramtsstudium und stellen einen wichtigen Schritt auf dem Weg zu deren längsschnittlicher Erfassung dar.

## Kompetenzentwicklung bei jungen Kindern in verschiedenen Bereichen

Chair(s): Miriam Leuchter (Westfälische Wilhelms-Universität)

Diskutant/in: Ilonca Hardy (Goethe-Universität Frankfurt)

Kompetenzen befähigen eine Person, komplexe Anforderungen einer konkreten Situation zu meistern (Furrer, 2000) und sind demnach in verschiedenen Situationen in unterschiedlichsten Lebensbereichen zentral. Zur Bewältigung der spezifischen Anforderungen müssen verschiedene Ressourcen herangezogen werden: Wissen (z.B. Wissen zu naturwissenschaftlichen Inhaltsbereichen), Fertigkeiten (z.B. Schreiben und Lesen), Fähigkeiten (z.B. Empathie) und Ressourcen des Umfeldes (z.B. Interaktionen mit Peers) (Furrer, 2000).

Junge Kinder im Alter zwischen vier und acht Jahren durchlaufen sowohl kognitiv als auch motivational große Entwicklungsfortschritte (Hasselhorn & Grube, 2008), die es ihnen ermöglichen, in verschiedenen Bereichen Kompetenzen aufzubauen.

Intuitives Wissen über Inhaltsbereiche, z.B. physikalisch-technische Phänomene, ist bereits bei sehr jungen Kindern vorhanden (Siegler, 1976). Es bietet sich daher an, frühzeitig an bestehende Konzepte der Kinder anzuknüpfen, um einen inhaltsbezogenen Kompetenzzuwachs anzuregen.

Die Anbahnung des Schriftspracherwerbs erfolgt stets in Zusammenhang mit inhaltlichem Wissen (Becker-Mrotzek & Schindler, 2007). Voraussetzung für das Erlernen von Kulturtechniken wie Schreiben und Lesen sind u.a. im Kindergartenalter stattfindende Veränderungen im phonologischen Arbeitsgedächtnis (Hasselhorn & Grube, 2008). Entscheidend für den Aufbau von kognitiven Kompetenzen in allen Bereichen sind die sozialen Kompetenzen (Hinsch & Pfungsten, 2007) einer Person. Als Einflussfaktor wirken soziale Interaktionen. Diese können über die Auseinandersetzung mit unterschiedlichen Ansichten die kognitive Entwicklung fördern als auch über Annäherungsprozesse soziale Kompetenzen unterstützen (Petillon, 2011).

Kompetenzzuwachs, unabhängig vom Bereich, kann durch Scaffolding angeregt werden. Dabei wird der Lernende durch die Unterstützung einer kompetenteren Person dazu befähigt, die Anforderungen einer konkreten Situation zu meistern, die er ohne solche Unterstützung nicht hätte lösen können (Wood, Bruner, & Ross, 1976). Diese Unterstützung kann sich auf unterschiedlichste Art und Weise manifestieren und durch das konkrete Umfeld, durch Peers oder durch Lehr-/Erziehungspersonen erfolgen. Scaffolding kann im alltäglichen Spiel (Leuchter, 2013) oder während einer strukturierten Situation (Hogan & Pressley, 1997) stattfinden.

Im Symposium wird die Kompetenzentwicklung von jungen Kindern in verschiedenen Bereichen in den Blick genommen.

1. Julia Stipp untersucht den Einfluss verschiedener Scaffoldingmaßnahmen auf das naturwissenschaftliche Lernen von 7-jährigen Kindern zum Thema potenzielle Energie.
2. Die Auswirkungen verbaler Scaffoldingmaßnahmen auf das Lernen sowie den Wissenstransfer von 7-jährigen Kindern zum naturwissenschaftlichen Thema einarmige Hebel werden von Britta Naber präsentiert.
3. Die Effekte eines spielerischen Alltagstrainings im Vergleich zu einem hochstrukturierten Trainingsprogramm auf schriftsprachliche Kompetenzen von Kindergartenkindern stellt Susanne Geyer dar.
4. Nicole Sturmhöfel befasst sich mit den Auswirkungen jahrgangsübergreifenden Lernens auf soziale und emotionale Kompetenzen von Kindern im Übergang vom Kindergarten zur Grundschule.

- Becker-Mrotzek, M., & Schindler, K. (2007). Schreibkompetenz modellieren. In M. Becker-Mrotzek & K. Schindler (Hrsg.), *Texte Schreiben* (S. 7-26). Duisburg: Gilles und Francke-Verlag.
- Furrer, H. (2000). Ressourcen-Kompetenzen-Performanz. *Kompetenzmanagement für Fachleute der Erwachsenenbildung*. Luzern.
- Hasselhorn, M., & Grube, D. (2008). Individuelle Voraussetzungen und Entwicklungsbesonderheiten des Lernens im Vorschul- und frühen Schulalter. *Empirische Pädagogik*, 22(2), 113-126.
- Hinsch, R., & Pfungsten, U. (2007). Gruppentraining sozialer Kompetenzen GSK. Weinheim: Beltz.
- Hogan, K., & Pressley, M. (1997). Scaffolding scientific competencies within classroom communities of inquiry. In K. Hogan, & M. Pressley (Eds.), *Scaffolding Student Learning: Instructional Approaches and Issues*. *Advances in Teaching and Learning Series* (pp. 74-107). Brookline Books.
- Leuchter, M. (2013). Die Bedeutung des Spiels in Kindergarten und Schuleingangsphase. *Zeitschrift für Pädagogik*, 59(4), 575-592.
- Petillon, H. (2011). Grundschulkind und ihre sozialen Beziehungen. In W. Einsiedler, M. Götz, A. Hartinger, F. Heinzl, J. Kahler & U. Sandfuchs (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (S.168-167). Stuttgart, Regensburg: Klinkhardt.
- Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, 8(4), 481-520.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89-100.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G165



## ***Der Einsatz von Scaffoldingmaßnahmen zur Unterstützung des Lernens von jungen Kindern zum physikalischen Phänomen der potenziellen Energie***

Potenzielle Energie wird definiert als das Arbeitsvermögen, das einem Körper aufgrund seiner Lage oder seines Zustandes zukommt (Stroppe, 1999). Direkt beobachtbar sind nur die Folgen von Energiespeicherung oder -transformation, nicht Energie selbst (Diakidoy, Kendeou, & Ioannides, 2003). Energie stellt somit ein abstraktes, physikalisches Phänomen dar. Es vergehen oft viele Jahre an naturwissenschaftlichem Unterricht, bis ein Kind das Konzept der Energie versteht (Vosniadou, 2007). Dabei besteht ein intuitives Verständnis für die Subkonzepte der potenziellen Energie, Höhe und Gewichtskraft, bereits im Säuglingsalter (Baillargeon & DeVos, 1991; Baillargeon et al., 2009).

Die schiefe Ebene stellt einen für potenzielle Energie relevanten Kontext dar. Konzepte von Laien zur schiefen Ebene sind häufig durch Oberflächenmerkmale wie Höhe oder Neigung charakterisiert, wohingegen die Konzepte von Experten relevante physikalische Prinzipien (z.B. Energieerhaltung) beinhalten (Chi, Feltovich, & Glaser, 1981). In Übereinstimmung damit haben auch Erstklässler Schwierigkeiten, die Auswirkungen von Variationen potenzieller Energie im Kontext einer schiefen Ebene zu beurteilen (Stipp & Leuchter, in Vorbereitung).

Naturwissenschaftlicher Unterricht hat zum Ziel, von den Vorstellungen der Kinder ausgehend einen konzeptuellen Wandel anzuregen. Häufig werden Scaffoldingmaßnahmen eingesetzt, um diesen Prozess zu unterstützen. Als „Scaffolds“ werden Unterstützungsmaßnahmen bezeichnet, die den Lernenden dazu befähigen, ein Problem oder eine Aufgabe zu lösen, die er ohne diese Unterstützung nicht hätte lösen können (Wood, Bruner, & Ross, 1976).

Grundlegend sind zum einen die Bereitstellung von konkreten Materialien und der aktive Umgang mit diesem (Möller, 2004). Weiterhin dienen verbale Scaffolds beispielsweise dazu, die Aufmerksamkeit des Lerners auf zentrale Merkmale zu lenken oder durch gezielte Anstöße das Denken anzuregen (Hogan & Pressley, 1997). Diese Aspekte können durch den Einsatz von externen Repräsentationen, z.B. Bilder, weiter unterstützt werden (Reid & Beveridge, 1986). Im Rahmen einer Interventionsstudie (Diakidoy et al., 2003) konnte gezeigt werden, dass Sechstklässler mehr über das physikalische Konzept der Energie lernen, wenn sie einen Text lesen, der Fehlkonzepte der Schüler aufgreift. Die Leistung der Schüler blieb jedoch gering, weshalb die Autoren annehmen, dass zusätzliche Scaffoldingmaßnahmen nötig sind, um einen Konzeptwandel anzuregen. In einer anderen Studie lernen 10- bis 11-jährige Schüler durch Fragen der Lehrperson, experimentelle Beobachtungen und externe Repräsentationen signifikant mehr zu den physikalischen Konzepten Kraft und Energie als Schüler, die regulären Unterricht erhalten (Vosniadou et al., 2001). Unklar ist, ob das Verständnis für Energie auch bereits bei jüngeren Kindern im Übergang zur Grundschule mithilfe von Scaffoldingmaßnahmen angeregt werden kann.

Die vorliegende Interventionsstudie mit Prä- und Posttest untersucht, durch welche Scaffoldingmaßnahmen das Lernen von sechs bis siebenjährigen Kindern zur potenziellen Energie angeregt werden kann.

In den Items der Tests beurteilen die Kinder die Auswirkungen der unterschiedlichen potenziellen Energie von Kugeln in unterschiedlichen Kontexten (schiefe Ebene, freier Fall). In einer 25-minütigen Intervention mit Gruppen von 16 Kindern werden die Situationen aus den Items durchgeführt und von den Kindern beobachtet. Jeweils  $\frac{1}{4}$  aller Kinder werden zusätzlich durch verbale und bildliche Scaffolds (EG1), verbale Scaffolds (EG2), bildliche Scaffolds (EG3) oder keine weiteren Scaffolds (EG4) unterstützt. Eine Kontrollgruppe erhält zeitgleich eine Intervention zu einem anderen Inhaltsbereich. Nach der Intervention folgt der Posttest. Um einen eventuellen Lernzuwachs auf Nachhaltigkeit zu testen, werden die Kinder vier Wochen nach dem Posttest erneut befragt.

Die Ergebnisse einer Erhebung mit 129 Kindern (Alter:  $M=6.71$  Jahre,  $SD=.38$ ) deuten an, dass eine Kombination verbaler und bildlicher Unterstützungsmaßnahmen den größten Lernzuwachs bewirkt, wobei sich Unterschiede je nach Itemkontext zeigen. Z.B. im Kontext „schiefe Ebene“ lernen die Kinder, die verbale und bildliche Scaffolds erhalten, signifikant hinzu ( $t(22)=-3.849$ ,  $p=.001$ ,  $d=.84$ ); In den anderen Experimentalgruppen und der Kontrollgruppe bestehen keine signifikanten Unterschiede zwischen Prä- und Posttest (alle  $t < 1.932$ ; n.s.). Auf der Tagung werden diese und weitere Ergebnisse vorgestellt und hinsichtlich der eingesetzten Scaffoldingmaßnahmen diskutiert.



## **Förderung von Wissenstransfer durch verbale Scaffolds bei einem physikalisch-technischen Thema in der Schuleingangsphase – Eine Interventionsstudie**

### Theoretischer Hintergrund

Ein wichtiges Bildungsziel besteht darin, dass Lernende Wissen über Prinzipien aufbauen, das sie in verschiedenen Kontexten anwenden können. Diese Fähigkeit, Wissen auf einen anderen Kontext zu übertragen (Transfer), ist insbesondere beim naturwissenschaftlichen Lernen wichtig (Schalk, Saalbach & Stern, 2011). Hebel als physikalisch-technisches Phänomen bieten eine gute Gelegenheit, die Rolle des Kontextes zu untersuchen. Alle Hebel unterliegen dem gleichen Funktionsprinzip (Hebelgesetz:  $\text{Last} \times \text{Lastarm} = \text{Kraft} \times \text{Kraftarm}$ ), weisen aber unterschiedliche Erscheinungsformen auf (z.B. Balkenwaage, Schubkarre). Aufgrund der unterschiedlichen Erscheinungsformen erkennen Lernende häufig nicht die strukturelle Ähnlichkeit zwischen verschiedenen Kontexten und es fällt ihnen schwer, ihr Wissen in einem neuen Kontext anzuwenden (Schalk et al., 2011).

Das Wissen junger Kinder zu Hebeln wurde bereits in unterschiedlichen Kontexten untersucht: Im Kontext der Balkenwaage (zweiseitiger Hebel) zeigt Siegler (1978), dass 5-Jährige ausschließlich das Gewicht beachten. 50% der 8-Jährigen berücksichtigen in ihrem Urteil den Abstand, wenn das Gewicht gleich ist. Stellen sich 5-Jährige eine Balkenwaage als Wippe vor, so wird der Abstand jedoch vor dem Gewicht berücksichtigt (Surber & Gzesh, 1984). Im Kontext Schubkarre (einseitiger Hebel) lösen nur 23.65% der Kinder Aufgaben zum Kraftarm korrekt (Naber & Leuchter, in Vorbereitung). Es gibt also Hinweise darauf, dass sich das Wissen junger Kinder zum Hebel in verschiedenen Kontexten unterscheidet und meist nicht mit dem wissenschaftlich korrekten Funktionsprinzip übereinstimmt.

Diese nicht korrekten Vorstellungen können durch Handlung am Material in Kombination mit verbalen Scaffolds in wissenschaftlich korrekte Vorstellungen überführt werden (Hsin & Wu, 2011). Bisherige Forschung hat insbesondere die Wirksamkeit verbaler Scaffolds gezeigt (Hogan & Pressley, 1997). Day und Cordón (1993) unterstützen 8-Jährige beim Lösen von Balkenwaagen-Aufgaben entweder durch kognitiv aktivierende verbale Scaffolds oder durch ausführliche Erklärungen. Nur der Gruppe mit kognitiv aktivierenden verbalen Scaffolds gelingt es, ihr Wissen auch auf komplexere Formen der Balkenwaage zu übertragen.

### Fragestellung

Die vorliegende Studie untersucht, a) ob eine handlungsorientierte Intervention mit verbalen Scaffolds zum Kraftarm im Kontext Schubkarre zu einem Wissenszuwachs führt und b) ob ein Wissenstransfer auf einen abstrakteren Kontext (Hebelmodell) möglich ist.

### Methode

In einer Interventionsstudie erleben die Kinder ( $N=104$ ; Alter:  $M=6.67$ ,  $SD=.32$ ) eine handlungsorientierte Intervention zum Kraftarm im Kontext Schubkarre mit oder ohne verbale Scaffolds. Eine Kontrollgruppe erhält eine Intervention zu einem anderen Thema. Das Wissen der Kinder zum Kraftarm im Kontext Schubkarre wird im Prä-, Post- und Follow-up-Test erfasst. Im Transfertest, direkt im Anschluss an den Follow-up-Test, bearbeiten die Kinder zusätzlich Aufgaben zu einem abstrakteren Kontext (Hebelmodell).

### Ergebnisse

#### a) Wissensentwicklung im Kontext Schubkarre

Kinder, die die Intervention mit verbalen Scaffolds erleben, lösen sowohl im Post- ( $t(40)=-7.639$ ;  $p<.001$ ;  $d=1.44$ ) als auch im Follow-up-Test ( $t(40)=-7.039$ ;  $p<.001$ ;  $d=1.23$ ) signifikant mehr Aufgaben zum Kraftarm im Kontext Schubkarre korrekt als im Prätest.

#### b) Transfer auf einen abstrakteren Kontext

Im Transfertest lösen Kinder, die verbale Scaffolds erhalten, im Kontext des Hebelmodells signifikant mehr Aufgaben richtig, als Kinder, die keine verbalen Scaffolds erhalten, ( $t(79.3)=-2.191$ ,  $p=.031$   $d=.48$ ). Diesen Kindern scheint ein Transfer vom konkreten Gegenstand Schubkarre auf einen abstrakteren Kontext zu gelingen.

## ***Schriftsprache im Kindergarten entwicklungsorientiert fördern***

Es gibt verschiedene Überlegungen, wie Schriftsprache bereichsspezifisch im Kindergarten gefördert werden kann. Vergleichende empirische Befunde aus den USA sprechen für eine Kombination von sprachlich-kommunikativen Elementen und Sequenzen eines sprachstrukturell-systematischen Trainings von Graphem-Phonem-Korrespondenzen im vorschulischen Bereich (z.B. Connor, Morrison, & Slominski, 2006). Zudem wurde in der amerikanischen „play and literacy“-Forschung Evidenz dafür gefunden, dass sich Kindergartenkinder durch Fiktions- und Rollenspiele zu Lese- und Schreibaktivitäten anregen lassen (im Überblick Ro-kos, Christie, Widman, & Holding, 2010). Im deutschen Sprachraum zeigen Forschungsbefunde, dass eine kombinierte Förderung der Phonologischen Bewusstheit mit gezielten Übungen zum Buchstabe-Laut-Zusammenhangs im letzten Jahr vor der Einschulung besonders effektiv ist für die schulische Leistungsentwicklung im Lesen und Schreiben (Schneider, 2008). Bisher gibt es dazu vor allem Trainingsprogramme für Risikokinder. Unklar ist, wie bereichsspezifische Kompetenzen im Schreiben und Lesen im Sinne einer entwicklungsorientierten Vorbereitung auf die spätere Bildungslaufbahn des Kindes angebahnt werden können. Dazu gehören neben der phonologischen Bewusstheit auch die Aufgeschlossenheit gegenüber Schrift und schriftsprachlichen Aktivitäten sowie eine emotional-motivationale Bereitschaft zum Weiterlernen.

Im Rahmen der BMBF-Studie „Förderung eines entwicklungsorientierten Schriftspracherwerbs im Kindergarten“ (FeS) soll u.a. folgende Frage überprüft werden: Unterscheiden sich Vorschulkinder in ihrer schriftsprachlichen Leistungsentwicklung in einem alltagsintegrierten Literacy-Setting im Vergleich zu einem hochstrukturierten Trainingsprogramm der Phonologischen Bewusstheit?

Dafür wurden 229 Kinder des Einschulungsjahrgangs 2012/2013 in zwei Treatmentgruppen und einer Vergleichsgruppe in einem Prätest-Posttest-Follow-up Design mit zwei Messzeitpunkten vor und einem weiteren nach Schuleintritt untersucht. In beiden Treatments wurden über eine neunmonatige Förderzeit Spiele zur Phonologischen Bewusstheit im Kindergartenalltag integriert, sowie Schrifterfahrungen in herausfordernden Alltagssituationen (z.B. Vorlesen) oder im Rollenspiel (z.B. „Literacy-center“) initiiert. Zusätzlich wurde in Treatment 2 eine Anlauttabelle als spezifisches Werkzeug aus dem Anfangsunterricht angeboten, um explizit den Buchstabe-Laut-Zusammenhang anzuregen. Die Vergleichsgruppe arbeitete ohne den Einbezug von Buchstaben mit dem Würzburger Trainingsprogramm „Hören, lauschen, lernen 1“.

Die Studie betrifft zwei Felder, die aktuell von großem Interesse sind. Das eine Feld ist der Bedarf einer Sprachfrühförderung von der Lauterfassung bis hin zur Schrift, um den späteren Anforderungen der Schule gewachsen zu sein. Das zweite Feld betrifft die sprachliche Unterstützung von Kindern im Hinblick auf ihre natürliche Neugierde und ihren Wissensdrang. Erste Ergebnisse deuten darauf hin, dass Vorschulkinder alltagsintegriert mindestens genau so effektiv hinsichtlich ihrer schriftsprachlichen Leistungsentwicklung gefördert werden können wie mit einem Trainingsprogramm. Die Anlauttabelle scheint hingegen keinen Einfluss auf den Lernzuwachs der Vorschulkinder zu haben. Auch die motivationale Entwicklung ist bei den alltagsintegriert geförderten Vorschulkinder vergleichbar mit denen, die das Förderprogramm durchliefen. Allerdings gelten diese Erkenntnisse mit der Einschränkung, dass zusätzlich die schriftsprachliche Anregungsqualität im Kindergarten hoch sein muss.

Wie Lerngelegenheiten im Alltag hergestellt werden, hängt maßgeblich von instruktionalen Unterstützungsstrategien der Erzieherin ab, sei es indirekt über die räumlich-materiale Gestaltung (z.B. Beschriftung von Regalen), sei es durch Erzieherin-Kind-Interaktionen (z.B. auf Reime in Liedern aufmerksam machen, Gespräche über den Zusammenhang gesprochene und geschriebene Sprache initiieren). Um der Heterogenität aller gerecht zu werden, kann es unter den oben genannten Bedingungen pädagogisch sinnvoll sein, Vorschulkinder alltagsintegriert, in ihrer natürlichen Umgebung auf die Schule vorzubereiten und isolierte Trainingsprogramme gezielt für Kinder mit Schwierigkeiten in der Phonologischen Bewusstheit anzubieten.

## **Soziale und emotionale Kompetenzen und Schulerfahrungen von Kindern im Modellprojekt „Bildungshaus 3 – 10“**

Das „Bildungshaus 3 - 10“ ist ein Modellprojekt des Landes Baden-Württemberg, in dessen Rahmen Kindergärten und Grundschulen neue Formen der Kooperation entwickeln und erproben. Ziel des Projekts ist es, die pädagogische Arbeit beider Einrichtungen stärker miteinander zu verzahnen und Kindern auf diese Weise eine kontinuierliche und möglichst bruchlose Bildungsbiographie zu ermöglichen. Neben einer intensivierten Kooperation zwischen Erzieherinnen und Erziehern und Lehrkräften stellen jahrgangs- und institutionsübergreifende Lern- und Spielgruppen ein weiteres Kernelement des Landesmodells dar. Schul- und Kindergartenkinder kommen dabei regelmäßig zu von den pädagogischen Fachkräften gemeinsam vorbereiteten und durchgeführten „Bildungshausangeboten“ zusammen und verbringen einen Teil ihres Bildungstags miteinander.

Beim jahrgangsübergreifenden Lernen handelt es sich um ein pädagogisches Konzept, das kontrovers diskutiert wird, in den letzten Jahrzehnten aber vor allem in der Schuleingangsphase zunehmend Verbreitung erfährt. Neben der Entlastung der Schuleingangsphase werden als Argument für die Jahrgangsmischung leichte Vorteile im sozialen Bereich angeführt, die sich empirisch belegt für Kinder in jahrgangsgemischten Gruppen ergeben (vgl. Kucharz & Wagener, 2007). Derartige Kompetenzen gewinnen im Zuge gesellschaftlicher Veränderungen immer mehr an Bedeutung für eine gelungene kindliche Entwicklung. Da es jedoch bis dato speziell an Schulen an stringenten Konzepten zur Förderung sozialer Kompetenzen fehlt, stellt sich die Frage, ob sich das „Bildungshaus 3-10“ förderlich auf die soziale Entwicklung von Kindern auswirkt (vgl. Brohm, 2009; Malti & Perren, 2008).

Mit der wissenschaftlichen Begleitung des Modellprojekts ist seit September 2008 das TransferZentrum für Neurowissenschaften und Lernen (ZNL) in Ulm betraut. Finanziert

durch das Bundesministerium für Bildung und Forschung und den Europäischen Sozialfond der Europäischen Union soll die wissenschaftliche Begleitung Erkenntnisse über Nutzen und Grenzen der intensivierten Verzahnung von Kindergarten und Grundschule an den 32 Modellstandorten hervorbringen. Um aufzeigen zu können, ob und wenn ja welche Auswirkungen die Bildungshausarbeit auf die Kinder hat, wird ein ganzheitlicher Forschungsansatz verfolgt. Innerhalb des kombinierten Querschnitt-Längsschnitt-Designs werden u. a. soziale Kompetenzen (z.B. prosoziales Verhalten) und emotionale Kompetenzen (z.B. Stressregulierung) der Kinder beforcht. Im Rahmen des mehrperspektivischer Ansatzes kommen verschiedene Beobachtungs- und Fragebögen zum Einsatz, welche die Sichtweisen von Eltern, pädagogischen Fachkräften und der Kinder selbst abbilden. Neben standardisierten Erhebungsinstrumenten wie dem „Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen“ (FEES 1-2) (Rauer & Schuck, 2004) setzt die Forschergruppe hierbei auch selbst entwickelte Verfahren wie den „Fragebogen zur sozial-emotionalen Entwicklung“ (FSEE) (Rehm, in Bearbeitung) ein.

Zur Untersuchung möglicher Auswirkungen der veränderten Form der Kooperation auf Kindergarten- und Schulkinder wird in einem kombinierten Querschnitt-Längsschnitt-Design verschiedenen Fragestellungen nachgegangen. Zwei dieser Fragestellungen sind Gegenstand des Vortrags:

Wie entwickeln sich Kinder im „Bildungshaus 3 – 10“ im Gegensatz zu Kindern in traditionellen Einrichtungen hinsichtlich sozialer und emotionaler Kompetenzen?

Wie entwickeln sich Kinder im „Bildungshaus 3 – 10“ im Gegensatz zu Kindern in traditionellen Einrichtungen hinsichtlich ihrer Einstellungen zum Selbstkonzept und zur Schule?

Die bisherigen Auswertungsergebnisse ergeben ein uneinheitliches Bild:

Aus Sicht der Eltern zeigen sich in den meisten Bereichen keine Unterschiede in der sozio-emotionalen Entwicklung zwischen Kindern in Modell- und Vergleichseinrichtungen. Hinsichtlich der sozialen und emotionalen Schulerfahrungen seitens der Kinder sind ebenfalls keine Unterschiede zwischen den Gruppen zu finden. Innerhalb der Modellgruppe variiert die Bewertung jedoch. Als Einflussfaktoren erweisen sich hier Kooperationshäufigkeit und Geschlecht. Auch Erzieherinnen und Erzieher beobachten Unterschiede zwischen Jungen und Mädchen. Des Weiteren spielen der familiäre Hintergrund und die Einrichtungsqualität eine Rolle.

Weitere Analysen der Längsschnittuntersuchungen werden es ermöglichen, zu einem klareren und differenzierteren Bild der sozial-emotionalen Entwicklung von Kindern im Modellprojekt „Bildungshaus 3 – 10“ zu gelangen.

## **Multimedia-Effekte beim Lernen und Testen: Eine Verbindung der Perspektiven**

Chair(s): Alexander Eitel (Leibniz-Institut für Wissensmedien Tübingen), Marlit Annalena Schmidt (IPN Kiel)

Diskutant/in: Alexander Renkl (Universität Freiburg)

Getreu dem GEBF Tagungsmotto „Die Perspektiven verbinden“ ist es Anliegen dieses Symposiums, Effekte der kombinierten Darstellung von Text- und Bildmaterial (Multimedia Effekte) aus verschiedenen bildungsrelevanten Perspektiven zu beleuchten – nämlich aus der Sicht des Lernens und des Testens (Lernabrufs). Ziel ist es, theoretische Annahmen und Methoden zur Untersuchung von Multimedia-Effekten aus der Instruktionspsychologie sowie der Kompetenzmessung (Testgestaltung/ Psychometrie) gewinnbringend für beide Disziplinen zu verbinden. So verfügt die Instruktionspsychologie über Theorien und empirische Befunde bezüglich der Funktion und Wirkweise von Bild-Text-Kombinationen beim Lernen. Im Fokus steht dabei herauszufinden, wie Texte und Bilder in der Lernphase gestaltet sein müssen, um Lernprozesse zu begünstigen. Methodisch wird typischerweise auf hochkontrollierte, laborexperimentelle Settings zurückgegriffen. Beispielhaft für diese Herangehensweise untersucht der erste Beitrag (T. Kühl), wie die systematische Variation räumlicher Inhalte in Text und Bild beim Lernen, die Ergebnisse in konzeptuell unterschiedlichen Wissenstest beeinflusst und wie dies mit inter-individuellen Unterschieden zusammenhängt. Zudem werden häufig Prozessmaße, wie Reaktionszeiten, Lautes Denken oder Blickbewegungsdaten herangezogen, um unterschiedliche Lernleistungen besser zu erklären. So zeigt der zweite Beitrag (A. Eitel), dass korrespondierende Hinweisreize zwischen Text und Bild im Lernmaterial zu einer besseren Lernleistung führen, wobei Blickbewegungsanalysen aufzeigen, dass diese durch eine höhere Aufmerksamkeit der Lernenden bezüglich des Bildes erklärbar ist.

Im Bereich der Kompetenzmessung mit multimedialem Material gibt es dagegen bislang kaum systematische Befunde. Die letzten beiden Beiträge des Symposiums demonstrieren daher wie Erkenntnisse aus der Instruktionspsychologie genutzt werden können, um Fragestellungen einer validen und fairen Testkonstruktion zu begegnen. Die Studien bedienen sich hierfür typisch psychometrischer Methoden in Kombination mit experimentellen Manipulationen und Prozessmaßen. So zeigt der dritte Beitrag (S. Saß), dass korrespondierende Hinweisreize zwischen Text und Bild auch in Testsituationen hilfreich sein und zu einer besseren Konstruktmessung führen können, wobei Blickbewegungsanalysen den Prozess abbilden. Schließlich zeigt der letzte Beitrag (M. Schmidt) anhand experimentell manipulierter Testaufgaben in einem Multimatrix-Design, dass Bilder einen deutlichen Einfluss auf psychometrische Parameter und Reaktionszeiten haben. So können Erkenntnisse testungsorientierter Multimediastudien langfristig wiederum der Instruktionspsychologie zugutekommen, um die Validität gemessener Lernleistungen durch verwendete Testaufgaben sicherzustellen.

Die Beiträge des Symposiums heben theoretische und methodische Stärken der beiden Forschungsfelder hervor, um Perspektiven für eine gegenseitige Befruchtung der bisweilen eher getrennt betrachteten Bereiche aufzuzeigen. Durch die im Symposium skizzierte Verbindung könnte die Forschung zur optimalen Gestaltung multimedialer Lern- und Testmaterialien, und somit nicht zuletzt die empirische Bildungsforschung im Allgemeinen, profitieren. Diesbezüglich freuen wir uns, dass Prof. Alexander Renkl einwilligte, die Diskussion des Symposiums zu übernehmen.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G131

## **Lernen mit zwei- und dreidimensionalen Visualisierungen: Der Einfluss räumlicher Textinformation und räumlichen Vorstellungsvermögens**

Theoretischer Hintergrund und Fragestellungen: Zweidimensionale (2D) und dreidimensionale (3D) Visualisierungen unterscheiden sich vor allem in der dargestellten räumlichen Information, die aus der 3D-Visualisierung unmittelbar herausgelesen werden kann (vgl. Scaife & Rogers, 1996). Eine Möglichkeit diese Informationsungleichheit von 2D- und 3D-Visualisierungen auszugleichen könnte darin bestehen, diese räumliche Information im Begleittext zu beschreiben. Während für 2D-Visualisierungen das Vorhandensein der räumlichen Textinformation lernförderlich sein sollte, könnte für 3D-Visualisierungen das Vorhandensein der räumlichen Textinformation sogar zu einem schlechteren Lernerfolg führen: Gemäß dem Redundanzprinzip (Sweller, 2005) führt die Darbietung redundanter Informationen in Text und Visualisierung zu einer unnötigen Belastung des Arbeitsgedächtnisses, was in schlechteren Lernleistungen resultieren sollte. Dies sollte speziell auf Personen mit einem geringeren räumlichen Arbeitsgedächtnis/Vorstellungsvermögen zutreffen.

Jedoch sind solche globalen Annahmen hinsichtlich des Einflusses unterschiedlicher Visualisierungsarten und Textinformationen gegebenenfalls zu unspezifisch und sollten sich vielmehr im Hinblick auf differenzierte Lernerfolgsmaße unterschiedlich stark auswirken.

Design/Methode: Der Einfluss räumlicher Textinformation und räumlichen Vorstellungsvermögens auf das Lernen mit 2D- und 3D-Visualisierungen wurde mittels eines 2x2-Between-Subject-Design mit den unabhängigen Variablen Visualisierungsart (3D vs. 2D) und Textinformation (vollständige räumliche Information vs. unvollständige räumliche Information) untersucht; zudem wurde räumliches Vorstellungsvermögen als kontinuierliche Moderatorvariable in das Design integriert.

Beim Instruktionsmaterial handelte es sich um drei verschiedenen Knoten, die anhand der jeweiligen Instruktionsbedingung erlernt werden sollten.

Als abhängige Variablen dienten das Vorführen der gelernten Knoten sowie ein verbaler und ein bildhafter Wissenstest, die jeweils spezifische Aspekte der unabhängigen Variablen adressierten. Der verbale Wissenstest war zweigeteilt: Der unspezifische verbale Wissenstest adressierte Aspekte, die in beiden Textbedingungen explizit gegeben wurden; der spezifische verbale Wissenstest adressierte Aspekte, die nur in den Textbedingungen mit vollständiger räumlicher Textinformation verbal gegeben wurde (und nonverbal in den 3D-Visualisierungen). Der Bildertest bestand in der korrekten Wiedererkennung von 2D- sowie 3D-Visualisierungen, wie sie in den jeweiligen Visualisierungsbedingungen dargestellt wurden.

Ergebnisse: Die Vorführen der Knoten gelang schneller, wenn mit 3D- statt mit 2D-Visualisierungen gelernt wurde ( $p < .05$ ) und tendenziell schneller, wenn mit vollständiger räumlicher Textinformation anstatt mit unvollständiger Textinformation gelernt wurde ( $p < .10$ ). Es gab keine Interaktion von Textinformation und Visualisierungsart und auch keinen moderierenden Einfluss vom räumlichen Vorstellungsvermögen.

Hinsichtlich des spezifischen verbalen Wissenstests zeigten erwartungsgemäß Lernende der vollständigen Textbedingungen tendenziell bessere Leistungen als Lernende der unvollständigen Textbedingungen ( $p < .10$ ), wohingegen es keinen Einfluss der Visualisierungsart und auch keine Interaktion von Textbedingung und Visualisierungsart gab. Für den unspezifischen verbalen Wissenstest gab es eine 3-fach-Interaktion von Visualisierungsart, Textinformation und räumlichem Vorstellungsvermögen ( $p < .05$ ): Die redundante Information von 3D-Visualisierungen und vollständiger räumlicher Textinformation war lernhinderlich (vgl. Sweller, 2005), allerdings nur für Probanden mit geringerem räumlichen Vorstellungsvermögen.

Entgegen der Erwartungen gab es sowohl für den 2D-Bildertest, als auch für den 3D-Bildertest keinen Haupteffekt der Visualisierungsart sowie keinen Haupteffekt von Textinformation und auch keine Interaktion von Visualisierungsart und Textinformation. Allerdings zeigte sich für beide Bildertests eine Interaktion von Visualisierungsart und räumlichem Vorstellungsvermögen ( $p < .10$  bzw.  $p < .05$ ), welche jeweils im Einklang mit der sogenannten „Ability-as-Enhancer“-Hypothese (z.B. Huk, 2006) steht. Gemäß der „Ability-as-Enhancer“-Hypothese ist ein hohes räumliches Vorstellungsvermögen notwendig, um 3D-Visualisierungen zu verstehen und um somit von ihnen profitieren zu können. Entsprechend zeigten die Ergebnisse, dass während das räumliche Vorstellungsvermögen für Lernende mit 2D-Visualisierungen keinen Einfluss auf die Leistung in den Bildertests hatte, es für Lernende mit 3D-Visualisierungen einen positiven Zusammenhang von räumlichem Vorstellungsvermögen und Leistung gab: Bei geringem räumlichen Vorstellungsvermögen führten 3D-Visualisierungen zu einem schlechteren Lernerfolg als 2D-Visualisierungen, während bei hohem räumlichen Vorstellungsvermögen 3D-Visualisierungen zu einem besseren Lernerfolg führten als 2D-Visualisierungen.

Zusammenfassend weisen diese unterschiedlichen Befundmuster auf die Wichtigkeit des Einsatzes differenzierter Lernerfolgstests hin, die unter anderem zu den jeweiligen Lernbedingungen korrespondieren sollten.

Huk, T. (2006). Who benefits from learning with 3D models? the case of spatial ability. *Journal of Computer Assisted Learning*, 22, 392-404.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies*, 45, 185-213.

Sweller, J. (2005). The Redundancy Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 159-168). New York: Cambridge University Press.



## **Müssen hervorgehobene Elemente in Texten und Bildern korrespondieren, um lernförderlich zu sein?**

Theoretischer Hintergrund und Fragestellungen: Ein gut bestätigter Effekt in der Forschung zum Lernen mit Text-Bild Kombination (Multimedia) ist, dass das Hinzufügen von Bildern zu einer rein aus Text bestehenden Lernumgebung den Lernerfolg verbessert (Multimedia-Effekt; Mayer, 2009). Bilder werden jedoch häufig nicht tief genug verarbeitet. Der Lernprozess ist textgesteuert und die Potenziale von Bildern werden wenig genutzt (Folker, Ritter, & Sichelschmidt, 2005). Um diesem Problem entgegenzuwirken, gibt es eine bewährte Methode – das Hinzufügen von korrespondierenden Hinweisreizen (z.B. farbliche Kodierungen von sich entsprechenden Elementen in Text und Bild; Ozcelik, Karakus, Kursun, & Cagiltay, 2009). Diese Methode führte in einer vorangegangenen Studie zu einer erhöhten Aufmerksamkeit für das Bild (ermittelt über Blickbewegungsanalysen; Scheiter & Eitel, 2010). Die erhöhte Bildaufmerksamkeit führte wiederum zu einer besseren Lernleistung (Mediation). Für diesen Befund gibt es zwei mögliche Erklärungsmuster. Einerseits könnten die Hinweisreize (z.B. farbliche Hervorhebungen) die Salienz des Bildes erhöht haben, so dass über eine allgemein höhere Bildaufmerksamkeit die lernförderliche Wirkung erklärbar ist (Salienzklärung). Andererseits könnte das Hervorheben der Korrespondenz zwischen Elementen in Text und Bild zu einer verbesserten Kohärenzbildung geführt haben, was die lernförderliche Wirkung ebenfalls erklären würde (Korrespondenzklärung, vgl. Seufert & Brünken, 2006). Um diese beiden Erklärungsmöglichkeiten voneinander abgrenzen zu können (Salienz vs. Korrespondenz), wurde in der vorliegenden Studie untersucht, ob das Hervorheben von Elementen auch zu höherem Lernerfolg führt, wenn diese in Text und Bild nicht korrespondieren. Nach der Salienzklärung müssten nicht-korrespondierende Hinweisreize lernförderlich sein, nach der Korrespondenzklärung nicht.

Design/Methode: Studierende (N = 78) wurden im Labor getestet und randomisiert einer von drei Gruppen zugeteilt: Ohne vs. mit nicht-korrespondierenden vs. mit korrespondierenden Hinweisreizen (Between-Subjects Design). In allen drei Gruppen wurde eine multimediale Lernumgebung (Text und statische Bilder) zur Struktur und Funktionsweise des Herz-Kreislauf-Systems dargeboten. Die Lernumgebung wurde computergestützt präsentiert und umfasste 15 Folien. In den beiden Gruppen mit Hinweisreizen wurden visuelle Hervorhebungen eingesetzt ohne dabei neuen Inhalt zu vermitteln. Diese bestanden aus Verweisen („siehe Pfeile“), kursiven Textteilen, Bildbeschriftungen sowie farblichen Kodierungen. In der Gruppe mit korrespondierenden Hinweisreizen wurde die Korrespondenz der Information in Text und Bild hervorgehoben, indem beispielweise sich entsprechende Elemente in Text und Bild dieselbe Farbkodierung aufwiesen. In der Bedingung mit nicht-korrespondierenden Hinweisreizen wurden die gleichen Hinweisreize verwendet. Lediglich die Reihenfolge, in der die Hinweisreize in den Bildern über die Folien hinweg erschienen, wurde systematisch verändert, so dass die Hinweisreize in Text und Bild nicht mehr korrespondierten.

Das Vorwissen für das Herz-Kreislauf-System wurde über einen Vorwissenstest und indirekt über die letzte Biologienote erfasst. Als abhängige Variablen dienten die Blickbewegungen der Probanden in der Lernphase (Dauer bis zur ersten Bildfixation, Relative Fixationshäufigkeit auf Bild vs. Text, Text-Bild Wechsel), sowie deren Leistungen in einem nachfolgenden Wissenstest über das Herz-Kreislauf-System (Subskalen für Erinnerung, Transfer, Text-Bild-Integration).

Ergebnisse und Interpretation: Die drei Gruppen unterschieden sich nicht hinsichtlich Vorwissen und letzter Biologienote ( $F_s < 1$ ). Entsprechend der vorangegangenen Studie führten korrespondierende Hinweisreize zu höherem Lernerfolg für die Text-Bild Integration ( $p < .05$ ), sowie zu früherer und häufigerer Bildbetrachtung ( $p_s < .05$ ). Häufigere Bildbetrachtung führte wiederum zu höherem Lernerfolg für die Text-Bild-Integration ( $p < .05$ ). Nicht-korrespondierende Hinweisreize führten weder zu einer Verbesserung in der Lernleistung ( $F_s < 1$ ) noch zu häufigerer Bildbetrachtung ( $p > .10$ ). Die Korrespondenz der hervorgehobenen Elemente war demnach relevant für höheren Lernerfolg (für Text-Bild Integration).

Diese Ergebnisse sprechen für die Korrespondenzklärung und gegen die Salienzklärung. Sie legen nahe, dass das Hervorheben der inhaltlichen Korrespondenz zu vermehrter Anstrengung nach Kohärenzbildung (höhere Bildaufmerksamkeit), und dadurch zu besserer Text-Bild Integration führte. Als mögliche praktische Implikation ließe sich ableiten, dass das Aufzeigen und Hervorheben von korrespondierenden Informationen in verschiedenen Medien bzw. externen Repräsentationen (Texte, Bilder, Animationen, Videos etc.) als lernförderliche Maßnahme fungieren kann.

Folker, S., Ritter, H., & Sichelschmidt, L. (2005). Processing and integrating multimodal material: The influence of color-coding. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), Proceedings of the 27th annual conference of the cognitive science society (pp. 690–695). Mahwah, NJ: Erlbaum.

Mayer, R. E. (2009). Multimedia learning. 2nd edition. Cambridge: Cambridge University Press.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. Computers & Education, 53, 445-453.

Scheiter, K., & Eitel, A. (2010). The effects of signals on learning from text and diagrams: How looking at diagrams earlier and more frequently improves understanding. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), Diagrammatic representation and inference – 6th International Conference, Diagrams 2010 (LNAI 6170, pp. 264-270). Heidelberg: Springer.

Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. Applied Cognitive Psychology, 20, 321-331.

## **Reduzierung der Leseanforderungen in multimedialen Testaufgaben zur naturwissenschaftlichen Kompetenzmessung durch das Signaling Prinzip**

Theoretischer Hintergrund und Fragestellungen: Testaufgaben in der naturwissenschaftlichen Kompetenzmessung sind typischerweise so aufgebaut, dass im Itemstamm eine naturwissenschaftliche Situation bzw. Problemstellung verdeutlicht und gegebenenfalls durch Verwendung ergänzender Texte oder Visualisierungen spezifiziert wird. Um die Testaufgabe erfolgreich zu lösen, müssen zunächst alle gegebenen Informationen im Stimulus in Form von Text und Bild verarbeitet und verstanden werden. Wie sich dabei verschiedene Präsentationsformen von Text und Bild auf den Verstehensprozess der Aufgaben auswirken, ist bislang wenig erforscht. Aus der Lehr-Lernforschung im Bereich des multimedialen Lernens ist jedoch bekannt, dass beispielsweise Integrationshilfen in Form von farblichen Hervorhebungen korrespondierender Informationen in Text und Bild ( Signaling ) den Aufbau eines mentalen Modells unterstützen und das Arbeitsgedächtnis entlasten können. So konnte eine Studie von Seufert (2003) zeigen, dass Lernende einen deutlichen Lernzuwachs erzielten, wenn die Integration von Informationen aus multiplen Quellen (Text und Bild) in Lernmaterialien instruktional unterstützt wurde. Weitere Studien liefern Hinweise darauf, dass durch instruktionale Maßnahmen zur Unterstützung der Integration multipler Quellen unnötige Belastungen des Arbeitsgedächtnis ( Extraneous Cognitive Load ; Sweller, 2005) reduziert werden können (z.B. Ciernak, Scheiter, & Gerjets, 2009). Es wird angenommen, dass sich der Anteil zur verarbeitender Informationen durch die Unterstützung von Selektions- und Integrationsprozessen reduziert und somit das Arbeitsgedächtnis entlastet wird (Florax & Plötzner, 2010). Die vorliegende Studie untersucht ausgehend von diesen Befunden aus der Lehr-Lern-Forschung, ob Mechanismen zur Unterstützung der Integration von Text- und Bildrepräsentationen auch in Testaufgaben zum Tragen kommen und durch welche weiteren Aufgaben- und Personenmerkmale ( Lesekompetenz und Vorwissen ) diese möglicherweise moderiert werden.

Methode und Design: Die Fragestellungen werden durch psychometrische und prozessorientierte Methoden in zwei aufeinander aufbauenden Studien beantwortet. In einer Paper-Pencil Studie ( N = 91 Schülerinnen und Schüler im 6. Schuljahr) wurde zunächst in einem 2x2 Design eine experimentelle Manipulation von 22 naturwissenschaftlichen Multiple-Choice Aufgaben bezüglich der Faktoren Informationsgehalt zwischen Text und Bild (kongruent vs. identisch) und der Verwendung von Integrationshilfen (Hilfen vs. keine Hilfen) vorgenommen. Zudem wurden unter anderem das domänenspezifische Vorwissen und die Lesekompetenz erhoben. Als abhängige Variable wurde die Anzahl korrekt gelöster Aufgaben herangezogen. Eine anschließende Eyetracking Studie ( N = 30 Schülerinnen und Schüler im 6. Schuljahr) mit äquivalentem Testdesign soll die zugrunde liegenden Prozesse der Integration von Text- und Bildinformationen durch die Analyse von Fixationszeiten und Transitionen auf und zwischen den Bild- und Textinformationen näher beleuchten.

Ergebnisse und Diskussion: Die Ergebnisse der Paper-Pencil Studie zeigen, dass Testaufgaben mit Integrationshilfen erwartungsgemäß leichter zu lösen waren als Testaufgaben ohne Integrationshilfen zwischen Text und Bild. Der Effekt zeigte sich deutlicher bei Aufgaben, in denen Text und Bild kongruente Informationen enthielten als bei Aufgaben mit identischem Inhalt in Text und Bild. Zusätzlich ließ sich in einer Moderatoranalyse ein Einfluss der Lesefähigkeit auf die Wirkung der Integrationshilfen identifizieren. So profitierten vor allem leseschwache Schülerinnen und Schüler von Integrationshilfen in naturwissenschaftlichen Testaufgaben und lösten Aufgaben mit Hilfen häufiger richtig als Aufgaben ohne Integrationshilfen. Die Eyetrackingdaten der zweiten Teilstudie werden aktuell erhoben, Ergebnisse liegen bis zur GEBF 2014 vor. Insbesondere die Veränderung des Bearbeitungsprozesses der Testaufgaben (z. B. Betrachtungsdauer von Bild und Textinformation) durch Integrationshilfen steht im Vordergrund der geplanten Blickbewegungsanalysen, um zugrunde liegende Mechanismen der sich in den Paper-Pencil Daten abzeichnenden Effekte besser zu verstehen. Die beiden Studien werden vor dem Hintergrund einer validen und fairen Gestaltung von Testaufgaben durch den Einsatz multimedialer Gestaltungstechniken diskutiert. Es wird zudem am Beispiel der vorgestellten Studien demonstriert, ob und wie es gelingen kann, bekannte kognitionspsychologische Theorien und Prinzipien der Bild- und Textverarbeitung aus der Instruktionspsychologie auf den Bereich des Testens zu übertragen.

Ciernak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25, 315-324.

Florax, M., & Plötzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labeling, and spatial proximity. *Learning and Instruction*, 20, 216-224.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13, 227-237.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). New York, NY: Cambridge University Press.

## **Der Multimediaeffekt beim Testen: Bilder beeinflussen psychometrische Parameter von Multiple-Choice Aufgaben**

Theoretischer Hintergrund und Fragestellungen: Im Large-Scale Assessment wie z. B. PISA (OECD, 2013) werden Bilder häufig eingesetzt, dennoch existieren bislang kaum systematische Befunde über Wirkungen des Hinzufügens von Bildern in Testaufgaben. Erste Studien liefern bereits Hinweise auf differenzielle Bildeffekte in Tests (Hartmann, 2012; Saß, Wittwer, Senkbeil, & Köller 2012) und legen zusammen mit vielen Multimediabefunden im Bereich des Lernens mit Text und Bild (vgl. Cognitive Theory of Multimedia Learning, Mayer, 2005) die Annahme nahe, dass Bilder auch in Testsituationen einen nicht zu vernachlässigenden Einfluss auf kognitive Verarbeitungsprozesse, sowie psychometrische Parameter haben könnten. Um Bildeffekte systematisch anhand einer großen Anzahl an Aufgaben zu untersuchen, wurde hier eine experimentelle Manipulation von Testaufgaben vorgenommen, um den Einfluss der Ergänzung eines veranschaulichenden Bildes im Itemstamm auf die psychometrischen Parameter (1) Itemschwierigkeit und (2) Itemdiskrimination zu untersuchen und die Generalisierbarkeit der Effekte über viele verschiedene Testaufgaben abzuschätzen. Darüber hinaus wurden (3) Lesefähigkeit und allgemeine kognitive Fähigkeiten als mögliche moderierende Faktoren der Bilderwirkung untersucht.

Design und Methode: Eine Manipulation von 48 naturwissenschaftlichen Aufgaben im Multiple-Choice Format mit kurzen Texten erfolgte, indem ein aufgabenrelevantes Bild zum Itemstamm hinzugefügt wurde oder nicht (Bild vs. kein Bild). Das Bild wurde jeweils so gewählt, dass es die Textinformation veranschaulicht, jedoch keine zusätzlichen lösungsrelevanten Informationen enthält. Die Aufgaben erforderten einen Transfer von Wissen auf Anwendungssituationen und waren abgesehen von der experimentellen Manipulation vollkommen parallel. Zur Sicherstellung externer Validität und praktischer Relevanz der Ergebnisse, erfolgte die Itemkonstruktion anhand des TIMSS-Rahmenkonzepts (Mullis et al., 2007). Ein Multi-Matrix-Design wurde eingesetzt, um Testparameter für die hohe Anzahl an Testitems durch den Einsatz von IRT-Modellen schätzen zu können und vergleichbare Parameter für Items mit und ohne Bild für jede Person und jedes Item zu erhalten. Die Stichprobe umfasst N = 304 Schülerinnen und Schüler der 5. und 6. Klassen verschiedener Schulformen.

Erste Ergebnisse und Diskussion: (1) Erste Analysen zeigen erwartungsgemäß einen signifikanten Multimediaeffekt ( $p = .001$ ;  $d = .24$ ) bezüglich der Itemschwierigkeit. So wurden 71% der Aufgaben durch ein Bild leichter. Dies könnte entweder ein Hinweis darauf sein, dass Bilder halfen, die Textinformationen besser zu verstehen oder dass vorhandenes Wissen durch das Bild einfacher auf die dargestellte Situation anwendbar war. 29% der Aufgaben wurden durch die Ergänzung eines Bildes allerdings auch schwieriger. (2) Für die Itemdiskrimination wurde kein globaler Multimediaeffekt signifikant, wobei sich auf Einzelitemebene große Veränderungen in der Diskrimination der Aufgaben durch die Ergänzung eines Bildes ergaben. (3) Die Korrelationen zwischen IRT-geschätzten Testwerten der Personen und den allgemeinen kognitiven Fähigkeiten ergab keinen Unterschied für Aufgaben mit ( $r = .43$ ) und ohne Bild ( $r = .42$ ), wohingegen Korrelationen von reinen Textaufgaben mit der Lesefähigkeit leicht höher ausfielen ( $r = .24$ ) als für Aufgaben mit Bild ( $r = .19$ ). Dies legt die Schlussfolgerung nahe, dass die Ergänzung eines Bildes die Konstruktmessung nicht grundlegend verändert, aber die Leseanforderung durch das Bild zumindest tendenziell gemindert wird.

Zusammengefasst zeigte sich kein vollständig generalisierbarer Multimediaeffekt für psychometrische Parameter sämtlicher Testaufgaben, vielmehr waren differenzielle Effekte für verschiedene Items und Parameter zu verzeichnen. Die zukünftige Forschung sollte sich daher noch stärker auf die Eigenschaften der Aufgaben und die Funktion einzelner Bilder im Lösungsprozess konzentrieren, um die Richtung und das Ausmaß psychometrischer Veränderungen durch multimediale Ergänzungen des Testmaterials besser vorherzusagen und mit kognitiven Operationen im Lösungsprozess in Verbindung zu bringen. Zusätzliche Analysen der vorliegenden Daten, die differenziertere Aussagen über die Effekte einzelner Bilder erlauben, werden für die GEBF 2014 vorliegen. Darüber hinaus präsentieren wir eine äquivalente, auf dieser ersten Studie aufbauende Studie mit einer computergestützten Erfassung durch Laptops (N = 398 Schülerinnen und Schüler), die neben den Effekten auf psychometrische Parameter auch Aussagen über Reaktionszeiteffekte durch die Ergänzung von Bildern zu den Testaufgaben erlaubt.

Hartmann, S. (2012). Die Rolle von Leseverständnis und Lesegeschwindigkeit beim Zustandekommen der Leistungen in schriftlichen Tests zur Erfassung naturwissenschaftlicher Kompetenz. Elektronische Dissertation, Universität Duisburg-Essen.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook Of Multimedia Learning* (pp. 31–48). New York: Cambridge University Press.

Mullis, I. V. S., Martin, M. O., Rueddick, G. J., O'Sullivan, C. Y., Arora, A., & Eberber, E. (2007). *TIMSS assessment framework 2007*. Chestnut Hill: Boston College.

OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing.

Saß, S., Wittwer, J., Senkbeil, M., & Köller, O. (2012). Pictures in Test Items: Effects on Response Time and Response Correctness. *Applied Cognitive Psychology*, 26, 70–81.

Knut Schwippert (Universität Hamburg)

## **Zum Einfluss von Sprachkompetenz im Fachunterricht – Entwicklung und empirische Erprobung fachsprachlicher C-Tests**

**GEBF** TAGUNG  
2014

F08

Symposium

Chair(s): Knut Schwippert (Universität Hamburg), Dominik Leiss (Leuphana Universität Lüneburg)

Diskutant/in: Johannes Hartig (DIPF)

Aus nationalen und internationalen Vergleichsstudien ist bekannt, dass insbesondere Schülerinnen und Schüler mit einem Migrationshintergrund erhebliche Probleme bei der Bearbeitung komplexerer Aufgabenstellungen haben (vgl. u.a. PISA, 2000). Insbesondere bei realitätsbezogenen Aufgaben bestimmt ein komplexes Wechselspiel von Merkmalen der – in der Regel schriftlich und in einem spezifischen Instruktionkontext – dargelegten Realsituation (Verschaffel et al., 2000) und von intrapersonellen Aspekten (Artelt et al., 2001) das mentale Modell der Aufgabenstellung in der Schülerkognition. Dabei scheint der Lesekompetenz bei den intrapersonellen Aspekten eine herausragende Rolle zuzukommen, was sich empirisch bei zahlreichen Studien in einer hohen Korrelation der Lesekompetenz mit der jeweiligen anwendungsorientierten Fachkompetenz widerspiegelt (vgl. u.a. Vilenius-Tuochimaa et al., 2008). Auch wenn zu diesem komplexen Zusammenhang eine Reihe von deduktiven Theorien existieren, so stellt die empirische Klärung der relevanten Prozessvariablen und deren Zusammenhänge sowohl in der nationalen als auch der internationalen Forschung ein Desiderat dar (siehe Überblick in Kintsch, 2004) und wird in einem von Wissenschaftlerinnen und Wissenschaftlern unterschiedlicher Disziplinen in einem überregionalen Projektverbund untersucht.

Konzeptionelle Grundlage dieses Forschungsverbundes ist die systematische Verbindung empirischer Bildungsforschung mit und zwischen den Fachdidaktiken Deutsch, Englisch, Mathematik, Physik, und Sport. Dabei stehen insbesondere die Zusammenhänge zwischen sprachlicher Expertise und Fachleistungen im Mittelpunkt. Nach einer angemessenen Dekomposition sprachlicher und fachlicher Kompetenzen sollen perspektivisch professionelle Handlungskompetenzen von Lehrkräften im Umgang mit sprachlicher Heterogenität im Fachunterricht in den Blick genommen werden, um diese langfristig im Projekt auf ihre Wirksamkeit zu prüfen und in Handlungsempfehlungen für Lehrkräfte zu übersetzen.

Zur Erfassung von sprachlichen und fachlichen Kompetenzen werden sprachlich systematisch variierte Fachaufgaben entwickelt und anhand mehrdimensionaler Analysemodelle untersucht. Als weitere Kontrollvariable sollen perspektivisch neben persönlichen und sozialen Hintergrundmerkmalen der gut administrierbare Sprachtest (C-Test) angewendet werden. Um hiermit jedoch nicht ausschließlich allgemein-sprachliche Komponenten zu erfassen, wurden in verschiedenen Vorstudien (unter Kontrolle von allgemeinen kognitiven Fähigkeiten) Varianten des C-Tests als Fach-Sprachen-Test durch die Zusammenarbeit von Vertretern verschiedener Fachdidaktiken und der (quantitativen) empirischen Schul- und Unterrichtsforschung entwickelt. Neben der Konstruktion von fachspezifischen Varianten stehen auch Fragen zu unterschiedlichen Konstruktionsprinzipien der C-Test im Rahmen der inzwischen abgeschlossenen Pilotstudien vergleichend im Fokus. Die fachspezifisch durchaus differenziert verlaufenen Testentwicklungen werden Gegenstand des Symposiums sein und in Einzelbeiträgen aus Perspektive der Testentwicklung, der Deutsch-, Mathematik, Sport & Physik- und der Englisch-Didaktik vorgestellt um schließlich von einem Experten der pädagogisch-psychologischen Diagnostik diskutiert zu werden.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G135

Artelt, C. et al. (2001): Lesekompetenz: Testkonzeption und Ergebnisse. In: Baumert, J. et al.: PISA 2000 - Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, 69-137.

Verschaffel, L. et al. (2000): Making sense of word problems. Lisse: Swets and Zeitlinger.

Vilenius-Tuohimaa, P., Aunola, K., Nurmi, J. (2008): The association between mathematical word problems and reading comprehension. Educational Psychologist 28 (4), 409-426.

Zum Einfluss von Sprachkompetenz im Fachunterricht

2

## ***C-Tests zur Messung global-sprachlicher Voraussetzungen für die adaptive Schreibunterstützung in SILLK***

Relevanter Kontext: Das Projekt SILLK (Schreibinterventionen in heterogenen Lern-Lehr-Kontexten) fokussiert den Erwerb der Schriftsprachlichkeit im Deutschen, die für den allgemeinen Lernerfolg sowie eigenständige Lebensführung als Schlüsselkompetenz gilt. Es handelt sich dabei – im Sinne von funktional-pragmatischen Ansätzen – um eine komplexe Problemlösefähigkeit, die zwar etappenweise bis zum Ende der regulären Schulausbildung angeeignet werden sollte, faktisch aber für einen nicht zu unterschätzenden Teil der Schüler\_innen unerreichbar bleibt (vgl. Neumann & Lehmann, 2008). In dem Projekt werden Maßnahmen an Grahams SRSD-Modell (vgl. Graham & Harris, 2003) angelehnte Scaffolding-Interventionen auf der Mikroebene erprobt. Eine wichtige Voraussetzung ist dabei die Ermittlung der global-sprachlichen Eingangsvoraussetzungen und der im Schreiben selbst. Dazu dienen einerseits C-Test, andererseits Pretests für die einzelnen Textformen.

Zielsetzung: Im Zentrum steht dabei ein hohes Interesse an den kognitiven Prozessen, die die Schreibenden bei der Auseinandersetzung mit verschiedenen kommunikativen Praxen (Fiehler et al., 2004, Bachmann & Becker-Mrotzek, 2010) vollziehen. Es gilt zu prüfen, ob und in welcher Form C-Tests dazu Hinweise liefern können. Können daraus Informationen gewonnen werden, interaktive, kollaborative Verfahren im Zentrum (Rijlaarsdam et al., 2008) eines prozessorientierten funktionalen Schreiberwerbs zu stützen? Es soll die Frage beantwortet werden, ob gemeinsames sprachliches Aushandeln in der Peergroup den Erwerb registeradäquaten Schreibens, gemessen an den Schreibprodukten verbessert.

Untersuchungsmethoden und erste Pilotierungsergebnisse: Die globale Sprachfähigkeit wurde anhand eines C-Tests erfasst, der in den Teiltextrn Kompetenzen für die zu schreibenden Textformen fokussierte. Eine differenziertere Erfassung der global-sprachlichen Fähigkeiten über einen Richtig-Falsch-Wert hinaus konnte nicht modelliert werden, die einzelnen Teiltests korrelieren im mittleren Bereich. Effekte auf die Schreibintervention wurden in einem Pre-Posttest-Design erfasst. Die Leistungsveränderung wurde anhand kodierter und dann skaliertes Schreibprodukte erfasst. Die Ergebnisse einer Pilotierungsstudie mit N=55 Schüler\_innen siebenter Klassen zeigen signifikante Interventionseffekte bildungssprachlich relevanter Textmuster. Es zeigt sich dabei, dass der Einfluss der global-sprachlichen Eingangsvoraussetzungen innerhalb der Intervention nivelliert wird.

Bachmann, T. & Becker-Mrotzek (2010). Schreibaufgaben situieren und profilieren. In T. Pohl & T. Steinhoff (Hrsg.), Textformen als Lernformen (S. 191-220). Köln: Gilles & Francke.

Fiehler, R., Barden, B. & Eltersmann, M. (2004). Eigenschaften gesprochener Sprache. Tübingen: Gunther Narr.

Graham S./Harris K. (2003): Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In: Swanson L. H./Harris K./Graham S. (Eds.): Handbook of learning disabilities. New York: Guilford, pp. 323–344.

Neumann, A. & Lehmann, R. H. (2008): Schreiben Deutsch. In: DESI-Konsortium (Hrsg.): Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim/Basel: Beltz, S. 89-103.

Rijlaarsdam, G.C.W. et al. (2008): Observation of peers in learning to write. Practice and research. Journal of Writing Research, 1 (1), 53-83.



## Was misst der C-Test?

### Rasch-Skalierung und kriteriumsorientierte Testwertinterpretation als Methoden der Konstruktvalidierung

Der C-Test ist hinsichtlich Testerstellung, -durchführung und -auswertung ein sehr einfach zu handhabendes Forschungsinstrument, das sich durch hohe Messgenauigkeit, Zuverlässigkeit und Robustheit auszeichnet (u.a. Grotjahn, 2006). Daher hat er in der Fremdsprachenforschung weite Verbreitung gefunden, ist – insbesondere in den letzten Jahren – in wichtigen large-scale-Studien wie z.B. DESI zum Einsatz gekommen und hat sich als Diagnostik- und Einstufungsinstrument bei öffentlichen und privaten Anbietern von Sprachkursen etabliert. Umso drängender ist es daher die nach wie vor allenfalls teilweise beantwortete Frage zu klären, was der C-Test erfasst. Zwar hat es in der Zwischenzeit, z.B. durch die Arbeiten von Stemmer (1991) und Sigott (2004) hier Fortschritte gegeben. Es gilt aber nach wie vor: „Dissens besteht in der Forschungsliteratur darüber, was ein derartiger Test eigentlich misst, - testtheoretisch gesprochen: was die Konstruktvalidität dieses integrativen Testformats ausmacht.“ (Vockrodtz-Scholz & Zydati, 2010, S. 5)

Die Klärung dieser Frage wird im Vortrag in drei Schritten unternommen. Zunächst wird der C-Test selbst vorgestellt und bzgl. seiner Konstruktionsprinzipien (z.B. verminderte Redundanz und Tilgungsprinzip) diskutiert. Außerdem wird theoretisch erörtert, in welchem Verhältnis der C-Test zu unterschiedlichen Konstrukten sprachlicher Fähigkeiten, wie z.B. dem Konzept der kommunikativen Sprachfähigkeit (z.B. Bachman, 1990) oder der allgemeinen Sprachfähigkeit steht. Insbesondere wird zu diskutieren sein, in welchem Verhältnis diese gewissermaßen ganzheitlichen Konstrukte zu den in rezenten Untersuchungen (z.B. DESI) und den Bildungsstandards verwendeten funktional-pragmatischen Kompetenzmodellen stehen, die sprachliche Teilfertigkeiten (Leseverstehen, Hörverstehen, Sprechen, Schreiben) ausweisen und in je eigenen Tests erfassen.

Anschließend wird anhand originaler C-Tests aus einem Projekt zu Kooperativem Lernen im Englischunterricht erläutert, wie der aufgeworfenen Frage danach, was der C-Test misst, auf zwei Wegen nachgegangen wurde. Zunächst wurden die C-Tests dazu auf der Basis der Item-Response-Theorie mittels Rasch-Skalierung (1PL-Modell) ausgewertet. Dazu müssen die C-Test-Superitems als polytome Items behandelt werden. Dazu werden alle Lücken als einzelne Items aufgefasst. Die dazu notwendige Annahme der lokalen stochastischen Unabhängigkeit der Lücken ist zwar theoretische angezweifelt worden, diese Zweifel konnten aber nicht empirisch erhärtet werden (z.B. Harsch & Hartig, 2010). Die Rasch-Skalierung liefert die Verteilung der Personen- und Itemparameter anhand derer sieben Niveaus bei der inhaltlichen Beschreibung der Itemschwierigkeit bestimmt werden konnten.

Auf dieser Basis wurde eine untere und obere Schwelle identifiziert, deren Items einer kriteriumsorientierten Testwertinterpretation unterzogen wurden. Dabei wurden die jeweiligen Lücken hinsichtlich der für ihre korrekte Lösung notwendigen kognitiven Operationen der Testteilnehmer analysiert. Dabei zeigte sich, dass die Lösung der Items der unteren Schwelle überwiegend lexikalisches Wissen, z.B. die Kenntnis einfacher Strukturwörter, Grundverben oder Substantive erfordert und dabei nahezu ausschließlich Informationen verarbeitet werden müssen, die in unmittelbarer Nähe des Zielwortes zu finden waren und daher kein strategisches Wissen erfordern. Im Gegensatz dazu zeigte sich, dass die Lösung der Items der oberen Schwelle nicht nur lexikalisches Wissen auf fortgeschrittenem Niveau, sondern auch eine Vertrautheit mit der englischen Basisgrammatik im Bereich der Zeiten, sowie ein globales Sinnverstehen auf Textebene, das durch Interpolation und Inferenzierung auf die Lücke zu beziehen war, erforderte. Diese Ergebnisse erlauben es erstens, die Schwierigkeit einzelner C-Test-Lücken quantitativ und qualitativ zu bestimmen und zweitens die dazu notwendigen deklarativen und prozeduralen Wissensbestände der Testteilnehmer/innen zu beschreiben. Die Ergebnisse zeigen, dass lediglich bei den einfachen Items eine Beschränkung auf isolierte Teilfertigkeiten gegeben, bei schwierigeren Items aber ein Kombination sehr unterschiedlicher Teilfertigkeiten notwendig ist.

Bachmann, L. F. (1990): Fundamental Considerations in Language Testing. Oxford: Oxford University Press.

Bockrodt-Scholz, V. & Zydati W. (2010): Sprachproduktive Faktoren und die Konstruktvalidität von C-Tests: Kompetenzniveaus und Fehlerquotient in textsortengebundenen Schreibaufgaben. In:

Grotjahn, Rüdiger (Hrsg.): Der C-Test. Beiträge aus der aktuellen Forschung. Frankfurt a.M.: Peter Lang.

Grotjahn, R. (Hrsg.) (2006): Der C-Test: Theorie, Empirie, Anwendungen. Frankfurt a.M.: Peter Lang.

Harsch, C. & Hartig, J. (2010): Empirische und inhaltliche Analyse lokaler Abhängigkeiten im C-Test. In: Grotjahn, R. (Hrsg.): Der C-Test Beiträge aus der aktuellen Forschung. Frankfurt a.M.: Peter Lang.

Sigott, G. (2004): Towards Identifying the C-Test Construct. Frankfurt a.M.: Peter Lang.

Stemmer, B. (1991): What's on a c-test taker's mind? Mental processes in c-test taking. Bochum: Universitätsverlag Dr. N. Brockmeyer.

## **Verstehen des Verstehens: Zum Einfluss (fach)sprachlicher Kompetenzen auf das Situationsmodell bei kompetenzorientierten Mathematikaufgaben**

Theoretischer Hintergrund: Durch die vermehrte Bearbeitung von Problemstellungen mit größeren Textanteilen im Mathematikunterricht ist das Textverständnis verstärkt zu einer zentralen Voraussetzung für die erfolgreiche Aufgabebearbeitung geworden (vgl. Duarte et al., 2011). Insbesondere bei sogenannten mathematischen Modellierungsaufgaben, also solchen Aufgaben die eine reale Problemstellung durch mathematische Modelle zu beantworten versuchen, zeigen empirische Studien, dass für das Bilden eines adäquaten Situationsmodells (Reusser, 1989) die Lesekompetenz von herausragender Relevanz ist (Leiss et al., 2010). Unklar ist dabei allerdings welche Faktoren genau bedeutsam für das Verstehen eines mathematischen Sachverhalts bzw. der Problemstellung sind. Ansätze liefert u.a. Cummins (2008), der die Konstrukte basic interpersonal communication skills (bics) und cognitive academic language proficiency (calp) unterscheidet. Während das Erstgenannte Sprachfähigkeiten beschreiben, die man im alltäglichen Leben benötigt, umfasst calp solche sprachlichen Fähigkeiten, die kognitiv anspruchsvoller sind und eher zur Kommunikation in fachspezifischen Kontexten (wie z.B. Schule) notwendig sind. Bezogen auf die Bearbeitung von Mathematikaufgaben, die auf Problemstellungen in alltägliche Situationen basieren, stellt sich somit die Frage inwieweit die Aspekte allgemeine Sprachkompetenz, fachsprachliche Kompetenz, Fachwissen und allgemeine kognitive Fähigkeiten den Verstehensprozess beeinflussen.

Fragestellungen: (1) Inwiefern handelt es sich bei der fachsprachlichen Kompetenz um ein mithilfe von spezifischen C-Tests messbares Konstrukt und inwiefern ist dieses trennbar von ebenfalls mit einem C-Test erfassten allgemeinen Sprachfähigkeiten?

(2) Wird das Verständnis einer Mathematikaufgabe eher durch (fach)sprachliche Kompetenzen oder eher durch Fachwissen und die allgemeine kognitive Leistungsfähigkeit beeinflusst?

(3) Lassen sich in diesem Kontext (mathematische) Inhaltsbereiche mit spezifischen Wirkungszusammenhängen identifizieren?

Methode: Im Rahmen der hier vorgestellten SITRE-Studie nahmen 190 Schülerinnen und Schüler des 7. Jahrgangs einer Hauptschule und einer integrierten Gesamtschule an einer 90minütigen Testung teil. Dabei umfasste der Test neben der Erfassung von Hintergrundvariablen (z.B. Alter, Geschlecht etc.) die folgenden Komponenten:

- DEMAT 6+ (zur Feststellung der allgemeinen mathematischen Leistungsstärke)
- Verständnistest Mathematik (zur Diagnose der Fähigkeit ein adäquates Situationsmodell zu bilden)
- Fachsprachliche C-Test (zur Messung fachsprachlicher Kompetenzen)
- Allgemeiner C-Test (zur Messung allgemeiner Sprachkompetenzen)
- Nicht-sprachlicher Teil des BEFKI (zur Messung der allgemeinen kognitiven Leistungsfähigkeit)

Eine Besonderheit bestand darin, dass für die Konstruktion des fachsprachlichen C-Tests unterschiedliche Vorgehensweisen gewählt worden, so dass dieser im rotierten Testdesign zwischen den Probanden variiert wurde und damit eine eigenständige methodologische Studie darstellt. Für die Auswertung wurde beim DEMAT 6+ auf die vorhandenen Daten der Normierungsstichprobe zurückgegriffen (Götz et al., 2013). Die Auswertung der Daten der verschiedenen Varianten des C-Tests werden im Rahmen des Symposiums vorgestellt -, wobei für eine auf probabilistische Testtheorien beruhende Auswertung der C-Tests insbesondere die Ergebnisse von Harsch & Hartig (2010) gesprochen haben, die darauf hinweisen, dass die einzelnen Lücken im C-Test als voneinander unabhängige Items betrachtet werden können.

Ergebnisse: In einem ersten Zugriff auf die verschiedenen C-Test-Varianten zeigte sich im OPL Modell ein zunächst als generalisierte Sprachkompetenz interpretierbares latentes Konstrukt. In weitergehenden Untersuchungen werden unterschiedliche Skalierungs-Modelle vorgestellt, die sowohl die verschiedenen Varianten des C-Tests (unterschiedlichen Tilgungsalgorithmen) als auch auf Einzel-Item-Ebene Facetten einer eher allgemeineren sprachlichen und einer eher (mathematisch-) fachlichen Dimensionen mittels mehrdimensionaler Skalierungsmodelle in den Blick nehmen.

Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In H. Hornberger (Hrsg.). Encyclopedia of Language and Education, Bd. 2, New York: Springer, 487-499.

Duarte, J., Gogolin, I. & Kaiser, G. (2011). Sprachlich bedingte Schwierigkeiten von mehrsprachigen Schülerinnen und Schülern bei Textaufgaben. In Özdil, E. & Prediger, S. (Hrsg.). Mathematiklernen unter Bedingungen der Mehrsprachigkeit. Stand und Perspektive der Forschung und Entwicklung in Deutschland. Münster: Waxmann, 35 – 54.

Götz, L., Lingel, K., Schneider, W. (2013). DEMAT 6+ - Deutscher Matheamtiktest für sechste Klassen. Göttingen: Hogrefe.

Harsch, C. & Hartig, J. (2010). Empirische und inhaltliche Analyse lokaler Abhängigkeiten im C-Test. In Grotjahn, R. (Hrsg.). Der C-Test: Beiträge aus der aktuellen Forschung. Frankfurt am Main: Lang, 193-204.

Leiss, D., Schukajlow, S., Blum, W., Messner, R. & Pekrun, R. (2010). The role of the situation model in mathematical modeling – Task analyses, student competencies and teacher interventions. JMD, 31(1), 119 – 141.

Reusser, K. (1989). Vom Text zur Situation zur Gleichung - Kognitive Simulation von Sprachverständnis und Mathematisierung beim Lösen von Textaufgaben. Bern.

## **Allgemeinsprachliche und fachsprachliche Fähigkeiten als Bedingungsfaktoren für Lernerfolg in Physik und Sport**

Der Umgang mit Fachsprache ist ein zentraler Aspekt im schulischen Unterricht. Bereits in der Sekundarstufe I werden im Physikunterricht hohe Anforderungen an die sprachlichen Fähigkeiten von Schülerinnen und Schülern gestellt, um beispielsweise Lehrbuchtexte zu lesen, schriftliche Aufgabenstellungen zu verstehen oder Ergebnisse aus Experimenten zu dokumentieren. Auch der Sportunterricht kommt nicht ohne fachsprachliche Begriffe aus, wenn es etwa darum geht, motorische Bewegungsabläufe zu verbalisieren oder Körpererfahrungen zu diskutieren. Inwieweit aber das Beherrschen von Fachsprache neben allgemein-sprachlichen Fähigkeiten ein wichtiger Bedingungsfaktor für den Lernerfolg im jeweiligen Unterrichtsfach ist (vgl. Rincke, 2010), ist bislang weitgehend ungeklärt und wird in dieser Studie für die Fächer Physik und Sport untersucht.

Die Stichprobe besteht aus N = 119 Schülerinnen und Schülern aus überwiegend sportbetonten Schulklassen der 6. und 7. Jahrgangsstufe. Die allgemeinsprachlichen Fähigkeiten und die fachsprachlichen Fähigkeiten in Physik und Sport werden anhand von neu konstruierten C-Tests erfasst.

C-Tests sind dabei globale Instrumente zur Bestimmung sprachlicher Kompetenzen (Baur, Grotjahn, & Spaettmann, 2006). Sie bestehen in der Regel aus mehreren kurzen Textblöcken, in denen Teile von ca. 15-25 Wörtern getilgt wurden. Die Testleistung besteht in der Rekonstruktion der getilgten Wortteile. Dies gelingt auf der Grundlage syntaktischen, semantischen und Kontext-Wissens. C-Tests basieren auf dem Prinzip der Reduktion von Redundanzen: Die Bedeutungen sprachlicher Mitteilungen werden von Hörern/Lesern erschlossen, indem sie möglichst viele Aspekte berücksichtigen, die zum Verstehen einer Mitteilung beitragen (z. B. Wortbedeutung, syntaktische Merkmale, Kollokationen). Ein Merkmal eines kompetenten Sprachnutzers besteht in der Fähigkeit, eine sprachliche Mitteilung selbst unter Bedingungen reduzierter Redundanz (Tilgung von Wortteilen) zu verstehen. Ein Indikator dieser Kompetenz ist die Fähigkeit einer Testperson, die Tilgungen im C-Test korrekt zu füllen.

Die C-Tests für Physik und Sport in dieser Studie umfassten jeweils zwei Teiltests à 15 Items, der allgemeinsprachliche C-Test bestand aus zwei Teiltests à 25 Items. Die Testreliabilitäten lagen auf hohem Niveau (Physik:  $r = 0.887$ , Sport:  $r = 0.880$ , Allgemein:  $r = 0.937$ ). Über einen Schülerfragebogen wurde neben soziodemographischen Angaben der Lernerfolg im Fach über die Schulnoten erhoben.

Die C-Tests wurden auf Basis der Item-Response-Theorie ausgewertet. Dabei wurde angenommen, dass die Annahme der lokalen stochastischen Unabhängigkeit nicht verletzt ist, da nach Harsch und Hartig (2010) in den von ihnen untersuchten C-Tests, keine nennenswerten und nicht-trivialen lokalen stochastischen Abhängigkeiten zwischen den Tilgungen (Items) bestehen.

Die Auswertung der C-Tests zeigte, dass alle drei C-Tests unterschiedliche sprachliche Teilfähigkeiten messen. Die Korrelationen zwischen den fachsprachlichen Fähigkeiten in Physik und Sport lagen auf mittlerem Niveau ( $r = .48$ ) ebenso wie die Zusammenhänge der beiden Fachsprachentests mit dem allgemeinsprachlichen C-Test (Physik:  $r = .48$ , Sport:  $r = .47$ ). Signifikante Zusammenhänge zeigten sich erwartungskonform mit der Deutschnote (C-Test-Physik:  $r = .37$ , C-Test-Sport:  $r = .34$ , C-Test-Allg.:  $r = .47$ ), in etwas geringerem Maße mit der Physiknote (C-Test-Physik:  $r = .30$ , C-Test-Sport:  $r = .33$ , C-Test-Allg.:  $r = .31$ ). Keine signifikanten Zusammenhänge ergaben sich zwischen allgemeinsprachlichen bzw. fachsprachlichen Fähigkeiten und der Sportnote. Die Ergebnisse legen den Schluss nahe, dass fachsprachliche Fähigkeiten als eigenständige Teilfähigkeiten aufgefasst werden müssen und im Zusammenhang mit dem Lernerfolg im Fach stehen. Das Ergebnismuster ist dabei in Physik stärker ausgeprägt als im Fach Sport.

Baur, R.S., Grotjahn, R. & Spettmann, M. (2006). Der C-Test als Instrument der Sprachstandserhebung und Sprachförderung im Bereich Deutsch als Zweitsprache. In J.-P. Timm (Hrsg.), Fremdsprachenlernen und Fremdsprachenforschung (S. 389-406). Tübingen: Narr.

Harsch, C. & Hartig, J. (2010). Empirische und inhaltliche Analyse lokaler Abhängigkeiten im C-Test. In R. Grotjahn (Hrsg.), Der C-Test: Beiträge aus der aktuellen Forschung/The C-Test: Contributions from Current Research (S. 205-232). Frankfurt/M.: Lang.

Rincke, K. (2010). Alltagssprache, Fachsprache und ihre besondere Bedeutung für das Lernen. Zeitschrift für Didaktik der Naturwissenschaften, 16, 235-260.



## **Soziale Herkunft und Schulerfolg.**

### **Erste Ergebnisse aus einer Panelstudie zur Erweiterung aktueller Erklärungsmodelle**

Chair(s): Jörg Siewert (Justus-Liebig-Universität Gießen)

Mit der vorzustellenden Studie verfolgen wir das Ziel, den Zusammenhang zwischen der sozialen Herkunft von Schülern und ihrem Schulerfolg besser zu verstehen. Dazu werden u.a. die Kompetenzentwicklungen der Schüler analytisch in ein Verhältnis zu individuellen Herkunftsvariablen und zu ihren operativen Codes gesetzt, die sich mit Bernstein als individuelle Grundlage pädagogischer Rahmung verstehen lassen (vgl. Bernstein, 2000, Sertl & Leufer, 2012). Ein spezifisches Design macht es möglich, dabei zudem die Rolle der (neuen) Schulformen (Gymnasium und Sekundarschule/NRW bzw. Oberschule/Niedersachsen) zu untersuchen.

Die in der Schulpädagogik zu verortende quantitativ-empirische Studie vereint soziologische und psychologische Modelle und entwickelt neben der Perspektive allgemeinbildender Schulen mittelfristig auch eine berufspädagogische. Im Vortrag werden Ergebnisse und Erkenntnisse aus der ersten Erhebung präsentiert.

Theoretischer Hintergrund: Der bildungspolitische Trend in Deutschland ist eindeutig: Gymnasium und eine weitere Sekundarschulform etablieren ein zweigliedriges Schulsystem. Dieser Einschnitt soll den Schulträgern ermöglichen, trotz zurückgehender Schülerzahlen ein attraktives Schulangebot mit allen Schulabschlüssen bereit zu stellen. Aber wie verhält sich das neue zweigliedrige System daneben zu dem bildungspolitischen Problem der Bildungsungerechtigkeit? Anders gefragt: Hat das avisierte zweigliedrige Schulsystem das Potential dazu, zur Entkopplung des Schulerfolgs (primär und sekundär) von Schülern und ihrer sozialen Herkunft beizutragen?

Neben der Suche nach möglichen Ursachen für den Zusammenhang zwischen der sozialen Herkunft von Schülern und ihrem Schulerfolg gehen wir mit der projektierten Panelstudie eben dieser Fragen nach. Dazu werden einerseits die Kompetenzentwicklungen der Schüler vor dem Hintergrund eines komplexen Modells sozialer Herkunft analysiert (vgl. etwa Baumert et al., 2006, Siewert, 2013), das neben strukturellen und prozessoralen Herkunftsvariablen auch die Codes nach B. Bernstein erfasst und in pädagogisch relevante überführt. Daneben wird in Betracht gezogen, dass herkunftsspezifische Kompetenzentwicklungen nicht nur durch außerschulische Faktoren, sondern auch durch die Schule als Bildungsinstitution und die Lernökologien der Schulen erzeugt oder wenigstens begünstigt sein können. Diese These wurde in Deutschland bisher kaum untersucht; sie stellt – wie die Berücksichtigung der pädagogischen Codes und Dispositive der Schüler, die in der jüngeren bildungswissenschaftlichen Debatte zumindest im deutschsprachigen Raum kaum aufgegriffen wurden – ein echtes Desiderat dar, obwohl die Ergebnisse solcher Untersuchungen fundamental für die Entwicklung von Interventionsmaßnahmen wären. Denn ließe sich diese These empirisch stützen – und dafür hat Siewert entsprechende Indizien gefunden (2013) – hätte das unmittelbare Auswirkungen darauf, dass Interventionsmaßnahmen beim Umbau des Schulsystems berücksichtigt werden und ggf. direkt in der Schule sowie in der Lehrerbildung ansetzen müssten.

Leitfragen der Studie: Zusammengefasst soll die projektierte Studie Antworten auf folgende fünf Leitfragen liefern:

1. Kommt es auch in dem neuen zweigliedrigen Schulsystem zu herkunftsspezifischen (primären und sekundären) Bildungsunterschieden?
2. Wenn ja: Auf welche herkunftstypischen Prozesse lassen sie sich zurückführen?
3. Inwiefern kann die Theorie der pädagogischen Codes und des Dispositivs von Bernstein zur Aufklärung beitragen?
4. Welche Rolle spielen dabei die Schulen und ihre Lernökologien?
5. Inwieweit ist der Bildungserfolg der Schülerinnen und Schüler in den „neuen“ Schulformen „Gymnasium“ und „Sekundar-/Oberschule“ schulformtypisch?

Untersuchungsdesign: Die Studie wird an jeweils sieben Sekundar- und Oberschulen und an zwölf Gymnasien (je sechs in NRW und Niedersachsen) durchgeführt. Der Panel bezieht sich auf ihre Schüler in der Klassenstufe 6 (N ≈ 2.000). Die Leistungsmessungen erfolgen in drei Erhebungswellen so, dass die Kompetenzentwicklungen der Schüler während der Schulzeit denen während der Ferien gegenübergestellt werden können. Hintergrund für dieses Design ist die Prämisse, dass während der Schulzeit der schulische Einfluss auf die Kompetenzentwicklungen der Schüler überwiegt, und während der Ferienzeit der außerschulische. Neben den Leistungsentwicklungen werden die Entwicklungen der pädagogischen Kodierungen und querschnittlich die Motivation und das kognitive Intelligenzniveau erhoben. Parallel werden die Kinder und die Eltern über ihre familiäre Strukturen und Prozesse befragt.

Baumert, J., Stanat, P., & Watermann, R. (Hg.) (2006). Herkunftsbedingte Disparitäten im Bildungswesen: Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bernstein, B. (2000). Pedagogy, Symbolic Control and Identity. Theory, Research, Critique. Revised Edition. - Lanham u.a.: Rowman & Littlefield.

Sertl, M. & Leufer, N. (2012). Bernsteins Theorie der pädagogischen Codes und des pädagogischen Diskurses. Eine Zusammenschau. In: Gellert, Sertl (2012): Zur Soziologie des Unterrichts. Basil Bernsteins Code-Theorie und Rekonstruktionen des pädagogischen Diskurses. Weinheim: Juventa-Verlag. S. 15-62.

Siewert, J. (2013). Herkunftsspezifische Unterschiede in der Kompetenzentwicklung: Weil die Schule versagt? Münster, New York, München, Berlin: Waxmann.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G147



## ***Bildungsaspirationen und die Entwicklung herkunftsspezifischer Bildungsungleichheiten im Verlauf der Sekundarstufe. Ein Vergleich zwischen Bundesländern***

Zahlreiche Studien belegen, dass der Bildungserfolg von Kindern und Jugendlichen stark mit der sozialen Herkunft variiert. Unsere Studie will klären, inwieweit herkunftsspezifische Bildungsungleichheiten durch Unterschiede in der Ausgestaltung des allgemeinbildenden Schulsystems in den Bundesländern beeinflusst werden. Wir konzentrieren uns hierzu zum einen auf die Bildungsaspirationen von Jugendlichen und zum anderen auf die faktischen Bildungsverläufe (Schultyp bzw. Abschluss) ab dem Übergang in die Sekundarstufe. Von besonderem Interesse ist für uns die Frage, wie gut es den verschiedenen gestalteten Sekundarschulsystemen gelingt, zu Beginn bestehende herkunftsspezifisch ungleiche Positionierungen im Bildungssystem abzubauen. Beobachten wir in unterschiedlich offenen Sekundarschulsystemen im Bildungsverlauf ein weiteres Auseinanderdriften oder aber eine Annäherung der Kinder unterschiedlicher sozialer Herkunftsgruppen?

Bislang gibt es wenig Forschung zu unseren Fragenstellungen bezüglich der Relevanz institutioneller Charakteristika für herkunftsspezifische Bildungsungleichheiten. Schuchart und Maaz (2007) finden erhöhte Bildungsaspirationen der unteren Schichten in Bundesländern mit offenen Bildungssystemen. In Bezug auf Bildungsverläufe gibt es Indizien, dass Kinder aus den oberen Schichten stärker von institutionellen Optionen zur Korrektur von Bildungsentscheidungen profitieren (Hillmert & Jacob, 2005; Jacob & Tieben, 2007; Glaeser & Cooper, 2011). Eine Studie findet jedoch Vorteile für Kinder unterer Schichten (Trautwein, Nagy, & Maaz, 2011).

Den theoretischen Hintergrund unserer Studie bildet ein Ressourcen-Ansatz, der argumentiert, dass Kinder/Jugendliche auf Basis der unterschiedlichen Verfügbarkeit von kulturellen und ökonomischen Ressourcen im Elternhaus unterschiedlich gut auf schulische Anforderungen vorbereitet werden und dass auf Basis dieser Ressourcen auch divergierende Bildungsentscheidungen gefällt werden (z.B. Kristen, 2008). Das theoretische Grundkonzept basiert also auf der Unterscheidung zwischen primären und sekundären Effekten der sozialen Herkunft nach Boudon (1974). Die (als rational angenommenen) Entscheidungen von Eltern und Jugendlichen werden innerhalb von institutionellen Gelegenheitsstrukturen mit unterschiedlichen Anreiz- und Kostenstrukturen getroffen. Damit ist zu erwarten, dass sich herkunftsspezifische Bildungsungleichheiten potentiell zwischen Schulsystemen unterscheiden. Dies sollte sowohl für die vorgelagerten Bildungsaspirationen als auch für die faktischen Übergänge im Bildungssystem und die resultierenden Bildungsverläufe gelten. Offenere Sekundarschulsysteme, d.h. solche, in denen es einfacher ist, zu Beginn der Sekundarstufe getroffene Weichenstellungen zu korrigieren (oder diese durch eine Gesamtschuloption offen zu halten), sollten zu höheren Bildungsaspirationen führen. Zudem ist zu erwarten, dass in solchen Systemen verstärkt auch Kinder/Jugendliche aus niedrigen Herkunftsgruppen höhere Bildungsaspirationen entwickeln und sich damit den anderen Herkunftsgruppen annähern. Gleichwohl ist auf Basis der theoretischen Überlegungen zu bezweifeln, dass offenere Sekundarschulsysteme tatsächlich zu einer Verringerung der Ungleichheiten in den Bildungsergebnissen über den Bildungsverlauf führen. Denn aufgrund der Motivation des Statuserhalts bleibt der Nutzen hoher Bildung bei hoher elterlicher Bildung immer höher als bei niedriger elterlicher Bildung.

Nach Ausarbeitung der theoretischen Überlegungen und Hypothesen entwickeln wir Indikatoren zur Charakterisierung der Sekundarschulsysteme im Hinblick auf Offenheit vs. Geschlossenheit. Darauf aufbauend prüfen wir anhand von Daten des Sozio-oekonomischen Panels (SOEP) unsere Hypothesen. Unsere Untersuchungsgruppe umfasst die Geburtsjahrgänge 1982-1992, für die Informationen über Bildungsaspirationen, Noten, Bildung der Eltern, Bundesland, Schulformen und Schulabschlüsse, Geschlecht und Migrationshintergrund vorliegen. Um der Frage nach der Entwicklung von Bildungsungleichheiten nachzugehen und herauszufinden, ob in unterschiedlich offenen Bildungssystemen ein mehr oder weniger großer Schereneffekt entsteht, rekonstruieren wir die Bildungsverläufe der Untersuchungspersonen. Zur Analyse der Unterschiede zwischen institutionellen Kontexten konzentrieren wir uns auf vier Bundesländer, für die genügend große Fallzahlen vorliegen, und zwar auf Nordrhein-Westfalen und Niedersachsen (relativ offene Bildungssysteme) sowie Bayern und Baden-Württemberg (relativ geschlossene Bildungssysteme). In den Analysen kommen primär multiple logistische Regressionsmodelle zum Einsatz. Die Ergebnisse unserer Analysen belegen, dass sich die Bildungsaspirationen zwischen den untersuchten Bundesländern in der erwarteten Weise unterscheiden: die Offenheit eines institutionellen Kontextes steigert die Chancen einer hohen Bildungsaspiration und reduziert herkunftsspezifische Unterschiede in den Bildungsaspirationen. Erste (vorläufige) Ergebnisse multipler Regressionsanalysen zur zweiten Fragestellung, zeigen allerdings, dass wie erwartet in offeneren Kontexten die Chancennachteile der Kinder niedriger sozialer Herkunft faktisch nicht reduziert werden.

Boudon, R. (1974). Education, Opportunity, and Social Inequality. New York: John Wiley & Sons.

Glaeser, J., Cooper, B. (2011). Selectivity and Flexibility in the German Secondary School System: A Configurational Analysis of Recent Data from the German Socio-Economic Panel. *European Sociological Review*, 27(5): 570-585.

Hillmert, S., Jacob, M. (2005). Zweite Chance im Schulsystem? Zur sozialen Selektivität bei ‚späteren‘ Bildungsentscheidungen. In: Berger, P./Kahlert, H. (Hrsg.). *Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert*. Weinheim und München: Juventa Verlag, S. 155-178.

Jacob, M., Tieben, N. (2007). Social Selectivity of Track Mobility in Secondary Schools. A Comparison of Intra-secondary Transitions in Germany and the Netherlands. ZEW Working Paper No. 105. Mannheim: Centre for European Economic Research.

Kristen, C. (2008). Schulische Leistungen von Kindern aus türkischen Familien am Ende der Grundschulzeit. Befunde aus der IGLU-Studie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Sonderband 48: 230-251.

Schuchart, C., Maaz, K. (2007). Bildungsverhalten in institutionellen Kontexten: Schulbesuch und elterliche Bildungsaspiration am Ende der Sekundarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59(4): 640-666.

Trautwein, U., Nagy, G., Maaz, K. (2011). Soziale Disparitäten und die Öffnung des Sekundarschulsystems: Eine Studie zum Übergang von der Realschule in die gymnasiale Oberstufe. *Zeitschrift für Erziehungswissenschaft*, 14: 455-463.





**Schichtspezifische Bildungsungleichheiten in der Argumentationskompetenz:  
Welche Rolle spielen Elternhaus, Intelligenz und Schulform in der Sekundarstufe I?**

Theoretischer Hintergrund. Die Fähigkeit zu argumentieren gilt als zentrale Voraussetzung gesellschaftlicher Partizipation und als wesentliche schulische Schlüsselkompetenz. Dennoch mangelt es an Studien zur Genese der Argumentationskompetenz und zu der Frage, welcher Stellenwert hierbei dem Elternhaus neben der Schulform und den kognitiven Fähigkeiten des Kindes zukommt. Dieser Frage wird im BMBF-geförderten Projekt „Die Rolle familialer Unterstützung beim Erwerb von Diskurs- und Schreibfähigkeiten in der Sekundarstufe I“ (FUnDuS) nachgegangen, in dem Familien von der fünften bis zur neunten Klasse begleitet werden. Methode. Im vorliegenden Beitrag sollen erste querschnittliche Befunde vorgestellt werden. N = 1670 SiebtklässlerInnen (63 % Gymnasiasten, 37 % HauptschülerInnen; 54 % männlich) wurden gebeten, verschiedene im Projekt konzipierte Argumentationsaufgaben sowie Teile des Kognitiven Fähigkeitstests (KFT 4-12+R; Heller & Perleth, 2000) zu bearbeiten. Zusätzlich wurden Einschätzungen der SchülerInnen zum sprachlichen Anregungsgehalt in Familienkonversationen (z.B. Toleranz von abweichenden Meinungen, Begründungen in der Eltern-Kind-Interaktion, elterliches Vorleben argumentativer Aushandlungsprozesse) sowie Elternangaben zum sozioökonomischen (HISEI) und kulturellen Kapital (häuslicher Bücherbestand) erhoben. Ergebnisse. In einem Strukturgleichungsmodell ( $\chi^2 [82, N = 1670] = 446.7, p = .00; CFI = .93; RMSEA = .05$ ) stellten sich unter Kontrolle der Mehrebenenstruktur die Schulform ( $\beta = .35$ ), die Intelligenz ( $\beta = .39$ ) sowie der sprachliche Anregungsgehalt in Familienkonversationen ( $\beta = .22$ ) als signifikante Prädiktoren der Argumentationskompetenz heraus und medierten den bivariaten Zusammenhang ( $r = .65$ ) zwischen dem SES und der Argumentationskompetenz ( $ps < .01$ ). Weiterhin korrelierte die Argumentationskompetenz signifikant mit den Schulnoten in Deutsch, Mathematik und Englisch ( $.25 \leq r_s \leq .36$ ). Diskussion. Die Ergebnisse stützen die Annahme, dass eine Verbesserung des sprachlichen Anregungsgehalts beispielsweise über Elterntrainings einen potentiellen Ansatzpunkt zur Entkopplung der sozialen Herkunft und der Argumentationskompetenz darstellt, um zur Verringerung von Bildungsungleichheiten beizutragen.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G147

## ***Diskriminierung im Bildungswesen:***

### ***Implikationen methodologischer, theoretischer und statistischer Überlegungen für die empirische Forschung***

Diskriminierung durch die Lehrkräfte ist eine mögliche Erklärung der vergleichsweise hohen ethnischen und sozialen Disparitäten im deutschen Bildungssystem. Tatsächlich gibt es aber immer noch erstaunlich wenig Beiträge, die Diskriminierung explizit untersuchen. Kein einziger Beitrag untersucht bisher systematisch die verschiedenen Formen und Definitionen von Diskriminierung, die unterschiedlichen Theorien und ihre Anwendbarkeit auf das deutsche Bildungssystem sowie die Konsequenzen für die empirische Überprüfung und statistische Modellierung. Mit dem vorliegenden Beitrag will ich helfen, diese Lücke zu schließen.

Zu diesem Zweck diskutiere ich verschiedene Definitionen von Diskriminierung aus Ökonomie, Sozialpsychologie und Soziologie. Zu den zwei wichtigsten Unterscheidungen gehört die zwischen differential effect discrimination und differential treatment discrimination (Blank, 2004) bzw. die Unterscheidung der Fragen ob z.B. türkische Kinder im deutschen Bildungssystem unter Diskriminierung leiden, oder ob sie das tun weil sie türkisch sind. Diese Unterscheidung hat Auswirkungen auf die Identifikations- und Schätzstrategie sowohl in Studien mit Surveydaten als auch mit experimentellen Daten. Die zweite sehr wichtige Unterscheidung ist die zwischen individueller Diskriminierung und Diskriminierung auf Gruppenebene. Aigner und Cain (1977) zeigen durch einfache Formalisierungen, dass es nicht notwendigerweise zu Diskriminierung auf Gruppenebene kommen muss, obwohl alle Angehörigen einer bestimmten Gruppe individuell diskriminiert werden. Ich zeige, dass dieses Argument allerdings auf sehr starken und also unrealistischen Annahmen beruht.

Als Zweites überprüfe ich, inwieweit verschiedene Theorien zur Erklärung von Diskriminierung auf das deutsche Bildungssystem anwendbar sind. Insbesondere die ökonomischen Theorien von Becker (1957) und die Theorie der statistischen Diskriminierung (Aigner & Cain, 1977) sind in empirischen Untersuchungen indirekt überprüft worden. Ich argumentiere, dass die bisherigen Anwendungen einerseits die Mechanismen von Beckers Taste Discrimination ignorieren und andererseits unrealistische Brückenannahmen treffen, die zur Ableitung von – folglich unsinnigen – Hypothesen aus der Theorie der statistischen Diskriminierung führen. Ich verweise außerdem auf Theorien und Modelle aus der Sozialpsychologie – wie z.B. der Social Identity Theory und Befunde der Social Cognition Forschung – die weitgehend ignoriert wurden.

Als Drittes zeige ich mittels einfacher Formalisierungen und Simulationen, wie die verschiedenen Formen von Diskriminierung identifizierbar und mittels unterschiedlicher Methoden schätzbar sind und unter welchen Bedingungen welche Theorie zur Erklärung von Diskriminierung auf der Gruppenebene und damit letztlich zur Erklärung der Bildungsungleichheit beitragen kann. Ich zeige außerdem, unter welchen Bedingungen Diskriminierung überschätzt wird – ein häufiger Vorwurf an den sogenannten Residualansatz. Ich zeige außerdem einen in der deutschen Literatur ignorierten Fall, nämlich die Unterschätzung von Diskriminierung im Residualansatz. Die Formalisierungen und Simulationen stützen sich unter anderem auf Daten der Studie PIRLS/IGLU 2006 und auf Ergebnisse experimenteller Untersuchungen.

Ich schließe mit einigen beispielhaften Befunden zu den Determinanten diskriminierendem Verhaltens, insbesondere Stereotype und Vorurteile deutscher Lehrkräfte. Die Befunde zu den Stereotypen der Lehrkräfte stammen aus ersten Analysen der unveröffentlichten Daten des NEPS-Großpilots in Klasse 2 (Startkohorte 2). Zudem lege ich die erste repräsentative Untersuchung ethnischer Vorurteile deutscher Lehrkräfte vor. Dazu verwende ich Daten aus unterschiedlichen Wellen des ALLBUS, in denen ich ausreichend Lehrkräfte mittels der ISCO-Berufskodierung identifizieren konnte.

Zusammenfassend argumentiere ich, dass es sehr wahrscheinlich ist, dass Kinder mit Migrationshintergrund und solche aus Familien niedriger sozialer Herkunft unter systematischer Diskriminierung im deutschen Bildungssystem leiden und also Diskriminierung in der Tat einen Beitrag zur Erklärung der Bildungsungleichheit leistet.

Aigner, D.J., & Cain, G. G. (1977). Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review* 30, no. 2: 175–187.

Becker, G.S. (1971). *The Economics of Discrimination*. Chicago, IL: University Of Chicago Press.

Blank, R.M., M. Dabady, and C.F. Citro. *Measuring Racial Discrimination*. National Research Council of the National Academies, 2004.

## Lehramtsstudiengänge als Aufstiegsfachrichtungen?

### Die Wahl eines Lehramtsstudiums aus Sicht der soziologischen Bildungsentscheidungsforschung

Ein systematischer Zusammenhang zwischen sozialer Herkunft und Bildungsentscheidungen ist vielfach nachgewiesen worden: beim Übergang von der Grund- in die Sekundarstufe I (u.a. Stocké, 2007), in die Sekundarstufe II (u.a. Stubbe, 2009) oder beim Hochschulzugang (Becker & Hecken, 2007; Müller & Pollak, 2010). Auch wenn die im Bildungsverlauf zunehmende Selektivität dazu führt, dass bei der relativ leistungshomogenen Gruppe der Studierenden nur geringere Herkunftseffekte sichtbar werden (Mare, 1980), wurde für die Studienfachwahl als „Spezialfall der Ausbildungsentscheidung“ (Becker et al., 2010) die soziale Selektivität wiederholt belegt (Davies & Guppy, 1997; van de Werfhorst et al., 2001; Heine & Willich, 2006; Reimer & Pollak, 2010). Während Akademikerkinder Studienrichtungen wie Medizin und Rechtswissenschaften präferieren, sind für Arbeiterkinder u.a. Lehramtsstudiengänge interessant (Reimer & Pollak, 2010). Daher gelten diese als Aufstiegsfachrichtungen. Die Wahl eines Lehramtsstudiums als eine spezifische Form der Studienfachwahl muss deshalb im Kontext des Systems der sozialen Stratifikation betrachtet werden.

Im Fokus der aktuellen Forschung zur Wahl eines Lehramtsstudiums stehen berufliche Interessen und Motive sowie individuelle Voraussetzungen der zukünftigen Lehrkräfte (Denzler & Wolter, 2008). Die soziale Herkunft der Studierenden wird selten berücksichtigt und aus den wenigen Forschungsergebnissen lassen sich keine differenzierten Effekte der Herkunft ableiten: Während also bei der gemeinsamen Betrachtung aller Lehramtsstudiengänge im Vergleich zu anderen Studiengängen eine schichtspezifische Präferenz nachgewiesen ist, gibt es keine eindeutigen Ergebnisse bezüglich verschiedener Lehramtsausbildungen. Während einige Studien belegen, dass insbesondere nicht-gymnasiale Lehramtsstudiengänge soziale Aufstiegskanäle darstellen (Rothland, 2011; Neugebauer, 2013), wird von anderen Autoren darauf verwiesen, dass es zu einer Angleichung der Rekrutierungsmuster zwischen Gymnasial- und Grundschullehrern gekommen ist (Enzelberger, 2001; Kühne, 2006). In bestehenden Studien werden allerdings meist mehrere Studiengänge gruppiert und keine differenzierte Betrachtung aller Lehramtsausbildungsgänge vorgenommen. Darüber hinaus stehen vorwiegend ausgewählte Regionen oder Fächergruppen im Fokus und die Analysen beziehen sich auf die nicht-reformierten Studiengänge.

Vor diesem Hintergrund wird in diesem Vortrag untersucht, welchen Einfluss die soziale Herkunft auf die Wahl eines Lehramtsstudiums in Deutschland aktuell hat und wie sich diese Herkunftseffekte zwischen verschiedenen Lehramtsstudiengängen (Primarstufe, Sekundarstufe I, Sekundarstufe II, Berufsschulen, Sonderpädagogik) unterscheiden.

Die Studie trägt dazu bei bestehende Forschungslücken zu schließen, indem die erziehungswissenschaftliche Forschung zu Lehrkräften mit der soziologischen Bildungsentscheidungsforschung verknüpft wird. Der Zusammenhang von Studienfachwahl und sozialer Herkunft lässt sich mit Hilfe von Boudons (1974) Konzeption von primären und sekundären Herkunftseffekte näher ergründen. Primäre Effekte treten demnach auf, wenn Kinder aus bildungsnahen Familien - bedingt durch den besseren Zugang zu bildungsrelevanten Ressourcen - die Anforderungen der jeweiligen Stufe des Bildungssystems besser erfüllen. Boudon spricht von sekundären Herkunftseffekten, wenn bei gleichen Leistungen herkunftsspezifische Unterschiede im Bildungsverhalten auftreten. Basierend auf Rational Choice-Annahmen, sind herkunftsspezifische Bildungsentscheidungen das Ergebnis von schichtspezifischen Abwägungen der Kosten, des Nutzens und der Erfolgswahrscheinlichkeit (Erikson & Jonsson, 1996; Breen & Goldthorpe, 1997). Datengrundlage der Analysen ist die Befragung der Startkohorte 5 des Nationalen Bildungspanels. Hierbei wurden die Bildungsbiographien, die soziale Herkunft, individuelle Charakteristika und Einstellungen von 18.252 StudienanfängerInnen des Wintersemesters 2010/11 erfasst ([www.neps-data.de](http://www.neps-data.de)). Der vollständige Datensatz liegt vor und wird zurzeit analysiert. Mittels binärer und multinomialer logistischer Regressionen werden unter Berücksichtigung der Mehrebenenstruktur des Datensatzes Herkunftseffekte auf die Wahl verschiedener Lehramtsstudiengänge identifiziert. Erste Analysen zeigen, dass primäre und sekundäre Herkunftseffekte auf die Studienfachwahl Lehramt wirken, die stark nach Lehramtsstudiengängen variieren. Die Ergebnisse weisen darauf hin, dass die Sekundarstufen I- und das Berufsschullehrämter als Aufstiegsfachrichtungen fungieren, während hingegen ein akademischer Bildungshintergrund die Wahrscheinlichkeit für eine Gymnasiallehramtswahl erhöht. Hieran anknüpfend, sollen mit differenzierteren Analysen die hinter den Herkunftseffekten stehenden Mechanismen beleuchtet werden, um Erkenntnisse über die spezifischen Kosten-Nutzen-Kalkulationen bei dieser Bildungsentscheidung zu erlangen. Von den Ergebnissen ausgehend werden in der Diskussion mögliche Implikationen für die Rekrutierung und Ausbildung von Lehramtsstudierenden und der Zusammenhang zwischen bestehenden Herkunftseffekten und der sozialen Selektivität des deutschen Schulsystems erörtert.

Becker, R., Haunberger, S. & Schubert, F. (2010). Studienfachwahl als Spezialfall der Ausbildungsentscheidung und Berufswahl. *Zeitschrift für ArbeitsmarktForschung*, 42(4), 292–310.

Becker, R. & Hecken, A. E. (2007). Studium oder Berufsausbildung? Eine empirische Überprüfung der Modelle zur Erklärung von Bildungsentscheidungen von Esser sowie von Breen und Goldthorpe. *Zeitschrift für Soziologie*, 36(2), 100–117.

Boudon, R. (1974). *Education, opportunity, and social inequality. Changing prospects in Western society*. New York: Wiley.

Breen, R. & Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and Society*, 9(3), 275–305.

Davies, S. & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social forces*, 75(4), 1417–1438.

Denzler, S. & Wolter, S. C. (2008). Selbstselektion bei der Wahl eines Lehramtsstudiums: Zum Zusammenspiel individueller und institutioneller Faktoren. *Beiträge zur Hochschulforschung*, 30(4), 112–141.

Enzelberger, S. (2001). Sozialgeschichte des Lehrerberufs. *Gesellschaftliche Stellung und Professionalisierung von Lehrerinnen und Lehrern von den Anfängen bis zur Gegenwart*. Weinheim [u.a.]: Juventa-Verlag.

Erikson, R. & Jonsson, J. (1996). Explaining Class Inequality in Education: The Swedish test Case. In R. Erikson (Hrsg.), *Can education be equalized?: the Swedish case in comparative perspective* (S. 1–63): Westview Press.

Heine, C. & Willich, J. (2006). Studienberechtigte 2005. Übergänge in Studium, Ausbildung und Beruf. HIS: Forum Hochschule 6.

Kühne, S. (2006). Das soziale Rekrutierungsfeld der Lehrer. *Zeitschrift für Erziehungswissenschaft*, 9(4), 617–631.

Mare, R. D. (1980). Social Background and School Continuation Decisions. *Journal of the American Statistical Association*, 75(370), 295–305.

Müller, W. & Pollak, R. (2010). Weshalb gibt es so wenige Arbeiterkinder an Deutschlands Universitäten? In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (4. Aufl., S. 305–344): VS Verlag für Sozialwissenschaften.

Neugebauer, M. (2013). Wer entscheidet sich für ein Lehramtsstudium – und warum? Eine empirische Überprüfung der These von der Negativelektion in den Lehrerberuf. *Zeitschrift für Erziehungswissenschaft*, 157–184.

Reimer, D. & Pollak, R. (2010). Educational Expansion and Its Consequences for Vertical and Horizontal Inequalities in Access to Higher Education in West Germany. *European Sociological Review*, 26(4), 415–430.

Rothland, M. (2011). Wer entscheidet sich für den Lehrerberuf? Forschung zum soziodemographischen Profil sowie zu Persönlichkeits- und Leistungsmerkmalen angehender Lehrkräfte. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 243–267). Münster: Waxmann.

Stocké, V. (2007). Strength, Sources, and Temporal Development of Primary Effects of Families' Social Status on Secondary School Choice. SFB 504 discussion paper 07-60. Sonderforschungsbereich 4, Universität Mannheim.

Stubbe, T. C. (2009). Bildungsentscheidungen und sekundäre Herkunftseffekte. *Soziale Disparitäten bei Hamburger Schülerinnen und Schülern der Sekundarstufe I*. Münster, New York, NY, München, Berlin: Waxmann.

van de Werfhorst, H. G., de Graaf, N., Dirik, & Kraaykamp, G. (2001). Intergenerational Resemblance in Field of Study in the Netherlands. *European Sociological Review*, 17(3), 275–293.

[www.neps-data.de](http://www.neps-data.de) (letzter Zugriff: 29. Oktober 2013)



## ***Bildungsverläufe vom Ende der Grundschulzeit bis zum jungen Erwachsenenalter im Kontext familialer Dynamiken***

Chair(s): Henning Lohmann (Universität Osnabrück)

Der Einfluss der sozialen Herkunft auf den Bildungserfolg ist in Deutschland und anderen Ländern breit belegt. In der Mehrzahl der vorliegenden Studien zum Thema wird dabei die soziale Herkunft über den Status der Eltern zu einem bestimmten Zeitpunkt im Bildungsverlauf der Kinder operationalisiert und somit als unveränderlich angenommen. Gängige Erklärungsansätze gehen dabei davon aus, dass der Einfluss der sozialen Herkunft sowohl auf die herkunftsspezifische Übertragung von kulturellem Kapital wie auch auf unterschiedliche Bildungsentscheidungen zurückzuführen ist. Daneben gibt es eine Reihe von Studien, die den Einfluss kritischer Lebensereignisse wie Arbeitslosigkeit, Trennung oder Tod der Eltern bzw. familiäre Armutsphasen als bedeutsam für den Bildungserfolg ansehen. Der Einfluss dieser Ereignisse wird zumeist auf zwei unterschiedliche Mechanismen zurückgeführt. Zum einen sind die genannten Ereignisse mit materiellen Einschränkungen verbunden, zum anderen mit anderen Belastungen der betroffenen Familien. Durch die Perspektive auf familiäre Ereignisse wird die Veränderlichkeit der Herkunftsbedingungen berücksichtigt und eine Zurechnung auf spezifische Mechanismen erleichtert. Gleichzeitig ist aufgrund der sozialen Selektivität des Auftretens bestimmter Ereignisse jedoch nicht immer klar, inwieweit der Einfluss familialer Ereignisse nicht doch auch auf allgemeine Aspekte der sozialen Herkunft zurückzuführen sind.

In dem Papier werden unter Berücksichtigung allgemeiner Aspekte der sozialen Herkunft sowohl die familialen Ereignisse als auch der kindliche Bildungsverlauf in einer Längsschnittperspektive untersucht. Wir betrachten Bildungsverläufe vom Zeitpunkt des in Deutschland so zentralen Übergangs in die mehrgliedrige Sekundarstufe bis ins junge Erwachsenenalter, in dem zumeist die allgemeine Schullaufbahn und teilweise auch Ausbildungen abgeschlossen bzw. nachschulische Bildungswege erst einmal eingeschlagen sind. Diese Perspektive ermöglicht es zu überprüfen, inwieweit die frühzeitige Platzierung im mehrgliedrigen Schulsystem spätere Bildungsentscheidungen bereits vorwegnimmt oder ob spätere familiäre Veränderungen Bildungswege noch in einschneidender Weise beeinflussen.

Um Veränderungen des familialen Status verlässlich zu erfassen, ist dabei die Verwendung von prospektiv erhobenen Daten notwendig wie sie in Haushaltspaneldata vorliegen. Wir verwenden daher das Sozio-oekonomische Panel (SOEP) der Jahre 1984-2012. Das SOEP bietet die Möglichkeit, Bildungsverläufe von Kindern und jungen Erwachsenen über einen längeren Zeitraum im Kontext des elterlichen Haushalts und auch nach dem Verlassen dieses Haushalts zu analysieren. Als Startpunkt des Beobachtungsfensters wählen wir den Zeitpunkt des Übergangs in die Sekundarstufe. Zu diesem Zeitpunkt betrachten wir vor allem den Einfluss des beruflichen Status und die Bildung der Eltern. Weiter betrachten wir familiäre Dynamiken – Ereignisse wie Armutsepisoden, Veränderungen der Partnerschaftskonstellation und Phasen der Arbeitslosigkeit – nach dem erfolgten Übergang. Wir erwarten, dass diese Dynamiken einen Einfluss auf spätere Bildungsergebnisse haben und – trotz einer immer noch recht geringen Mobilität zwischen Bildungsgängen - Veränderungen eines einmal eingeschlagenen Bildungsweges bewirken können. Wir betrachten dabei ein möglichst positives und ein möglichst negatives nachschulisches Bildungsergebnis: Zum einen die Aufnahme eines Universitätsstudiums bis spätestens zum Alter von 23 Jahren, zum anderen Ausbildungslosigkeit (bzw. Abbruch einer Ausbildung nach 1 Jahr oder weniger) bis zum selben Alter. Als weitere Faktoren werden der Migrationshintergrund, das Geschlecht des Kindes, Geschwisterangaben, die Region, das Jahr des Übergangs (unterteilt in drei Perioden) und der Schultyp im ersten Jahr nach dem Übergang berücksichtigt. Für eine Substichprobe liegen weiterhin Angaben zu Schulnoten und zur Klassenwiederholung vor, die in zusätzlichen Analysen verwendet werden.

Neben deskriptiven Analysen schätzen wir multivariate Modelle (lineares Wahrscheinlichkeitsmodell und Logit-Modelle). Erste Ergebnisse mit Blick auf die Erklärung der beiden gegenläufigen Bildungsergebnisse Studienbeginn und Ausbildungslosigkeit zeigen, dass unterschiedlichen Einflussfaktoren eine Rolle spielen und sich nicht einfach ein spiegelbildliches Ergebnis zeigt. Während die Wahrscheinlichkeit ein Studium aufzunehmen sehr stark auf den in herkömmlicher Weise gemessenen familialen Bildungshintergrund zurückzuführen ist, scheint die Ausbildungslosigkeit stärker von elterlichen Armuts- und Arbeitslosigkeitsphasen geprägt zu sein.

## Differenzielle Lern- und Entwicklungsmilieus – ein domänenspezifisches Phänomen?

Die Zuweisung von Schülerinnen und Schülern zu institutionell getrennten Bildungsgängen führt einerseits zu sozialer Stratifizierung und Segregation (Baumert, Stanat & Watermann, 2006; Baumert, Trautwein & Artelt, 2003). Andererseits eröffnen „differenzielle Lern- und Entwicklungsmilieus“ (Baumert & Schümer, 2001, S. 462) den Schülerinnen und Schülern je nach besuchter Schulform unabhängig von und zusätzlich zu individuellen Lernvoraussetzungen unterschiedliche Entwicklungschancen: So können schulformspezifische Unterschiede in den institutionellen Arbeits- und Lernbedingungen oder pädagogisch-didaktischen Traditionen dazu führen, dass die individuelle Entwicklung und Förderung von Schülerinnen und Schülern ungleich verläuft, erschwert oder sogar verhindert wird (ebd.).

Effekte differenzieller Lern- und Entwicklungsmilieus wurden bislang vielfach, für verschiedene Altersgruppen und fachspezifische Kompetenzdomänen nachgewiesen (z. B. Baumert, Becker, Neumann & Nikolova, 2009, 2010; Becker, Lüdtke, Trautwein & Baumert, 2006; Nikolova, 2011): Nicht erst gegen Ende der Sekundarstufe I (Baumert et al., 2006), sondern bereits am Ende der Beobachtungsstufe (Bonsen, Gröhlich & Bos, 2009) konnten Kompositionseffekte der sozialen und leistungsbezogenen Schülerzusammensetzung sowie institutionelle Effekte der Schulformgliederung auf fachspezifische Kompetenzen der Schülerinnen und Schüler nachgewiesen werden. Gröhlich, Guill, Scharenberg und Bos (2010a, 2010b) konnten Kompositions- und Institutionseffekte auf die fachspezifische Leistungsentwicklung über die gesamte Sekundarstufe I hinweg belegen.

Dieser Beitrag soll über den bisherigen Forschungsstand hinausgehen, indem zum einen die Entwicklung in zwei unterschiedlichen fachspezifischen Kompetenzdomänen (Leseverständnis und Mathematik) und zum anderen auch ein fachunabhängiger Kompetenzbereich, d.h. die Intelligenzentwicklung, längsschnittlich untersucht werden. Anknüpfend an die Analysen von Becker, Lüdtke, Trautwein, Köller und Baumert (2012) ist das Ziel unseres Beitrages, die Bedeutung von Kompositions- und Institutionseffekten für die Entwicklung fachspezifischer und fachunabhängiger Kompetenzen an nur einem Längsschnittdatensatz für dieselbe Schülerpopulation vergleichend zu analysieren. Im Mittelpunkt stehen dabei folgende Forschungsfragen: Lassen sich differenzielle Lern- und Entwicklungsmilieus domänenübergreifend nachweisen oder sind sie domänenspezifisch? Welche Bedeutung haben die verschiedenen Kompositions- und Institutionsmerkmale für die Entwicklung in den verschiedenen Kompetenzdomänen? Lassen sich Faktoren identifizieren, denen eine domänenübergreifende oder aber bloß domänenspezifische Bedeutung zukommt?

Im Rahmen der Hamburger Schulleistungsstudie KESS („Kompetenzen und Einstellungen von Schülerinnen und Schülern“) wurde eine vollständige Schülerkohorte (N = 9.628) längsschnittlich über drei Messzeitpunkte von Jahrgangsstufe 4 bis 8 untersucht. Neben fachspezifischen Kompetenzen im Leseverständnis und in Mathematik wurden auch die non-verbale kognitive Grundfähigkeit (KFT; Heller & Perleth, 2000) als fachunabhängige Kompetenz sowie weitere soziokulturelle Herkunftsmerkmale der Schülerinnen und Schüler erfasst. Methodisch wird ein Mehrebenenmodell mit zwei Analyseebenen in HLM (Raudenbush, Bryk & Congdon, 2009) spezifiziert, das auf der Individualebene Hintergrundmerkmale der Eingangsselektivität von Schulformen (Vorwissen, kognitive Grundfähigkeit, soziale Lage, Geschlecht und Migrationshintergrund), auf der Schulebene das mittlere Vorwissen und die mittlere soziale Lage als Kompositionsmerkmale sowie die Schulform kontrolliert. Die Analysen beschränken sich zunächst auf die Entwicklungsverläufe in den Jahrgangsstufen 5 und 6. Die Ergebnisse zeigen, dass die Schulformgliederung in der Hamburger Sekundarstufe in allen untersuchten Kompetenzdomänen zu differenziellen Entwicklungsverläufen in substantiellem Ausmaß führt: Gerade die Gymnasiastinnen und Gymnasiasten können deutlich von ihrem schulischen Lernmilieu profitieren. Sowohl für die fachspezifische Leistungsentwicklung als auch für die Entwicklung der non-verbalen kognitiven Grundfähigkeit erweist sich die Leistungs- bzw. Fähigkeitskomposition der Schülerschaft als relevanter Prädiktor. Allerdings kommt es nur bei der Entwicklung der non-verbalen kognitiven Grundfähigkeit auch auf die soziale Komposition der Schülerschaft an.

Mit Bezug auf die Forschungsfragen lässt damit feststellen, dass es sich bei differenziellen Lern- und Entwicklungsmilieus um ein domänenübergreifendes Phänomen handelt. Neben Institutionsmerkmalen kommt auch der Leistungs- und Fähigkeitskomposition eine domänenübergreifende Bedeutung für die Entwicklung in allen drei Domänen zu. Die Auswirkungen der sozialen Komposition erweisen sich in unseren Längsschnittanalysen als domänenspezifisch und schlagen sich unter Kontrolle des mittleren Ausgangsniveaus nur in der Entwicklung der kognitiven Grundfähigkeit, aber nicht in der fachspezifischen Kompetenzentwicklung nieder.

- Baumert, J., Becker, M., Neumann, M. & Nikolova, R. (2009). Frühübergang in ein grundständiges Gymnasium – Übergang in ein privilegiertes Entwicklungsmilieu? Ein Vergleich von Regressionsanalyse und Propensity Score Matching. *Zeitschrift für Erziehungswissenschaft*, 12 (2), 189–215.
- Baumert, J., Becker, M., Neumann, M. & Nikolova, R. (2010). Besondere Förderung von Kernkompetenzen an Spezialgymnasien? Der Frühübergang in grundständige Gymnasien in Berlin. *Zeitschrift für Pädagogische Psychologie*, 24 (1), 5–22.
- Baumert, J. & Schümer, G. (2001). Schulformen als selektionsbedingte Lernmilieus. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 454–467). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten – Institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Opladen: Leske + Budrich.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie*, 20 (4), 233–242.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104 (3), 682–699.
- Bonsen, M., Gröhlich, C. & Bos, W. (2009). Differentielle Lern- und Entwicklungsmilieus in der Hamburger Beobachtungsstufe? In W. Bos, M. Bonsen & C. Gröhlich (Hrsg.), *KESS 7 – Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (S. 113–122). Münster: Waxmann. (HANSE – Hamburger Schriften zur Qualität im Bildungswesen; Bd. 5).
- Gröhlich, C., Guill, K., Scharenberg, K. & Bos, W. (2010a). Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (S. 100–106). Münster: Waxmann. (HANSE – Hamburger Schriften zur Qualität im Bildungswesen; Bd. 6).
- Gröhlich, C., Guill, K., Scharenberg, K. & Bos, W. (2010b). Kumulative Effekte differentieller Lern- und Entwicklungsmilieus innerhalb der Sekundarstufe I am Beispiel der Mathematikleistung. In B. Schwarz, P. Nenniger & R. S. Jäger (Hrsg.), *Erziehungswissenschaftliche Forschung – nachhaltige Entwicklung. Beiträge zur 5. DGE-Sektionstagung „Empirische Bildungsforschung“/AEPF-KBBB im Frühjahr 2009* (S. 473–479). Landau: Verlag Empirische Pädagogik. (Erziehungswissenschaft; Bd. 28).
- Heller, K. A. & Perleth, C. (2000). KFT 4–12-R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz.
- Nikolova, R. (2011). Grundschulen als differenzielle Entwicklungsmilieus. Objektive Kontextmerkmale der Schülerzusammensetzung und deren Auswirkung auf die Mathematik- und Leseleistungen. Münster: Waxmann. (Pädagogische Psychologie und Entwicklungspsychologie; Bd. 81).
- Raudenbush, S. W., Bryk, A. S. & Congdon, R. (2009). *HLM 6.08 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G150





## **Die sozioemotionale Entwicklung von Klassenwiederholern in der Sekundarstufe I – ein mehrbenenanalytischer Beitrag im Rahmen des Nationalen Bildungspanels (NEPS)**

Klassenwiederholungen sind sowohl in Hinblick auf ihre Effektivität als auch hinsichtlich der Gerechtigkeit umstritten. Die Debatte um ihre Beibehaltung oder Abschaffung ist in Deutschland in vollem Gange. In den Bundesländern Hamburg und Berlin finden derzeit Schulstrukturentwicklungsmaßnahmen statt, mit deren Umsetzung die Klassenwiederholung vollständig abgeschafft werden soll. Andere Bundesländer wie Bayern nutzen die Klassenwiederholung weiterhin überdurchschnittlich häufig (vgl. Statistisches Bundesamt, 2012). Ungenügend geklärt ist in diesem Rahmen noch immer die Frage, ob und inwiefern Repetenten überhaupt von dieser Maßnahme profitieren.

Im englischsprachigen Raum wurden Effekte von Klassenwiederholungen eingehend untersucht, in der Bundesrepublik Deutschland ist die Forschungslage weniger umfangreich: Viele Studien weisen darauf hin, dass Schülerinnen und Schüler nicht von einer Klassenwiederholung profitieren (vgl. Ehmke, Drechsel & Carstensen, 2010; Fabian, 2012), ein Großteil der Studien zeigt gar einen negativen Effekt auf. Dennoch gibt es auch Forschungsarbeiten, die positive Effekte nahe legen (vgl. Schwerdt & West, 2013).

Die am häufigsten rezipierten Studien befassen sich jedoch hauptsächlich mit der Entwicklung der Schulleistung von Repetenten. Effekte von Klassenwiederholungen auf sozioemotionale Faktoren wie das schulische Selbstkonzept, die Leistungsmotivation oder Schul- und Leistungsangst sind dagegen deutlich seltener Gegenstand empirischer Forschungsarbeiten.

Von Befürwortern der Klassenwiederholung wird auf der einen Seite unterstellt, dass sich durch eine Klassenwiederholung das Selbstkonzept und Selbstwertgefühl der Repetenten dadurch aufbaut, dass sie in der aufnehmenden Klasse nun zu den Schülerinnen und Schülern mit erweitertem Vorwissen gehören (vgl. Bless, Schüpbach & Bonvin, 2004). Die Klassenwiederholung soll sich außerdem durch neue Erfolge positiv auf die Lern- und Schulfreude sowie auf die Motivation auswirken (vgl. Bless et al., 2004). Gegner der Klassenwiederholung merken konträr an, dass Stigmatisierungen auftreten, Versagensgefühle entstehen, die Motivation sinkt und sich der Verlust der gewohnten Bezugsgruppe negativ auf die sozioemotionale Lage der Repetenten auswirkt.

Es stellt sich also die Frage: Wie wirkt eine Klassenwiederholung auf die Entwicklung sozioemotionaler Faktoren der Schülerinnen und Schüler? Grundlage der Analysen sollen die Daten der National Educational Panel Study (NEPS) sein. Der Datensatz umfasst Schülerinnen und Schüler, die auch nach einer Klassenwiederholung längsschnittlich begleitet wurden. Außerdem liegen die Daten von Schülerinnen und Schülern mit ungebrochenem Bildungsverlauf (Promovierte) vor, was einen interindividuellen Vergleich sowohl mit der alten (abgebenden) Lerngruppe, also den promovierten Schülerinnen und Schülern, als auch mit der neuen (aufnehmen) Lerngruppe sowie intraindividuelle Vergleiche im Zeitverlauf bezüglich der Entwicklung sozioemotionaler Faktoren bei den Repetenten zulässt.

Die Daten sollen anhand hierarchischer linearer Modelle (HLM, Raudenbush et al., 2009) ausgewertet werden. Dazu sollen Effekte individueller Merkmale (Geschlecht, Migrationshintergrund, soziale Lage, Vorwissen, kulturelles Kapital) auf die sozioemotionale Entwicklung zusätzlich zu einer erfolgten Klassenwiederholung berücksichtigt werden. Auf aggregierter Ebene werden die mittlere kognitive Grundfähigkeit, die mittlere soziale und leistungsbezogene Komposition der Schulklassen sowie die Schulform kontrolliert.

Es wird erwartet, dass eine Klassenwiederholung unter Kontrolle anderer individueller und klassenbezogener Merkmale einen negativen Effekt auf die sozioemotionale Entwicklung der Repetenten hat.

Bless, G., Schüpbach, M. & Bonvin, P. (2004). Klassenwiederholung: Determinanten, Wirkungen und Konsequenzen. Bern: Haupt Verlag.

Ehmke, T., Drechsel, B. & Carstensen, C. H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. *Studies in Educational Evaluation*, 36, (1-2), 27-35.

Fabian, P. (2012). Die Lernentwicklung in der Sekundarstufe I von Kindern mit und ohne Klassenwiederholung in der Grundschule: Ein mehrbenenanalytischer Beitrag zur Diskussion um die Effektivität von Klassenwiederholungen. Unveröffentlichte Diplomarbeit. Technische Universität Dortmund.

Raudenbush, S. W., Bryk, A. S. & Congdon, R. (2009). HLM 6.08 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Statistisches Bundesamt (Hrsg.) (2012). Schulen auf einen Blick. Wiesbaden: Statistisches Bundesamt.

Schwerdt, G. & West, R. (2013): The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. Harvard Kennedy School: Cambridge.

## ***Einzelwahl im Fokus: Die Folgen der freien Schulwahl und der Zweigliedrigkeit auf das Schulwahlverhalten von Eltern am Beispiel Berlins***

Gegenwärtig erfolgt in vielen Bundesländern eine Umgestaltung der Schulstruktur in der Sekundarstufe auf ein zweigliedriges Schulsystem. In Berlin tritt neben das Gymnasium eine integrierte Sekundarschule. Beide Schulen werden als „gleichwertig“ bezeichnet, da auf beiden Schulformen das Abitur erreicht werden kann. Zeitgleich findet eine Umstellung des Übergangsverfahrens statt, bei dem Eltern sich drei Wunschschulen frei wählen können. Sollten Schulen mehr Anmeldungen als verfügbare Plätze haben, dann gelten sie als übernachgefragt und können sich einen Teil der Schüler auswählen. Die restlichen Plätze werden verlost. Diese Änderungen der Schulstruktur und Übergangsregelungen beeinflussen das elterliche Schulwahlverhalten am Ende der Primarstufe. Während in einem strukturell nach Leistungsanforderungen und -abschlüssen differenzierten Schulsystem die Wahl der Schulform im Zentrum von elterlichen Bildungsentscheidungen stand, ändert sich die Angebotssituation, wenn man nun auf allen Schulen das Abitur erreichen kann. Die Einzelschule rückt in den Mittelpunkt und ergänzt, wenn nicht sogar ersetzt, die Wahl einer Schulform. Die internationale Forschung zeigt uns, wie beim Wegfall des unterscheidenden Merkmales der Schulform, zunehmend die Schülerzusammensetzung, das schulische Profil, der Standort, die Qualität der Schule und die Wohnortnähe bedeutsam für die Wahlentscheidung werden. Die wenigen deutschen Studien, die beim Bildungsübergang die Einzelschule in den Blick genommen haben, fokussieren die Entscheidung auf Ebene der individuellen Merkmale mit starker Konzentration auf Präferenzen der Eltern. Schulen stellen aber unterschiedliche Lernkontexte dar, die gerade in nach sozio-ökonomischen Merkmalen stark polarisierten Ballungszentren, förderlich oder nachteilig wirken können. Eltern wählen also nicht nur nach ihren Präferenzen, sondern müssen auch aus dem städtischen Bildungsraum, die Schule finden, die ihrer Meinung nach den größten Bildungserfolg für das Kind verspricht. Deshalb widmet sich dieser Beitrag der Frage, welche Eigenschaften der Einzelschule in Kombination mit den Merkmalen der Familie die Auswahl von Einzelschulen bestimmt. Die Schätzungen werden mit Conditional Logit Modellen durchgeführt, die einerseits eine Modellierung der Schuleigenschaften zusammen mit den Eigenschaften der Eltern erlauben und es andererseits ermöglichen die Einzelschulen als untersuchendes Kriterium zu verwenden. Die Daten stammen aus dem Projekt „Schulwahl in Berlin“ und umfassen 360 Familien aus vier Berliner Bezirken, von denen über zwei Drittel eine Zuwanderungsgeschichte besitzen. Zusätzlich wurden Daten über die Eigenschaften von Schulen und ihrer Umgebung gesammelt und mit Hilfe des Geoinformationsprogrammes ArcGis die Distanzen der Familien zu allen Berliner Schulen errechnet. Die Analysen zeigen, dass die bisherigen Indikatoren der Übergangsforschung, wie z.B. das Vorhandensein einer gymnasialen Oberstufe vor Ort immer noch ein wichtiger Prädiktor für die Schulwahl ist. Während Prädiktoren, die erst durch das neue Übergangsverfahren in den Blick genommen werden können, wie die Nachfrage einer Schule, je nach Noten der Kinder und Migrationsstatus unterschiedlich bedeutsam werden. Insgesamt zeigen die Ergebnisse, wie die individuellen Merkmale der Schüler und die strukturellen und kontextuellen Merkmale der Einzelschule miteinander interagieren und dadurch die Schulwahlentscheidungen von Eltern vorhersagen können.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G150

## ***Modeling Predictions of Secondary School Track Appropriateness***

### Theoretical background

Making a decision or judgment often involves integrating diverse pieces of information. For instance, teachers may place students in a specific ability group depending on their past achievement, behavior, and perceived motivation. But how exactly do people integrate such cues in order to make judgments? Are there ways of improving cue integration to make better judgments?

One approach to answering these questions lies in mathematically modeling the cue-integration process. The underlying assumption is that creating an algorithm connecting cues and judgment may help predict judgment in new situations, estimate the importance of specific cues, and make judgments more consistent, efficient, accurate, and fair (e.g. Meehl, 1954, 1963).

### Research questions

One of the most important judgments affecting children's educational career is how they are grouped in the school system. In Luxembourg, children are grouped into three separate school tracks at the end of sixth grade. This grouping helps determine the academic degree with which they leave school, and thus their future job opportunities. The present study drew upon real data showing Luxembourgish students' placements and subsequent achievement in order to explore whether formal models of cue integration can be applied to improve the tracking decision in predicting future school success.

### Method

Using the data available to teachers and administrators making track recommendations, the study calculated a logistic regression model of the tracking decision (optimal weights). Additionally, a model which weighted all these sixth grade predictors equally (equal weights) and a model which included only predictors shown to be significant in the optimal regression model weighted equally (reduced equal weights) were calculated. All three models were applied in order to create "alternative" track recommendations. These alternative recommendations were indicators of how well each model was able to predict real track recommendations in a holdout sample. Calculated for the whole sample, however, they also allowed for a comparison of the appropriateness of different model predictions and real track recommendations. Recommendations' appropriateness was measured in terms of their accuracy in predicting students' academic success in the three years following the tracking decision.

### Results

Results showed that all models predicted actual track recommendations quite accurately (>90% correct classifications), though the optimal regression model slightly outperformed the two equal weights models. In terms of track appropriateness, however, the equal weights model predictions correlated more strongly with later student grade point average than both the optimal weights model and real track recommendations. Comparison of mean grades for "controversial" cases (=students for whom models disagreed on correct track placement) echoed these results, with the equal weights model showing superior or equivalent performance in comparison to all other models as well as to real recommendations. Predictive validity was also inferred on the basis of "incorrect" tracking decisions (i.e. placement of students in tracks whose requirements they failed to meet by either repeating a grade or transferring to another track). Pairwise comparisons of models in terms of unique incorrect tracking decisions (i.e. incorrect decisions made by one model but not the other) again showed slight superiority of the equal weights model in avoiding incorrect placement recommendations.

As a whole, results suggested that formally modeling track recommendations using an equal weights combination of sixth grade predictor variables may lead to slightly better tracking decisions than sole reliance on human judges. This conclusion echoes research showing that almost any type of formalized integration process leads to better judgments than informal integration of disparate cues—even among experts (e.g. Grove, Zald, Lebow, Snitz, & Nelson, 2000). Furthermore, simply standardizing and summing theoretically determined predictor variables may be a better modeling strategy than trying to determine "optimal" linear weights for individual predictors.

Chair(s): Daniela J. Jäger (Institut für Schulentwicklungsforschung Dortmund)

Diskutant/in: Ariane Willems (Institut für Schulentwicklungsforschung Dortmund)

#### Theoretischer Hintergrund

Länder mit einem hohen Grad an Standardisierung und Stratifizierung, wie Deutschland, weisen einen engen Zusammenhang zwischen sozialer Herkunft und Bildungserfolg von Schüler/-innen auf (vgl. Becker & Lauterbach, 2010). Im Projekt „Ganz In - Mit Ganzttag mehr Zukunft. Das neue Ganzttagsgymnasium NRW“ führen 30 Gymnasien den gebundenen Ganztagsbetrieb ein und reformieren ihre Lernkultur. Das Projekt zielt darauf, Schüler/-innen mit ungünstigen äußeren Rahmenbedingungen den Übergang und das Bestehen am Gymnasium zu ermöglichen. Durch den Ganzttag sollen die Vereinbarkeit von Familie und Beruf erleichtert, Ressourcen für individuelle Förderung bereitgestellt (vgl. Holtappels, 2006) und die Leistungen verbessert werden (vgl. Berkemeyer et al., 2010). Ganz In, verstanden als Schul- und Unterrichtsentwicklungsprojekt, erprobt einen inhaltlich entwickelten Ganzttag am Gymnasium.

#### Fragestellung

Wie können Organisations-, Personal-, Unterrichtsentwicklung und das außerunterrichtliche Angebot an Schulen hinsichtlich individueller Förderung und verbesserter Chancengerechtigkeit gefördert werden?

#### Methode

Über sechs Jahre werden die Kompetenzen der Schüler/-innen, ihre Wahrnehmung der Unterrichtsqualität, das Lehrberufswissen und die Entwicklung von Förderkonzepten kontinuierlich evaluiert. Zudem werden Schulleitungen, Eltern und weiteres Personal befragt. Im Schuljahr 2010/11 wurden Schüler/-innen der Klasse 5 erstmals und im Schuljahr 2012/13 in der 7. Klasse erneut befragt. Eine zweite Kohorte der Klasse 5 nahm im Schuljahr 2012/13 an der Befragung teil. Zusätzlich fanden Interventionsstudien und Erhebungen der Fachdidaktiken, statt.

#### Perspektivenvielfalt

Um Schul- und Unterrichtsentwicklung adäquat zu fördern, sind Akteure verschiedener Disziplinen beteiligt:

Die Einstellungen von Lehrkräften zum gebundenen Ganzttag und ihre Bedeutung für die Umsetzung individueller Förderung Susanne Lindemann, Birte Glesemann, Daniela J. Jäger und Wilfried Bos. Eine quantitative Erhebung aus pädagogischer und governancetheoretischer Sicht.

Ein Training für Lehrpersonen zur Vermittlung des selbstregulierten Lernens aus Sachtexten und durch Experimentieren Ferdinand Stebner, Silke Schiffhauer, Benjamin Klein, Annett Schmeck, Detlev Leutner und Joachim Wirth. Eine Interventionsstudie aus pädagogisch-psychologischer Sicht.

Mehr lernen durch eine bessere Sequenzierung des Physikunterrichts bei verlängerter Unterrichtstaktung Simon Zander, Heiko Krabbe und Hans E. Fischer. Eine Interventionsstudie der Fachdidaktik Physik.

Außerunterrichtliche Fördermöglichkeiten der naturwissenschaftlich-experimentellen Arbeitsweisen Jenna Koenen, Markus Emden und Elke Sumfleth. Eine Interventionsstudie der Fachdidaktik Chemie.

#### Ergebnisse

Die Ergebnisse der ersten beiden Vorträgen liegen erst zur Tagung vor, die Fachdidaktik Physik zeigt auf, dass im Physikunterricht durch eine lernprozessorientierte Sequenzierung nach den Oserschen Basismodellen eine Verbesserung der Unterrichtsqualität bei einer auf 90 Minuten verlängerten Taktung erreicht werden kann. In der Studie der Chemiedidaktik ist es gelungen, Lernmaterialien zu entwickeln und aufzuzeigen, welche Art von Lernmaterial in Bezug auf unterschiedliches Vorwissen der Schüler/-innen lernförderlich ist.

Becker, R. & Lauterbach, W. (Hrsg.) (2010). Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit. Wiesbaden: VS Verlag.

Berkemeyer, N., Bos, W., Holtappels, H.G., Meetz, F. & Rollet, W. (2010). "GanzIn": das Ganzttagsgymnasium in Nordrhein-Westfalen. In N. Berkemeyer, W. Bos, H. G. Holtappels, N. McElvany, & R. Schulz-Zander (Hrsg.): Jahrbuch der Schulentwicklung Band 16 (S. 131-152). Weinheim und München: Juventa.

Holtappels, H. G. (2006). Ganzttagsschule – ein Beitrag zur Förderung und Chancengleichheit? In K. Höhmann & H. G. Holtappels (Hrsg.), Ganzttagsschule gestalten: Konzeption, Praxis, Impulse (S. 10-33). Seelze-Velbert: Kallmeyer Verlag in Verbindung mit Klett.

Mittwoch,

05.03.2014,

11:15 - 13:00,

IG-Farben-Haus

Raum 0.251



## **Die Einstellungen von Lehrkräften zum gebundenen Ganzttag und ihre Bedeutung für die Umsetzung individueller Förderung**

Der Ausbau von Schulen zu Ganztagschulen, welcher mittlerweile sukzessive gestiegen ist, ist eine bildungspolitische Konsequenz in Bezug auf die Ergebnisse deutscher Schüler/-innen in großen Schulleistungstudien, wie PISA (vgl. Berkemeyer et al., 2010). Dieser Schritt ist bedeutend, da mit dem Ausbau des Ganztagsbetriebs an Schulen das Ziel verbunden ist, eine neue Lernkultur zu entwickeln, welche unter anderem auf die individuelle Förderung von Schüler/-innen ausgerichtet ist (vgl. ebd.).

### Theoretischer Hintergrund

Die Einführung des Ganztags ist durchaus als eine Reform zu bezeichnen, die als Top-down-Prozess durchgesetzt wurde. Wie bei vielen solcher Reformen im schulischen Bereich ist ihre Umsetzung abhängig von der Akzeptanz aller beteiligten Akteure (vgl. Oerke, 2012; Gräsel, 2010; Pant et al., 2008), hier insbesondere der Lehrpersonen. Ein Modell, welches die Einstellungen von Lehrkräften in Innovationsprozessen oder auch im Umgang mit neuen Reformen betrachtet, ist das Concerns-Based Adoption Modell von Hall und Hord (2006). Grundannahme des Modells ist ein stufenweiser Prozess der Übernahme einer Innovation (vgl. Gräsel, 2010; Hall & Hord, 2006), welcher sich in sieben aufeinanderfolgenden Stufen (Bewusstsein, Information, persönliche Betroffenheit, Aufgabenmanagement, Auswirkungen auf die Lernenden, Kooperationsmöglichkeiten, Revision und Optimierung) vollzieht (vgl. ebd.). Mit den Stages of Concern liegt der Bildungsforschung ein Instrument vor, mit welchem die Einstellungen von Lehrer/-innen erfasst (vgl. Gräsel, 2010) und Muster in der Auseinandersetzung mit Innovationen ermittelt werden können. Im Idealfall bilden sich nach Hall und Hord (2006) dabei zwei Typen ab, und zwar die Innovationsgegner und Kooperationsbereiten. Je nachdem, welche Innovation mit den Stages of Concern verbunden ist, werden die Einstellungsmuster in der empirischen Arbeit weiter spezifiziert. In der Studie von Pant und Kollegen (2008) wurden beispielsweise die Profile der selbstorientierten Kooperierer und der informationssuchenden Kooperierer ohne Aufgabenfokus eruiert.

### Fragestellungen

Vor dem Hintergrund, dass die individuelle Förderung von Schüler/-innen als wichtiges Element und auch Ziel der Ganztagschulentwicklung bezeichnet werden kann (vgl. Holtappels, 2006), liegt diesem Beitrag die Annahme zugrunde, dass je höher die Auseinandersetzung mit dem Ganzttag von Lehrkräften beurteilt wird, auch die Verwendung von Instrumenten individueller Förderung steigt. Der Beitrag geht somit den Frage nach: 1) Welche Typen von Lehrkräften lassen sich hinsichtlich ihrer Einstellung zum gebundenen Ganzttag identifizieren? 2) Bleiben diese Lehrertypen über zwei Jahre Ganztagsentwicklung konstant? Und letztlich soll aufgezeigt werden, 3) ob ein Zusammenhang zwischen der Einstellung von Lehrpersonen zum Ganzttag, gemessen an den Stages of Concern, und ihrer Verwendung von Instrumenten zur individuellen Förderung von Schüler/-innen besteht.

### Methode

Im Rahmen des Projekts Ganz In wurden Lehrkräfte von 30 ausgewählten Gymnasien zu ihren Einstellungen zum gebundenen Ganzttag befragt. An der ersten Befragung im Sommer 2011 nahmen 564 Lehrpersonen teil. Bei der Erhebung des zweiten Messzeitpunktes im Sommer 2013 wurden insgesamt 632 Lehrer/-innen befragt. Teil dieses Fragebogens sind unter anderem die an die Thematik der Unterrichtsgestaltung im gebundenen Ganzttag angepassten Stages of Concern, welche sich auf insgesamt sieben Stufen mit je vier Items verteilen. Mittels einer latenten Klassenanalyse (LCA), welche der Eruiierung von Typologien dienen (vgl. Bacher & Vermunt, 2010), werden Stages-of-Concern-Profile ermittelt und ihre Stabilität im zeitlichen Verlauf betrachtet. In einem nächsten Schritt gilt es, die unterschiedlichen Lehrertypen den Anwendungshäufigkeiten von Instrumenten individueller Förderung gegenüber zu stellen. Damit ist es möglich, Aussagen darüber zu treffen, wie sich die Einstellung zum Ganzttag und die Gestaltung der individuellen Förderung von Schüler/-innen zueinander verhalten.

### Ergebnisse

Die Berechnung einer LCA zum ersten Messzeitpunkt zeigt, dass sich vier Lehrertypen in den Daten abbilden. In einem nächsten Schritt gilt es, die Berechnungen zur zweiten Befragung vorzunehmen und die Verbindung zur Anwendung individueller Fördermaßnahmen herzustellen. Die Ergebnisse liegen zur GEBF-Tagung vor.

Bacher, J. & Vermunt, J. K. (2010). Analyse latenter Klassen. In C. Wolf & H. Best (Hrsg.), Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: VS Verlag für Sozialwissenschaften.

Berkemeyer, N., Bos, W., Holtappels, H.G., Meetz, F. & Rollet, W. (2010). "GanzIn": das Ganztagsgymnasium in Nordrhein-Westfalen. In N. Berkemeyer, W. Bos, H. G. Holtappels, N. McElvany, & R. Schulz-Zander (Hrsg.): Jahrbuch der Schulentwicklung Band 16 (S. 131-152). Weinheim und München: Juventa.

Gräsel, C. (2010). Stichwort Transfer und Transferforschung im Bildungsbereich. Zeitschrift für Erziehungswissenschaft 13 7-20.

Hall, G. E. & Hord, S. M. (2006). Implementing change: Patterns principles and pottholes (2nd ed.). Boston: Pearson Education.

Oerke, B. (2012). Auseinandersetzung der Lehrpersonen mit der Einführung des Zentralabiturs: Stages of concern. In K. Maag Merki (Hrsg.), Zentralabitur. Die längsschnittliche Analyse der Wirkungen der Einführung zentraler Abituprüfungen in Deutschland (S. 207-236). Wiesbaden: VS-Verlag.

Holtappels, H. G. (2006). Ganztagschule – ein Beitrag zur Förderung und Chancengleichheit? In K. Höhmann & H. G. Holtappels (Hrsg.), Ganztagschule gestalten: Konzeption, Praxis, Impulse (S. 10-33). Seelze-Velbert: Kallmeyer Verlag in Verbindung mit Klett.

Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. Zeitschrift für Pädagogik 54 (6) 827-845.





Ferdinand Stebner (Universität Bochum), Silke Schiffhauer (Universität Duisburg Essen), Benjamin Klein (Universität Bochum), Annett Schmeck (Universität Duisburg Essen), Detlev Leutner (Universität Duisburg Essen), Joachim Wirth (Universität Bochum)

## **Ein Training für Lehrpersonen zur Vermittlung des selbstregulierten Lernens aus Sachtexten und durch Experimentieren**

### Theoretischer Hintergrund

Selbstreguliertes Lernen wird als ein aktiver, konstruktiver Prozess verstanden, bei dem sich der Lernende selbst Ziele für sein Lernen setzt und zudem seine Kognitionen, seine Motivation und sein Verhalten in Abhängigkeit von diesen Zielen und den gegebenen äußeren Umständen beobachtet, reguliert und kontrolliert (Pintrich, 2000). Erkenntnisse vorangegangener Studien zeigen, dass es lernförderlich ist, kombiniert die Strategien zur Förderung des selbstregulierten Lernens mit kognitiven und metakognitiven Strategien wie den Methoden naturwissenschaftlichen Arbeitens (bspw. Experimentieren oder Lesen aus Sachtexten), zu trainieren (Leutner et al., 2001; Stebner et al., 2013).

### Fragestellung und Methode

Ziel des Projekts ist die Evaluation eines Trainings zum selbstregulierten Lernen aus Sachtexten und durch Experimentieren, das im Rahmen des Projektes Ganz In entwickelt und von vier Lehrkräften nordrhein-westfälischer Ganztags gymnasien im regulären Schulalltag in Form einer AG durchgeführt wurde.

In der ersten Phase dieses Projekts wurde je ein Training zum Lernen aus Sachtexten und zum Lernen durch Experimentieren für Fünft- und Sechstklässler entwickelt. Diese zwei Trainings wurden von Mitarbeitern der Universität durchgeführt und erfolgreich evaluiert (Klein et al., 2013; Stebner et al., 2013). Inhaltlich wurden die zwei Trainings in der hier vorgestellten zweiten Projektphase zu einem Training zusammengeführt, wobei die Vermittlung selbstregulativer Strategien beibehalten wurde. Das Training wird von Naturwissenschaftslehrkräften durchgeführt. Schülerinnen und Schüler (SuS) werden angeleitet, selbstreguliert aus einem Sachtext zu lernen, daraus eine Fragestellung zu entwickeln und diese anschließend selbstreguliert experimentell zu untersuchen.

In diesem Training werden mit den SuS zunächst Grundlagen des kooperativen Lernens erarbeitet, bevor eine neu entwickelte Testbatterie zur Diagnose als Prä-Erhebung eingesetzt wird. Diese misst in zwei aufeinanderfolgenden Doppelstunden à 90 Minuten schülerfreundlich – unter dem Vorwand, eine reguläre, interessante Unterrichtsstunde durchzuführen –, wie gut die SuS bereits Selbstregulations-, Lese- und Experimentierstrategien anwenden können. Darauf aufbauend werden in den folgenden vier Doppelstunden metakognitive Prozesse aktiviert, Motivationsregulationsstrategien gelernt, sowie Lese- und Experimentierstrategien in Kombination mit metakognitiven Strategien selbstregulierten Lernens trainiert. In den folgenden acht Doppelstunden, in denen die naturwissenschaftlichen Fachinhalte variieren, erhalten die SuS stets einen Text, welchen sie selbstreguliert erarbeiten, um eine Fragestellung zu entwickeln, die anschließend experimentell untersucht werden soll. Zum Abschluss jeder Doppelstunde ist es Aufgabe der SuS, die zurückliegende Stunde und ihre Leistung anhand von offenen Fragen wie „Was hat heute beim Lernen gut geklappt?“ zu reflektieren. Zudem werden dabei u. a. Daten zum Verständnis, der Motivation und der kognitiven Belastung mit geschlossenen Fragen erhoben. Im Sinne individueller Förderung werden diese Daten den Lehrkräften vor der nächsten Doppelstunde zurückgemeldet, damit sie dann die Möglichkeit haben, sensibler auf Bedürfnisse sowie Stärken und Schwächen der SuS einzugehen. Am Ende des Halbjahres finden ein Transfertraining sowie die Post-Erhebung der Testbatterie zur Diagnose statt.

Die vier teilnehmenden Naturwissenschaftslehrkräfte wurden vor Beginn des Halbjahres in zwei Workshops fortgebildet und während der Trainingslaufzeit im Sinne der Prozessbegleitung von Mitarbeitern des Projekts intensiv (auch vor Ort im Unterricht bspw. durch Teamteaching) betreut. Das Training wird im ersten Halbjahr des Schuljahres 2013/2014 mit vier Klassen (etwa 100 SuS) der fünften und sechsten Jahrgangsstufe durchgeführt. Zur Kontrollgruppe gehören etwa 100 SuS derselben Altersgruppe, die an anderen AG-Angeboten der jeweiligen Schulen teilnehmen. Zur Evaluation des Trainings hinsichtlich Erfolg und Praktikabilität wurden zusätzlich zu der Testbatterie zur Diagnose und den Daten zur individuellen Förderung vor Beginn des Trainings standardisierte Fragebögen eingesetzt. Diese im Prä-Post-Design sowohl bei Kontroll- als auch bei der Trainingsgruppe eingesetzten Fragebögen beinhalten Motivationsregulationsstrategiewissen, Fähigkeitsselbstkonzept, Interesse an Naturwissenschaften, Lern- und Leistungsmotivation und Wissen über selbstregulative sowie Lese- und Experimentierstrategien.

### Ergebnisse

Die Evaluationsergebnisse werden zur GEBF-Tagung vorliegen. Es wird erwartet, dass die SuS der Trainingsgruppen auf den eingesetzten Erhebungsinstrumenten eine positivere Entwicklung aufweisen als die SuS der Kontrollgruppen und somit die Wirksamkeit des Trainings nachgewiesen werden kann.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Handbook of Self-Regulation* (S. 451–502). San Diego: Academic Press.  
Leutner, D., Barthel, A. & Schreiber, B. (2001). Studierende können lernen, sich selbst zum Lernen zu motivieren. Ein Trainingsexperiment. *Zeitschrift für Pädagogische Psychologie*, 15 (3/4), 155–167.  
Klein, B., Stebner, F., Schmeck, A., Marschner, J., Leutner, D., & Wirth, J. (2013, April). Evaluation of a reading strategy training in a field setting. Poster presented at the annual meeting of the American Educational Research Association (AERA), 27.4.-1.5.2013, San Francisco, USA.  
Stebner, F., Klein, B., Marschner, J., Schmeck, A., Leutner, D. & Wirth, J. (2013, März). Die Wirkung selbstregulativer Strategien beim naturwissenschaftlichen Experimentieren. Paper presented at the 1st Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), 11.-13.3.2013, Christian-Albrechts-Universität, Kiel.

Simon Zander (Universität Duisburg Essen), Heiko Krabbe (Universität Duisburg-Essen),  
Hans E. Fischer (Universität Duisburg-Essen)

## **Mehr lernen durch eine bessere Sequenzierung des Physikunterrichts bei verlängerter Unterrichtstaktung**

Theoretischer Hintergrund: In Verbindung mit dem Ganztagsunterricht werden auch veränderte Zeitstrukturmodelle diskutiert (Bergmann & Fiegenbaum, 2009). Anstelle der üblichen 45-Minuten-Taktung der Stunden tritt z.B. eine Doppelstunden-Taktung, die zu veränderten Unterrichtschoreographien führt (Haenisch, 2011). In 45 Minuten ist es kaum möglich, Lernprozesse von Schülern durch den Transfer oder die Vernetzung von Inhalten abzuschließen (Borowski, Fischer, Trendel, & Wackermann, 2010). Eine verlängerte Taktung ermöglicht eine verbesserte lernprozessorientierte Sequenzierung des Unterrichts. Die Basismodelle von Oser und Baeriswyl (2001) sind eine mögliche Beschreibung einer auf die verlängerte Taktung angepassten lernprozessorientierten Sequenzierung. Rektor und Wackermann (2011) haben jedoch festgestellt, dass die Verlängerung der Taktung nicht unbedingt zu einer vollständigeren Realisierung der Basismodelle führt, die insbesondere Transfer-, Vernetzungs- oder Reflexionsphasen enthält. In einer Lehrerfortbildung im Rahmen des Ganz-In-Projekts wurden Lehrkräfte bezüglich einer optimalen Sequenzierung im Sinne der Basismodelle gecoacht und die Auswirkung auf die Lernzuwächse der Schüler erfasst (Zander, Krabbe, & Fischer, 2013). Dabei konnte auch Unterricht mit unterschiedlicher Taktung verglichen werden. Eine verbesserte Strukturierung des Unterrichts sollte insbesondere schwächeren Schülern ein besseres Lernen ermöglichen (Helmke, 2009, S. 209).

Fragestellungen: 1. Inwiefern ermöglicht eine verlängerte Taktung eine verbesserte lernprozessorientierte Sequenzierung im Sinne der Basismodelle?

2. Welche Auswirkung hat eine verbesserte lernprozessorientierte Sequenzierung auf den Lernzuwachs insbesondere bei schwächeren Schülern?

Methode: Im Schuljahr 2011/12 nahmen 15 Gymnasiallehrkräfte an einer einjährigen Lehrerfortbildung zur lernprozessorientierten Sequenzierung mit Videofeedback teil. Zehn der 15 Lehrpersonen unterrichteten in 90-minütigen, die anderen fünf Lehrkräfte in 45- oder 60-minütigen Physikstunden. Je Lehrkraft wurde in jedem Quartal des Schuljahres ein Unterrichtsvideo im Physikunterricht einer 8. Klasse aufgezeichnet. Nach dem ersten Video fand ein eintägiger Fortbildungstag über Basismodelle statt. Außerdem wurden die ersten drei Videos mit einem Coach nachbesprochen und die Unterrichtsstunden für das 2. und 3. Video gemeinsam geplant. Alle Videos wurden manualbasiert kodiert (Cohens Kappa zwischen 0,92 und 1). Veränderungen in der Sequenzierung nach Basismodellen bei den Lehrkräften wurden anhand des ersten Videos (Vortest) und des letzten Videos (Nachtest) analysiert. Zu Beginn und zum Ende des Schuljahres wurde sowohl in der Fortbildungsgruppe als auch in einer Vergleichsgruppe mit 15 Parallelklassen, deren Lehrkräfte nicht an der Fortbildung teilgenommen hatten, ein Fachwissenstest zur Mechanik durchgeführt, um den Lernzuwachs zu bestimmen. Der Fachwissenstest wies für die Schüler der Fortbildungsgruppe (n=380) und der Vergleichsgruppe (n=366) zusammen im Vortest eine Reliabilität von Cronbach's  $\alpha=0.48$  und im Nachtest von  $\alpha=0.66$  auf. Ergebnisse: Die Lehrkräfte der Fortbildungsgruppe mit 90-minütigen Unterrichtsstunden zeigten nach der Fortbildung eine vollständigere Umsetzung der Basismodelle (Median = 100%) als Lehrkräfte mit einer geringeren Unterrichtszeit (Median = 80%),  $W=3$ ,  $p<0,05$ ,  $\Phi=0,75$ . Für die Leistungszuwächse zeigen sich ähnlich große Unterschiede zwischen den beiden Gruppen (Median 45/60 Minuten = 7%, Median 90 Minuten = 10,5%),  $W=8,5$ ,  $p<0,05$ ,  $\Phi=0,51$ . In der Vergleichsgruppe zeigte sich kein signifikanter Unterschied im Lernzuwachs bzgl. der Taktung.

Die Vollständigkeit der Videos im Nachtest korrelierte mit den Lernzuwächsen der Klassen,  $\rho=0,52$ ,  $p<0,05$ . Dies weist darauf hin, dass eine vollständigere Strukturierung zu besserem Lernen führt.

Mit einer latenten Klassenanalyse konnten für Schüler der Vergleichs- und Fortbildungsgruppe jeweils in zwei Leistungsgruppen aufgeteilt werden. Dabei war der Anteil an Schülern in der leistungsstärkeren Gruppe für die Fortbildungsgruppe größer (54%) als für die Vergleichsgruppe (47%). Vor allem die leistungsschwächeren Schüler der Fortbildungsgruppe schnitten besser ab als in der Vergleichsgruppe.

Die Ergebnisse weisen für den Physikunterricht darauf hin, dass durch eine lernprozessorientierte Sequenzierung nach den Basismodellen eine Verbesserung der Unterrichtsqualität bei einer auf 90 Minuten verlängerten Taktung erreicht werden kann. Wie erwartet konnten besonders schwächere Schüler von der verbesserten Sequenzierung profitieren.

Bergmann, H.-P., Fiegenbaum, D. (2009) Rhythmisierung und Zeitstrukturmodelle im Ganztag. In K. Althof (Hrsg.) Der Ganztag in der Sekundarstufe I. Der Ganztag in NRW, Heft 12, S. 9-16. Münster: Serviceagentur „Ganztägig lernen in Nordrhein-Westfalen“.

Borowski, A., Fischer, H.E., Trendel, G. & Wackermann, R. (2010). Guter Fachunterricht braucht seine Zeit. Pädagogik, 62 (3), 26-29.

Haenisch, H. (2011). Gebundene Ganztagschule – Ansätze zur Gestaltung. Der Ganztag in NRW, Heft 12, S. 9-16. Münster: Serviceagentur „Ganztägig lernen in Nordrhein-Westfalen“.

Helmke, A. (2009). Unterrichtsqualität und Lehrerprofessionalität, Evaluation und Verbesserung der Unterrichtsqualität. Seelze: Klett-Kallmeyer.

Oser, F.K. & Baeriswyl, F.J. (2001). Choreographies of Teaching: Bridging Instruction of Learning. In V. Richardson (Ed.), Handbook of research on teaching (4th ed.) (pp. 1031-1065). Washington: American Educational Research Association.

Rektor, J. & Wackermann, R. (2012). Der Einfluss von 60-Minuten-Stunden auf den Physikunterricht – Vorstellung des Forschungsvorhabens und erste Ergebnisse. In: S. Bernholt (Hrsg.), Konzepte fachdidaktischer Strukturierung für den Unterricht. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Oldenburg 2011. Berlin: LIT-Verlag, 176-178.

Reyer, T. (2004). Oberflächenmerkmale und Tiefenstrukturen im Unterricht. Berlin: Logos.

Wackermann, R., Trendel, G. & Fischer, H.E. (2010). Evaluation of a Theory of Instructional Sequences for Physics Instruction. International Journal of Science Education, 32 (7), 963-985.

Zander, S., Krabbe, H. & Fischer, H. E. (2013). Lernzuwächse in der Mechanik im Rahmen der Lehrerfortbildung „Sequenzierung von Lernprozessen“. In: S. Bernholt (Hrsg.), Inquiry-based Learning – Forschendes Lernen. Kiel: IPN, 503-505.



## **Außerunterrichtliche Fördermöglichkeiten der naturwissenschaftlich-experimentellen Arbeitsweisen**

Theoretischer Hintergrund: Innerhalb der Struktur der Ganztagschulen gibt es viele Gelegenheiten, welche Schülerinnen und Schülern Möglichkeiten zum individuellen und selbstregulierten Lernen bieten. Dies geschieht vor allen Dingen in zusätzlichen Lernzeiten oder Nachmittagsangeboten. Oft mangelt es den Lehrpersonen jedoch an entsprechendem Material, um diese Lerngelegenheiten zu gestalten. Dieses Projekt soll einen Beitrag zu Gestaltung dieser Lerngelegenheiten leisten. Es wurden verschiedene Lernmaterialien entwickelt und evaluiert (z.B. experimentunterstützte Lösungsbeispiele). Da die PISA-Ergebnisse gezeigt haben, dass Schülerinnen und Schüler in Deutschland Wissensrückstände im Bereich des naturwissenschaftlich-experimentellen Arbeitens besitzen (OECD, 2007), liegt der Fokus des entwickelten Lernmaterials auf diesem Bereich. Es kann als Vorbereitung auf den Chemieunterricht eingesetzt werden, sodass die Schülerinnen und Schüler selbstständig erstes Wissen in diesem Bereich erarbeiten können. Fragestellung: Die zentrale Frage, die sich ergibt, ist, welche Lernmaterialien sich besonders gut für Schülerinnen und Schüler mit bestimmten Voraussetzungen (z.B. Vorwissen) eignen.

Methode: Die Materialien wurden im Rahmen eines Prä-, Post- und Follow-up-Designs mit fünf Versuchsgruppen und  $N = 285$  Schülerinnen und Schülern evaluiert. Die Intervention erfolgte über einen Zeitraum von sechs Wochen. Zunächst erhielten alle Lernenden ein Training zum naturwissenschaftlich-experimentellen Arbeiten. Innerhalb der folgenden fünf Wochen lernten die Gruppen mit verschiedenen Lernmaterialien mit Fokus auf den naturwissenschaftlichen Experimentierprozess. Grundsätzlich lernten die Schülerinnen und Schüler entweder mit einem Lösungsbeispiel, einer Problemstellung mit Musterlösung (vgl. Renkl et al., 2000), oder einer offenen Problemlöseaufgabe. Je nach Gruppenzugehörigkeit wurden die Lernmaterialien mit hands-on Experimenten verknüpft. Die Kontrollgruppe lernte mit zwei experimentellen Problemlöseaufgaben, die Experimentalgruppe 1 mit zwei experimentunterstützten Lösungsbeispielen. Die Schülerinnen und Schüler der Experimentalgruppe 2 bearbeiteten ein theoretisches Lösungsbeispiel und anschließend ein experimentunterstütztes Lösungsbeispiel. Zwei weitere Gruppen arbeiteten mit Example-Problem-Pairs, die sich nur darin unterschieden, ob das Lösungsbeispiel theoretisch (EG 3) oder experimentunterstützt (EG 4) war. Im Anschluss folgte jeweils eine experimentelle Problemlöseaufgabe.

In den Testungen wurden Multiple-Choice-Single-Select-Items zur Erhebung des Wissens im Bereich des naturwissenschaftlich-experimentellen Arbeitens (Strukturierungstest: adaptiert nach Wahser, 2007; NAW-Test: adaptiert nach Mannel, 2011) und des Fachwissens eingesetzt. Leseverständnis, kognitive Fähigkeiten und Interesse an Naturwissenschaften wurden als Kontrollvariablen erhoben ebenso wie das situationale Interesse (vgl. Fechner, 2009; Klos, 2008) und der Cognitive Load (Perceived Item Difficulty: Kalyuga et al., 1999; Invested Mental Effort: Paas, 1992).

Ergebnisse: Die Ergebnisse der Studie zeigen, dass das Lernmaterial einen Einfluss auf den Lernerfolg hat. Des Weiteren scheinen verschiedene Lernmaterialien unterschiedlich gut für Lernende mit viel bzw. wenig Vorwissen geeignet zu sein. So konnte gezeigt werden, dass experimentunterstützte Lösungsbeispiele (EG 1) sowie ein Example-Problem-Pair (EG 3), das aus einem traditionellen Lösungsbeispiel mit einer sich anschließenden experimentellen Problemlöseaufgabe besteht, in Bezug auf den Lernerfolg im Bereich des naturwissenschaftlich-experimentellen Arbeitens (NAW-Test) besser für Lernende mit wenig Vorwissen geeignet sind als für Lernende mit höherem Vorwissen. Dies zeigen die Ergebnisse zweier Kovarianzanalyse mit drei Messwiederholungen unter Kontrolle von Vorwissen und Leseverständnis. Zur Auswertung wurden die Probanden je nach Vorwissenslevel in zwei Gruppen eingeteilt (wenig Vorwissen:  $F(4,198)=7.80$ ,  $p<.001$ ,  $etap2=.242$ ; größeres Vorwissen:  $F(4,110)=6.83$ ,  $p<.001$ ,  $etap2=.199$ ). Des Weiteren zeigt die Post hoc-Analyse, dass sich experimentelle Problemlöseaufgaben sowohl für Lernende mit wenig Vorwissen als auch für Lernende mit mehr Vorwissen zu eignen scheinen. In Bezug auf das Erlernen von Fachwissen zeigt sich ein ähnliches Bild (wenig Vorwissen:  $F(4,100)=7.94$ ,  $p<.001$ ,  $etap2=.241$ ; größeres Vorwissen:  $F(4,106)=10.39$ ,  $p<.001$ ,  $etap2=.282$ ).

Insgesamt konnte im Rahmen dieser Studie Lernmaterial für zusätzliche Lerngelegenheiten in Ganztagschulen für die Naturwissenschaften entwickelt und evaluiert werden. Darüber hinaus ist es gelungen Aussagen darüber zu machen, welche Art von Lernmaterial besonders gut für Schülerinnen und Schüler mit wenig bzw. mit viel Vorwissen geeignet zu sein scheint, so dass diese Ergebnisse einen Beitrag zur individuellen Förderung leisten können.

Fechner, S. (2009). Effects of context oriented learning on student interest and achievement in chemistry education. Studien zum Physik- und Chemielernen: Vol. 95. Berlin: Logos.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing Split-attention and Redundancy in Multimedia Instruction. Applied Cognitive Psychology, (13), 351–371.

Klos, S. (2008). Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht: Der Einfluss eines integrierten Unterrichtskonzepts. Studien zum Physik- und Chemielernen: Vol. 89. Berlin: Logos.

Mannel, S. (2011). Assessing scientific inquiry: Development and evaluation of a test for the low-performing stage. Studien zum Physik- und Chemielernen: Vol. 111. Berlin: Logos Verlag.

OECD. (2007). PISA 2006: Science competencies for tomorrow's world. Volume 1 - Analysis. Paris: OECD.

Paas, G. F. W. C. (1992). Training strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. Journal of Educational Psychology, 84(4), 429–434.

Renkl, A., Atkinson, R. K., & Maier, U. H. (2000). Research Report No. 140: From Example Study to Problem Solving: Smooth Transitions Help Learning.

Wahser, I. (2007). Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie. Studien zum Physik- und Chemielernen: Vol. 73. Berlin: Logos-Verlag.



## Arbeitsgedächtniskapazität und Disfluency Effekt: Eine Aptitude-Treatment-Interaktionsstudie

Chair(s): Christina Goussios (Universität Ulm)

Theoretischer Hintergrund und Fragestellung: Gemäß der Cognitive Load Theory (CLT) sollte Lernmaterial so gestaltet werden, dass eine unnötige Arbeitsgedächtnisbelastung vermieden wird. Ziel ist somit die Reduzierung des extraneous load durch eine günstige Darstellung sowie eine Erhöhung des germane load durch die Förderung lernrelevanter Aktivitäten (Sweller, van Merriënboer, & Paas, 1998). Neuere Studien (Diemand-Yauman, Oppenheimer, & Vaughan, 2011; Kühl, Eitel, Scheiter, & Gerjets, 2011; Scheiter & Schüler, 2011) zeigen jedoch, dass eine zusätzliche Belastung in Form von schwer leserlichem Material den Lernerfolg positiv beeinflusst. Dieser sogenannte Disfluency Effekt soll darauf basieren, dass die Lernenden mehr Anstrengung in den Lernprozess investieren und das Lernmaterial tiefer verarbeiten. Allerdings gibt es auch Studien (Pieger, Mengelkamp, & Bannert, 2013), die den Disfluency Effekt nicht replizieren konnten, was darauf hindeutet, dass nicht alle Lernenden von Disfluency profitieren. Gemäß der CLT sollte die zusätzliche kognitive Belastung in Form von disfluentem Material durch eine hohe Arbeitsgedächtniskapazität kompensiert werden können. Um den kompensatorischen Effekt der Arbeitsgedächtniskapazität zu prüfen, wurde in der vorliegenden Studie die Aptitude-Treatment-Interaktion zwischen der Arbeitsgedächtniskapazität und Disfluency untersucht. Je höher die Arbeitsgedächtniskapazität ist, desto höher sollte die Lernleistung in der Disfluency-Bedingung sein. Außerdem wurden die Effekte für die verschiedenen cognitive load (CL) Arten untersucht, um anhand der CLT den genauen Wirkmechanismus des Disfluency Effekts zu prüfen. Da Lernhilfen - und wir verstehen Disfluency als eine solche - insbesondere bei Lernenden mit mittlerem Vorwissen wirksam sind (Seufert, 2003; Seufert & Brünken, 2004), sollten nur Probanden mit mittlerem Vorwissen berücksichtigt werden.

Methode: Insgesamt wurden 65 Studierende untersucht, von denen 47 ein mittleres Vorwissen aufwiesen und bei der Datenauswertung berücksichtigt wurden. Die Probanden wurden zufällig auf die Fluency- und Disfluency-Bedingung der unabhängigen Variablen Lernmaterial (Treatment-Faktor) aufgeteilt. Nach dem Lernen eines fluenten (Schrift in Arial, 12pt, schwarz) oder disfluenten (Schrift in Haettenschweiler, 12pt, grau 35%) Textes über „Zeit- und Datumsunterschiede auf der Erde“ (nach Schnotz & Bannert, 1999) wurden die Wissens-, Anwendungs- und Transferleistung sowie der intrinsic, extraneous und germane load (gemessen nach Klepsch, Wagner, & Seufert, in preparation) als abhängige Variablen erhoben. Abschließend wurde die Arbeitsgedächtniskapazität (stetiger Aptitude-Faktor) als zweite unabhängige Variable mit dem Numerical Memory Updating Test (Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000) erfasst. Als Kontrollvariablen dienten das Vorwissen, die Motivation, der Spaß und das Interesse.

Ergebnisse: Um die Interaktionseffekte auf die Lernleistung und den CL zu untersuchen, wurden multiple Regressionsanalysen mit der Recentering-Methode nach Aiken und West (1991) durchgeführt, die es ermöglicht, den spezifischen Einfluss der Arbeitsgedächtniskapazität auf die einzelnen Bedingungen zu untersuchen. Als Prädiktoren dienten die Versuchsgruppe (fluent, disfluent), die Arbeitsgedächtniskapazität und der Interaktionsterm Versuchsgruppe $\times$ Arbeitsgedächtniskapazität.

Der Interaktionsterm Versuchsgruppe $\times$ Arbeitsgedächtniskapazität sagte die Wissens- ( $\beta = 11.66$ ,  $t(44) = 2.46$ ,  $p = .02$ ) und Anwendungsleistung ( $\beta = 12.62$ ,  $t(46) = 2.33$ ,  $p = .03$ ) signifikant vorher. In der Disfluency-Bedingung beeinflusste die Arbeitsgedächtniskapazität die Wissens- ( $\beta = 10.66$ ,  $t(44) = 2.86$ ,  $p = .01$ ) und Anwendungsleistung ( $\beta = 10.66$ ,  $t(46) = 2.86$ ,  $p = .05$ ) signifikant: Je höher die Arbeitsgedächtniskapazität der Probanden, desto höher die Wissens- und Anwendungsleistung. In der Fluency-Bedingung beeinflusste die Arbeitsgedächtniskapazität die Wissens- und Anwendungsleistung nicht signifikant. Weder die Transferleistung noch die verschiedenen CL-Arten wurden durch den Interaktionsterm Versuchsgruppe $\times$ Arbeitsgedächtniskapazität signifikant vorhergesagt, sodass der Wirkmechanismus des Disfluency Effekts unklar bleibt.

Diskussion: Während Personen mit einer hohen Arbeitsgedächtniskapazität ein schlechtes Design in Form von disfluentem Lernmaterial kompensieren können und von Disfluency profitieren, scheint ein Mangel an Anregung in Form von fluentem Lernmaterial sogar hinderlich für diese Personen zu sein. Schließlich erzielten sie in der Fluency-Bedingung schlechtere Lernleistungen als in der Disfluency-Bedingung. Künftige Studien sollten zudem untersuchen, wann ein Text als disfluent aber noch nicht als unleserlich gilt. Lediglich disfluente Texte können die Lernleistung verbessern, während vollkommene Unleserlichkeit den Lernprozess behindert.

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition*, 118, 111-115.

Klepsch, M., Wagner, F., & Seufert, T. (in preparation). Differentiated Measurement of Cognitive Load: possible or not?

Kühl, T., Eitel, A., Scheiter, K., & Gerjets, P. (2011, August). Disfluency research meets cognitive load theory in multimedia learning: When a bad layout leads to good performance. 14th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI). Exeter, Great Britain. [Talk]

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity - facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017-1045.

Pieger, E., Mengelkamp, C., & Bannert, M. (2013, August). If it's hard to read - regulate yourself! When disfluency leads to better monitoring and performance. Poster presented at the JURE Preconference of the 15th Biannual EARLI Conference for Research on Learning and Instruction, Munich, Germany.

Scheiter, K., & Schüler, A. (2011, August). Fortune favors the bold (and the italicized), but why? Revisiting the disfluency effect. 14th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI). Exeter, Great Britain. [Talk]

Schnotz, W., & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für Experimentelle Psychologie*, 46, 217-236.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13, 227-237.

Seufert, T., & Brünken, R. (2004). Supporting coherence formation in multimedia learning. In H. Niegemann, R. Brünken, & D. Leutner (Hrsg.), *Instructional Design for Multimedia Learning*. Proceedings of the EARLI SIG 6 Biannual Workshop 2002 in Erfurt (S. 138-147). Münster: Waxmann.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
IG-Farben-Haus  
Raum 0.254





## **Beeinflusst die Arbeitsgedächtniskapazität das Navigationsverhalten von Grundschulkindern in einer iPad-Lernumgebung?**

Theoretischer Hintergrund und Fragestellung: Digitale Lerntechnologien sind aus dem Schulalltag nicht mehr wegzudenken. Viele dieser Lerntechnologien basieren dabei auf Hypermedia-Umgebungen, d.h. auf nicht-linearen multimedialen Informationsnetzwerken (Scheiter & Gerjets, 2007). Laut der Cognitive Flexibility Theory (Spiro & Jehng, 1990) sind solche hypermedialen Systeme aufgrund ihrer nichtlinearen Netzwerke besonders gut dafür geeignet, komplexe, facettenreiche Themengebiete darzustellen, die man aus verschiedenen Perspektiven betrachten kann (z.B. naturwissenschaftliche Phänomene). Auf der anderen Seite stellt die hypermediale bzw. multiperspektivische Darstellung von Lernmaterial auch hohe kognitive Anforderungen an den Lerner. So beanspruchen besonders die ungewohnten Navigationsanforderungen, die der Wechsel zwischen verschiedenen Perspektiven und Informationsangeboten erfordert, Arbeitsgedächtnisressourcen (Diamond, 2013). Dementsprechend wurde bei Schulkindern bereits festgestellt, dass sich hohe Arbeitsgedächtniskapazität positiv auf Hypermedia-Lernen auswirkt (Pazzaglia, Toso & Cacciamani, 2008). Ob dieser Zusammenhang auf ein spezifisches Navigationsverhalten zurückgeführt werden kann, das bei Kindern mit höherer Arbeitsgedächtniskapazität zu höherem Lernerfolg führt, ist allerdings unklar. So sollte in der vorliegenden Studie untersucht werden, ob sich das Navigationsverhalten von Kindern mit hoher Arbeitsgedächtniskapazität systematisch von denen mit geringerer Arbeitsgedächtniskapazität unterscheidet, und dadurch möglicherweise Unterschiede im Lernerfolg erklären kann.

Methode: Viertklässler/innen (N = 97) aus verschiedenen Grundschulen in Baden-Württemberg nahmen an der aus zwei Teilen bestehenden Studie teil. Im ersten Teil wurde das Arbeitsgedächtnis der Kinder anhand drei verschiedener Tests, dem Listening Span Test ( $\alpha = .74$ ), dem Spatial Span Test ( $\alpha = .77$ ) und dem N-back Test ( $\alpha = .76$ ), erfasst. Im zweiten Teil der Studie bearbeiteten die Kinder 45 min lang eine multiperspektivische Hypermedia-Lernumgebung auf dem iPad. Hierbei sollte den Kindern das komplexe Thema ‚Biodiversität von Fischen‘ vermittelt werden. Die Lernumgebung enthielt neben einer Übersichtsseite über alle verfügbaren Fisch-Informationen spezifische Informationen über die in der Lernumgebung vorhandenen Fische (als Text und Video) und verschiedene Fisch-Dimensionen (Größe, Gemeinschaftsverhalten, Speiseplan, Schwimmverhalten, Lebensraum). Die Fisch-Dimensionen bestanden aus jeweils drei Subkategorien (z.B. Lebensraum: Fluss – Mittelmeer – Tropisches Korallenriff), auf die sich alle vorhandenen Fische je nach Zugehörigkeit verteilten, sodass es dem Lerner möglich war, die Verteilung der Fische aus unterschiedlichen Blickwinkeln zu betrachten. Anhand von Explorationsfragen (z.B. „Worin unterscheiden sich der Nasenfisch und der Doktorfisch?“) wurden die Kinder durch die Lernumgebung geführt. Nach Beendigung der Lernphase bearbeiteten die Kinder einen Posttest. Dieser sowie die Antworten in den Explorationsfragen dienten als abhängige Variablen.

Ergebnisse und Interpretation: Arbeitsgedächtniskapazität konnte die Leistung sowohl in den Explorationsfragen als auch im Posttest signifikant vorhersagen ( $ps < .001$ ) und unterstreicht damit erneut die Relevanz von Arbeitsgedächtnisressourcen für erfolgreiches Lernen mit Hypermedia bei Kindern. Analysen in Bezug auf das Navigationsverhalten konnten zeigen, dass (i) die Anzahl an Klicks zwischen den verschiedenen Fisch-Dimensionen signifikant positiv ( $p = .008$ ) und (ii) die Dauer auf der Übersichtsseite signifikant negativ ( $p = .039$ ) mit Arbeitsgedächtniskapazität zusammenhängen, d.h. dass Kinder mit höherer Arbeitsgedächtniskapazität häufiger zwischen den verschiedenen Dimensionen hin- und herwechselten und gleichzeitig weniger Zeit auf der Übersichtsseite verbrachten als Kinder mit geringerer Arbeitsgedächtniskapazität. Darüber hinaus ergab eine Mediationsanalyse (Preacher & Hayes, 2008), dass der Zusammenhang zwischen Arbeitsgedächtniskapazität und Lernerfolg durch die beiden Navigationsmuster signifikant mediiert wurde (CI 95% [0.03, 0.16]). Dies lässt darauf schließen, dass der auf Arbeitsgedächtniskapazität zurückgehende Lernerfolg zumindest teilweise durch Unterschiede im Navigationsverhalten erklärt werden kann.

In weiteren Analysen wurde das Navigationsverhalten noch detaillierter betrachtet, wobei z.B. verschiedene Handlungsmöglichkeiten zur Beantwortung der Explorationsfragen kontrastiert wurden (z.B. ob Kinder eher Informationen gelesen oder ein Video angeschaut haben bzw. die Information über eine der Fisch-Dimensionen inferiert haben). Mögliche Navigationscluster sowie deren Zusammenhang mit Arbeitsgedächtniskapazität und Lernerfolg werden ebenfalls präsentiert. Als praktische Implikation ergibt sich, dass das Fördern effizienter Navigationsstrategien im Umgang mit Hypermedia-Umgebungen den damit verbundenen Lernerfolg im Schulkontext verbessern könnte.

Diamond, A. (2013). Executive Functions. *The Annual Review of Psychology*, 64, 135-168.

Pazzaglia, F., Toso, C., & Cacciamani, S. (2008). The specific involvement of verbal and visuospatial working memory in hypermedia learning. *British Journal of Educational Technology*, 39, 110-124.

Preacher, K. J. & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879-891.

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, 19, 285-307.

Spiro, R. J. & Jehng, J.-C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163-205). Hillsdale, NJ, England: Lawrence Erlbaum Associates.





## **Neue außerschulische Lernorte: Wie gehen Schüler und Schülerinnen mit einem Mangel an Struktur bzw. der Bereitstellung von Struktur in einer Mitmachausstellung um?**

Theoretischer Hintergrund: Außerschulische Lernorte stellen Lernumgebungen zu Verfügung, die erkundendes Verhalten unterstützen (Falk & Dierking, 2000). Der Erfolg von Besuchen in z.B. Museumsausstellungen wird unter anderem durch den „novelty-effect“ beeinflusst (Falk, Martin & Balling, 1978). Nach Gottfried (1980) zeigten Schüler und Schülerinnen (SuS) dort unterschiedlichste entdeckende Verhaltensweisen. Dabei suchten sie ihre Umgebung ungerichtet nach Stimulation ab. Das führt zu einem Zeitverzug bis SuS sich mit Ausstellungsstücken oder Aufgaben beschäftigen. Wird zu diesem Zeitpunkt keine Struktur bereitgestellt, werden SuS neben der Beschäftigung mit Ausstellungsstücken und dem Lösen von zugehörigen Fragen auch mit der Neuheit der Lernumgebung belastet. Die kognitive Belastung der SuS lässt sich als Cognitive Load (CL; Paas & Merriënboer, 1994) bestimmen. Skinner, Furrer, Marchand und Kindermann (2008) zufolge sind zur Unterstützung von Motivation neben optimaler Struktur auch autonomieunterstützende Instruktionen notwendig. Wahrgenommene Autonomieunterstützung, wahrgenommene Kompetenz und soziale Eingebundenheit sind die Voraussetzungen für das Entstehen von intrinsischer Motivation (Deci & Ryan, 2002). Struktur wird oft durch Lehrkräfte bereitgestellt (Jang & Reeve, 2006) und wird nicht selten kontrollierend dargeboten. In der vorliegenden Studie werden die Konstrukte Struktur und Kontrolle getrennt, indem die Bereitstellung von Struktur möglichst weitgehend vom Verhalten der Lehrkräfte separiert wird.

Forscherverfragen: Führt die Bereitstellung von Struktur zu...

- einer Erhöhung der wahrgenommenen Autonomie?
- einer Reduktion des wahrgenommenen CL in der Ausstellung?
- einer Beeinflussung der Schüler-Lehrer Interaktion?

Methode: 223 SuS der fünften und sechsten Jahrgangsstufe im Alter von 11,7(±0,6) Jahren besuchten eine Mitmachausstellung rund um das Thema Bewegung an der Universität für den Zeitraum von vier Stunden. Die Ausstellung unterteilte sich in drei Räume, wobei jeder Raum einem anderen Themenschwerpunkt zugeordnet war. Diese quasi-experimentelle Untersuchung nutzte ein Prä-Post-Test Design. Die Experimentalgruppe bestand aus 112 SuS (65 Jungen), die Kontrollgruppe aus 111 SuS (60 Jungen). Die Kontrollgruppe erhielt nur ein Basisniveau an Struktur, wohingegen die Experimentalgruppe zusätzliche Struktur in Form einer Vorbereitung auf den Besuch, strukturunterstützender Anweisungen, Erklärung der Arbeitsmaterialien und ihrer Benutzung erhielt. In beiden Gruppen arbeiteten SuS in unabhängigen Kleingruppen an einer Reihe von Arbeitsstationen. Bei den Lehrern handelte es sich um Biologiestudenten fortgeschrittener Semester. Alle Lehrer verhielten sich autonomieunterstützend. Sie griffen nur auf Nachfrage in die Aktivitäten der SuS ein, oder wenn es z.B. aufgrund von Sicherheitsgründen nötig war. Ansonsten hielten sie sich im Hintergrund. Als Testinstrumente wurde eine adaptierte Version der Basic Psychological Needs Scale (BPNS) für wahrgenommene Autonomie (Baard, Deci & Ryan, 2004) sowie eine adaptierte Version der Subskala "Hilfe/Unterstützung" des Teacher as Social Context Questionnaire (TASCQ; Wellborn, Connell, Skinner & Pierson, 1988). Der CL wurde gemäß Paas und Merriënboer (1994) gemessen.

Ergebnisse: Eine ANOVA mit Messwiederholung zeigte signifikante Unterschiede in der Autonomiesubskala des BPNS ( $F(1,195)=9,38, p<0,01, \eta^2=.05$ ) zwischen Prä- und Post-Test und, wie erwartet, keine Interaktion zwischen Zeit und Treatment ( $F(1,195)=1,29, p=ns$ ). Die Bereitstellung von Autonomieunterstützung führt also zu einem erhöhten wahrgenommenen Autonomieempfinden in beiden Treatmentgruppen. Die Subskala wahrgenommene Hilfe/Unterstützung des TASCQ zeigte signifikante Unterschiede zwischen den beiden Gruppen ( $F(1,210)=22,90, p<0,001, \eta^2=.10$ ). Die Kontrollgruppe nahm mehr Hilfe und Unterstützung durch die Lehrer wahr. In zwei der drei Räume berichteten SuS der Kontrollgruppe einen erhöhten CL ( $F(1,135)=3,77, p<0,05, \eta^2=0,03$ ;  $F(1,151)=9,18, p<0,01, \eta^2=0,06$ ;  $F(1,166)=0,05, p=ns$ ).

Die Bereitstellung eines höheren Maßes an Struktur bewirkte bei den SuS eine geringere Inanspruchnahme von Hilfsangeboten, sowie geringeren CL als in der Kontrollgruppe. Struktur könnte die Selbstständigkeit und die Eigenaktivität von SuS fördern.

Baard, P. P., Deci, E. L., Ryan, R. M. (2004). Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings. *Journal of Applied Social Psychology*, 34, 2045-2068.

Falk, J.H., Dierking, L.D. (2000). *Learning from museums: Visitor experience and the making of meaning*. Walnut Creek, CA: Altamira Press

Falk, J.H., Martin W.W., Balling, J.D. (1978). Environmental effects on learning: The outdoor field trip. *Science Education*, 65 (3), 301-309

Gottfried, J.L., (1980). Do children learn on school field trips? *Curator*, 23, 165-174.

Paas, F., Merriënboer, J. van (1994). Measurement of cognitive load in instructional research. In: *Perceptual and Motor Skills*, 79, 419-430

Reeve, J. & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98, 209-218.

Skinner, E., Furrer, C., Marchand, G., Kindermann, T. (2008). Engagement and Disaffection in the Classroom: Part of a Larger Motivational Dynamic? *Journal of Educational Psychology*, 100 (4), 765-781

Wellborn, J., Connell, J., Skinner, E.A., & Pierson, L.H. (1988). Teacher as social context: A measure of teacher provision of involvement, structure and autonomy support. (Tech.Rep. No. 102).

Rochester. NY: University of Rochester.



Udo Käser (Rheinische Friedrich-Wilhelms-Universität Bonn),  
Meta Miriam Bönniger (Rheinische Friedrich-Wilhelms-Universität Bonn)

## **Die Entwicklung strategischer Kompetenzen im Grundschulalter: Wie entwickelt sich der Einsatz von Rechenstrategien für das Kopfrechnen bei Schülerinnen und Schülern der vierten Klasse?**

Eine Kernaufgabe von Schule besteht darin, Schülerinnen und Schülern beizubringen, wie man lernt. Die Vermittlung von metakognitivem Wissen und von metakognitiven Fähigkeiten sollte daher eine zentrale Bedeutung bei der Organisation schulischer Lernprozesse einnehmen. Lernende müssen um Personenmerkmale, Aufgabenmerkmale und Strategiemerkmale des Wissenserwerbs wissen und dazu in der Lage sein, kognitive Prozesse zu kontrollieren und zu regulieren. Denn es sind letztlich derartige deklarative und prozedurale metakognitive Kompetenzen (Hasselhorn, 2006), die es uns erlauben effektiv zu lernen sowie durch Strukturierung und (Re-) Organisation von Informationen ein tieferes Verständnis über Phänomene der Welt zu gewinnen. Bildung kann insofern nur gelingen, wenn Schülerinnen und Schüler in der Schule auch strategische Kompetenzen (Hasselhorn & Gold, 2006) erwerben, die über bloßes Fachwissen hinausgehen.

Solche strategischen Kompetenzen können fachübergreifenden Charakter aufweisen wie Prinzipien des selbstregulierten Lernens, aber auch fachspezifische Charakteristika besitzen. Beispiele hierfür sind etwa die Strategie der Pluralbildung, um die korrekte Rechtschreibung von Endkonsonanten zu überprüfen, oder die Bildung der Quersumme ganzer Zahlen, um zu untersuchen, ob sie ohne Rest durch 3 oder 9 teilbar sind.

Auch im Zuge von Schulevaluationsforschung etwa durch Studien im Kontext von PISA wurde bereits versucht, entsprechende Kompetenzen zu messen. So untersuchten Artelt, Demmrich und Baumert (2001) die Verwendung unterschiedlicher Gedächtnisstrategien. Häufig basieren entsprechende Untersuchungen jedoch auf Selbstauskünften und greifen daher methodisch zwangsläufig zu kurz (Bortz & Döring, 2006). Valide Aussagen über Strategienutzung können letztlich nur gewonnen werden, wenn der Strategieeinsatz direkt gemessen wird. Vor diesem Hintergrund wird der Strategieeinsatz von Grundschulern exemplarisch an Kopfrechenaufgaben untersucht. Kopfrechenaufgaben sind aus mehreren Gründen von besonderem Interesse. Einerseits lassen sich strategische Entscheidungen hinsichtlich der Art und Weise, wie Rechenoperationen ausgeführt werden – zum Beispiel Anwendung des Distributivgesetzes bei Multiplikationsaufgaben oder der Auffülltechnik bei Subtraktionsaufgaben (Padberg & Benz, 2011) –, auf Grundlage des Rechenwegs analysieren, der als Lösungsweg gewählt wurde: Insofern kann Strategieeinsatz beim Kopfrechnen durch die Verbalisierung des Rechenwegs gemessen werden (Funke, 1996). Andererseits stellen Kopfrechenfähigkeiten im Sinne des Spiralprinzips (Bruner, 1980) eine wichtige Basiskompetenz für komplexere mathematische Probleme wie algebraische Operationen dar: Insofern ist der Strategieeinsatz bei Kopfrechenaufgaben langfristig bedeutsam. Des Weiteren ordnen die aktuellen Ergebnisse der TIMSS-Studie (2011) die mathematischen Fähigkeiten deutscher Schülerinnen und Schüler im internationalen Vergleich im oberen Drittel an, konstatieren aber für den Bereich der Arithmetik vergleichsweise schwache Leistungen. Insofern stellt sich die Frage, ob dies durch Mängel beim Kopfrechnen oder durch unzureichende strategische Kompetenzen erklärt werden kann.

Die vorliegende Studie zur Entwicklung der Kopfrechenfähigkeit wurde längsschnittlich über drei Messzeitpunkte an einer Stichprobe von 56 Viertklässlern realisiert. Zu jedem Messzeitpunkt wurde die Kopfrechenfähigkeit der Probanden quantitativ und qualitativ in Einzeltestungen durch 20 Kopfrechenaufgaben in elementaren Rechenoperationen erfasst. Des Weiteren wurden neben Alter, Geschlecht und Schulnote das schulische Selbstkonzept via SESSKO (Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2002) sowie die Lern- und Intelligenzleistung via KLI4-5R (Schröder, 2005) erhoben.

Quantitativ zeigt sich, dass die Leistungen der Schülerinnen und Schüler stark variieren, im Mittel nur mäßig ausfallen und sich im Verlauf des vierten Schuljahrs nur noch geringfügig verbessert. Es liegt ein signifikanter, positiver Zusammenhang zwischen der Leistung zum ersten Zeitpunkt und dem Entwicklungsverlauf der Kopfrechenleistung vor. Weiterhin ergeben sich Zusammenhänge zwischen Teilaufgaben des KLI4-5R (Zahlenumwandeln und Rechenaufgaben) und der Kopfrechenleistung. Qualitativ zeigen sich ebenfalls nur unzureichende Verbesserungen hinsichtlich der effektiven Nutzung mathematischer Strategien. So unterlaufen selbst zum letzten Messzeitpunkt noch drei von vier Schülern systematisch Fehler bei der Anwendung des Distributivgesetzes.

Diskutiert werden die Befunde angesichts der methodisch-didaktischen Anforderungen an die Gestaltung und den Anspruch des Grundschulunterrichts (beispielsweise Gasteiger & Paluka-Graham, 2013). Nimmt im Grundschulunterricht, im Fach Mathematik im Speziellen ebenso wie im Allgemeinen, die Vermittlung strategischer Fähigkeiten ausreichend Raum ein? Und wie kann die Vermittlung strategischer Fähigkeiten in der Grundschule verbessert werden?

Artelt, C., Demmrich, A. & Baumert, J. (2001). Selbstreguliertes Lernen. In Deutsches PISA-Konsortium (Hrsg.), PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich (S. 271-298). Opladen: Leske + Budrich.

Bortz, J. & Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 4. Auflage. Berlin: Springer.

Bruner, J. S. (1980). Der Prozess der Erziehung. 5. Auflage. Berlin: Cornelsen.

Funke, J. (1996). Methoden der Kognitiven Psychologie. In E. Erdfelder, R. Mausfeld, T. Meiser & G. Rudinger (Hrsg.), Handbuch Quantitative Methoden (S. 515-528). Weinheim: Psychologie Verlags Union.

Gasteiger, H. & Paluka-Graham, S. (2013). Strategieverwendung bei Einmaleinsaufgaben – Ergebnisse einer explorativen Interviewstudie. Journal für Mathematik-Didaktik, 34, 1-20.

Hasselhorn, M. (1996). Kategoriales Organisieren bei Kindern. Zur Entwicklung einer Gedächtnisstrategie. Göttingen: Hogrefe.

Hasselhorn, M. & Gold, A. (2009). Pädagogische Psychologie. Erfolgreiches Lernen und Lehren. 2. Auflage. Stuttgart: Kohlhammer.

Padberg, F. & Benz, C. (2011). Didaktik der Arithmetik für Lehrerbildung und Lehrerfortbildung. München: Spektrum Akademischer Verlag.

Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2002). SESSKO. Skalen zur Erfassung des schulischen Selbstkonzepts. Göttingen: Hogrefe.

Schröder, H. (2005). KLI4-5R. Kombiniertes Lern- und Intelligenztest für 4. und 5. Klassen – Revidierte Form. Göttingen: Hogrefe.



## Effekte zwischen metakognitivem Strategiewissen und Lernstrategien?

Die Fähigkeit das Lernen erfolgreich selbst zu regulieren, stellt im Gymnasium eine wichtige Kompetenz dar. Hierbei nimmt das metakognitive Strategiewissen – das Wissen über Lernstrategien und deren Eigenschaften sowie das Wissen über Aufgabenanforderungen – und die Nutzung von Lernstrategien eine zentrale Rolle ein. Zum einen ist die angemessene Anwendung von Lernstrategien ohne metakognitives Strategiewissen kaum denkbar. Zum anderen werden der Erwerb und die Ausdifferenzierung des metakognitiven Strategiewissens von der Lernstrategienutzung beeinflusst. In der theoretischen Literatur finden sich deshalb Hinweise, dass das metakognitive Strategiewissen und die Lernstrategienutzung eng miteinander verbunden sind. In Anlehnung an das Modell des erfolgreichen Strategienutzers (Borkowski et al., 2000) ist davon auszugehen, dass zwischen dem metakognitiven Strategiewissen und den Lernstrategien eine wechselseitige Beziehung besteht. Aus empirischer Sicht ist jedoch ein inkonsistentes Bild bezüglich dem Zusammenhang zwischen metakognitivem Strategiewissen und Lernstrategien vorhanden (z.B. Artelt et al., 2000; Sperling et al., 2004). Bisherige Resultate stammen zudem vorwiegend aus Querschnittstudien oder aus Korrelationsberechnungen. Daraus lassen sich nur beschränkt Hinweise zu den Effektrichtungen ableiten. Im Fokus dieser Studie steht deshalb die Fragestellung nach den Effekten zwischen dem metakognitiven Strategiewissen und den kognitiven und metakognitiven Lernstrategien. Eine repräsentative Stichprobe von Lernenden aus dem 10. und 11. Schuljahr wurde zu Beginn (t1) und am Ende (t2) des Schuljahres 2010/11 in mehreren Gymnasien im Kanton Zürich, Schweiz befragt. In dieser Längsschnittstudie mit zwei Messzeitpunkten nahmen N = 1'272 Gymnasiastinnen und Gymnasiasten zu t1 und N = 1'126 zu t2 teil. Bei 897 Gymnasiastinnen und Gymnasiasten konnten die Informationen aus den beiden Befragungen kombiniert werden. Die befragten Gymnasiastinnen (58.6%) und Gymnasiasten (41.4%) hatten dabei ein Durchschnittsalter von 16.6 Jahren (SD = 1.06). Die Online-Erhebungen wurden im Zeitraum von neun Monaten jeweils während zwei Schulstunden im Regelunterricht unter Aufsicht einer Lehrperson durchgeführt.

Das Testinstrument zur Erfassung des metakognitiven Strategiewissens basiert auf dem Konstruktionsprinzip ähnlicher Tests für andere Bereiche und Schulstufen (z.B. Schlagmüller & Schneider, 2007). Mit dem eingesetzten metakognitiven Strategiewissenstest wurde das Wissen im Kontext des Verfassens von grösseren und komplexen Arbeiten und Projekten erfasst (Maag Merki et al., 2013). In sieben verschiedenen Lernszenarien, die einen vollständigen Selbstregulationsprozess umfassen, mussten die Lernenden die jeweils aufgelisteten Lernstrategien hinsichtlich ihrer Nützlichkeit auf einer sechsstufigen Skalen (1 = nicht nützlich bis 6 = sehr nützlich) zur Lösung der Aufgabenanforderungen einschätzen. Zur Erfassung der selbstberichteten Nutzung von kognitiven (Transformations- und Elaborationsstrategien) und metakognitiven (Evaluations- und Planungsstrategien, Monitoring) Lernstrategien wurde ein standardisierter Fragebogen eingesetzt. Latente cross-lagged Modelle wurden zur Untersuchung der Effekte zwischen dem metakognitiven Strategiewissen und den verschiedenen Lernstrategien modelliert.

Die Ergebnisse weisen auf unterschiedliche Effekte zwischen den Lernstrategien und dem metakognitiven Strategiewissen hin. Einerseits zeigt sich ein unidirektionaler Effekt zwischen der selbstberichteten kognitiven Lernstrategienutzung (Elaborationsstrategien) zu t1 auf die Entwicklung des metakognitiven Strategiewissens zu t2. Andererseits fand sich ein unidirektionaler Effekt des metakognitiven Strategiewissens zu t1 auf die selbstberichtete Nutzung von Transformationsstrategien zu t2. Des Weiteren wurde eine reziproke Beziehung zwischen dem metakognitiven Wissen und den Evaluationsstrategien gefunden. Zusammenfassend liess sich ein erster Hinweis für die im theoretischen Modell von Borkowski und Kollegen (2000) postulierte wechselseitige Beziehung zwischen metakognitivem Strategiewissen und der Lernstrategienutzung finden. Des Weiteren zeigen die Resultate, dass es zwischen den unterschiedlichen Lernstrategien zu differenzieren gilt. Diese Ergebnisse werden vor dem Hintergrund möglicher Implikationen für die Praxis sowie zukünftigen Analysen diskutiert.

Artelt, C., Naumann, J., & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), PISA 2009. Bilanz nach einem Jahrzehnt (S. 73-112). Münster: Waxmann.

Borkowski, J. G., Chan, L. K. S., & Muthukrishna, N. (2000). A process-oriented model of metacognition: Links between motivation and executive functioning. In G. Schraw & J. C. Impara (Hrsg.), Issues in the measurement of metacognition (S. 1-42). Lincoln: Buros Institute of Mental Measurements.

Maag Merki, K., Ramseier, E., & Karlen, Y. (2013). Reliability and validity of a newly developed test to assess learning strategy knowledge. *Journal of Cognitive Education and Psychology*, 12(3), 391-408.

Schlagmüller, M., & Schneider, W. (2007). Würzburger Lesestrategie – Wissenstest für die Klassen 7–12 (WLST 7-12). Göttingen: Hogrefe.

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and Self-Regulated Learning Constructs. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 10(2), 117-139.



## ***Mehrperspektivische empirische Blicke auf die Durchlässigkeit von Berufsausbildungen mit geringerem Anforderungsniveau in dualen Berufsbildungssystemen***

Chair(s): Lars Balzer (Eidgenössisches Hochschulinstitut für Berufsbildung Schweiz)

Diskutant/in: Frank Musekamp (Institut Technik und Bildung (ITB), Universität Bremen)

Durchlässigkeit im Bildungssystem ist ein schillernder Begriff, der mit vielen Erwartungen und Chancen, aber auch Ernüchterungen verbunden ist. Dabei ist schon der Begriff keineswegs eindeutig, selbst wenn man ihn auf die Berufsausbildung fokussiert. So sprechen wir von horizontaler Durchlässigkeit beim Übergang in eine ausbildungsadäquate Berufstätigkeit (Fach- und Niveauadäquanz). Vertikale Durchlässigkeit ist mit den Möglichkeiten des Übertritts in eine anspruchsvollere Ausbildung verbunden. Die Eintrittsdurchlässigkeit beschreibt die Eintrittsschwelle und den Zugang zu einer qualifizierenden Ausbildung, z.B. für bestimmte Personengruppen. Mit interner Durchlässigkeit ist die Durchlässigkeit innerhalb eines Ausbildungssystems gemeint; mangelnde Durchlässigkeit führt dagegen zu Segmentierungen.

In diesem Symposium wird das Thema Durchlässigkeit am Beispiel von Berufsausbildungen mit geringerem Anforderungsniveau in dualen Berufsbildungssystemen aus unterschiedlichen disziplinären und theoretischen Perspektiven empirisch beleuchtet. Dabei präsentiert das Symposium seine Perspektivenvielfalt auf unterschiedlichen Ebenen:

Vier Beiträge aus Deutschland, Österreich und der Schweiz, d.h. Länder, die ein duales Berufsausbildungssystem kennen, erlauben einen internationalen Blick. Sie untersuchen verschiedene Aspekte von Durchlässigkeit mit unterschiedlichen disziplinären Zugängen (berufswissenschaftlich, soziologisch, erziehungswissenschaftlich, psychologisch) sowie unterschiedlichen methodischen Ansätzen (bi- und multivariate Analysen von Voll- und Stichprobenerhebungen (Bildungsstatistiken und Fragebogenerhebungen), qualitative Auswertungen von Fallstudien und Leitfadeninterviews). Die Beiträge eint, Durchlässigkeit empirisch und theoretisch fundiert in der o.g. thematischen Breite zu untersuchen:

- Berufswissenschaftliche Untersuchung 2-jähriger Berufe und deren Wirkung auf die Beschäftigungsstrukturen am Beispiel Kfz-Service-Mechaniker/in in Deutschland.

Untersucht wird, ob mit der Ausbildung eine Anlage auf dauerhafte, vom Lebensalter unabhängige berufliche Tätigkeit geschaffen wurde, ob eine Grundlage für Fortbildung und beruflichen Aufstieg gegeben ist und ob ein hinreichender Bedarf an entsprechenden Qualifikationen vorhanden ist, der zeitlich unbegrenzt und einzelbetriebsunabhängig ist.

- Zweijährige Ausbildung: Zeichen einer weiteren Segmentierung im System der dualen Ausbildung in Deutschland?

Hier interessiert, inwieweit die unterschiedliche Dauer von Ausbildungsberufen eine (weitere) Segmentationslinie im dualen Ausbildungssystem darstellt, entlang derer sich Chancen verschieden verteilen.

- Berufseinmündung von AbsolventInnen der Integrativen Berufsausbildung (IBA) in Österreich - eine Analyse der Beschäftigungsverläufe.

Es wird die Frage beantwortet, wie Berufseinmündung und Beschäftigungsverläufe von IBA-AbsolventInnen verlaufen und inwieweit das Ziel der IBA (Verbesserung der Eingliederung von benachteiligten Personen in das Berufsleben) erreicht wurde.

- Berufsbiografien nach einer zweijährigen beruflichen Grundbildung mit eidgenössischem Berufsattest (EBA) in der Schweiz.

Thema ist, ob EBA-Absolvierende 5 Jahre nach Ausbildungsabschluss einen existenzsichernden Platz im Arbeitsmarkt gefunden und die Chancen nach anspruchsvolleren Ausbildungen ergriffen haben sowie, wie sich ihre Berufsverläufe entwickelt haben sowie wie zufrieden sie mit ihren Berufskarrieren sind. Ein abschliessender Diskussionsbeitrag würdigt kritisch die Bedeutung der empirischen Ergebnisse zu Durchlässigkeit aus internationaler Perspektive.

Mittwoch,

05.03.2014,

11:15 - 13:00,

IG-Farben-Haus

Raum 1.411

Matthias Becker (Berufsbildungsinstitut Arbeit und Technik, Universität Flensburg),  
Georg Spöttl (Institut Technik und Bildung (ITB), Universität Bremen)

## ***Berufswissenschaftliche Untersuchung 2-jähriger Berufe und deren Wirkung auf die Beschäftigungsstrukturen am Beispiel Kfz-Service-Mechaniker/in***

Theoretischer Hintergrund: Der zweijährige Ausbildungsberuf Kraftfahrzeugservice-Mechaniker/ Kraftfahrzeugservice-Mechanikerin wurde in Deutschland 2004 für eine Erprobung verordnet und bis zum 31. Juli 2013 beibehalten. In den Jahren 2010 bis 2012 wurde der Beruf durch die beiden Vortragenden im Bundesgebiet evaluiert. Dabei standen Forschungsfragen zur Ausbildung schwächerer Schulabsolventen und deren späteren Beschäftigungschancen im Zentrum, deren Beantwortung zur Beratung des Sachverständigenbeirats bezüglich der Beurteilung der Ausbildungs- und Arbeitsmarkteffekte genutzt wurde. Der Forschungsstand (vgl. Kelle/Kluge, 1999; Solga, 2005; Musekamp, 2010) findet in inhaltlicher wie methodischer Hinsicht Berücksichtigung, wobei vor allem aufgezeigt wird, wie sich die Art der Untersuchung auf die Ergebnisqualität auswirkt. Die unterschiedlichen Ergebnisse werden charakterisiert, verglichen und bewertet. Die Notwendigkeit einer berufswissenschaftlichen Betrachtung wird begründet. Zum einen wird Bezug genommen zum Spannungsfeld beruflicher Anforderungen in Kfz-Betrieben mit seinen komplexen Aufgaben und einem Ausbildungsprofil im Service- und Reparatursektor mit einem reduzierten Qualifikationsniveau. Zum anderen wird die Eignung und das Verhältnis quantitativer und qualitativer Erhebungsmethoden aufgezeigt.

Fragestellung: Kernfragen der Etablierung und Evaluation von Ausbildungsberufen sind in Deutschland durch die „Kriterien und Verfahren für die Anerkennung und Aufhebung von Ausbildungsberufen“ (BWP, 1974) benannt und bis heute relevant. Die Frage, ob mit der Ausbildung eine „Anlage auf dauerhafte, vom Lebensalter unabhängige berufliche Tätigkeit“ geschaffen, eine „Grundlage für Fortbildung und beruflichen Aufstieg“ gegeben und ein „hinreichender Bedarf an entsprechenden Qualifikationen, der zeitlich unbegrenzt und einzelbetriebsunabhängig ist“ steht im Mittelpunkt. Solche Fragen lassen sich nur schwer durch quantitativ ausgerichtete Befragungsmethoden beantworten, weil damit Wechselbeziehungen zwischen verschiedenen Einflussfaktoren kaum aufzudecken sind. Auf letztere kommt es aber an, weil Berufe als gesellschaftlich ausgehandelte Konstrukte vielfältige Abhängigkeiten von Arbeitsorganisationsformen, Arbeitsmarktentwicklungen, Aufgabenstrukturen, Weiterbildungsmöglichkeiten sowie technologischen, ökonomischen und ökologischen Bedingungen haben. Am Beispiel der deutschlandweiten Evaluation des zweijährigen Ausbildungsberufes Kfz-Service-Mechaniker/-in lässt sich zeigen, dass diese Wechselbeziehungen durch berufswissenschaftlich ausgerichtete Fallstudien aufgedeckt und zur Beantwortung von Gestaltungsfragen nutzbar gemacht werden können.

Methode: Im Beitrag wird die Reichweite und der jeweilige Erkenntnisgewinn mehrerer Befragungsmethoden mit Blick auf die Relevanz der Ausbildung für den Arbeitsmarkt und die Durchlässigkeit in andere Ausbildungsgänge und Berufe diskutiert (vgl. Becker et al., 2012). Aufgezeigt wird, welche Wirkungen die Ergebnisse jeweils auf die Entscheidungsfindung zur Weiterentwicklung des untersuchten Berufes hatten. Die Forschungsfragen wurden durch ein Forschungsdesign, bestehend aus quantitativen Erhebungen zur Erfassung von Trends und zur Erhöhung der Reliabilität und Objektivität und aus qualitativen Erhebungen (berufswissenschaftliche Fallstudien) zur vertiefenden Beantwortung und Aufdeckung von Wechselbeziehungen untersucht. Dadurch ließen sich insbesondere die betrieblichen und schulischen Bedingungen aufdecken, unter denen Kfz-Service-Mechaniker/innen eine den Berufskriterien entsprechende Beschäftigung finden können.

Ergebnisse: Das Fallstudienkonzept wurde durch insgesamt 23 Fallstudien in allen mit der Ausbildung beteiligten Institutionen und Personen umgesetzt. Es ließ sich zeigen, dass durch eine kriteriengestützte Auswahl von Fällen eine hohe Validität und Reliabilität berufswissenschaftlicher Methoden erreicht werden kann. Insbesondere konnten so die Wechselbeziehungen zwischen Ausbildungs- und Beschäftigungsbedingungen aufgedeckt werden. Zentrale Ergebnisse der Untersuchung sind etwa, dass der Abschluss der zweijährigen Ausbildung nicht in ausreichendem Maß für eine Berufstätigkeit qualifiziert und dass die Ausbildung trotz formaler Durchstiegsmöglichkeiten problematische Folgen hinsichtlich Umsetzung der Ausbildung wie auch hinsichtlich der Anerkennung und Verwertung der Qualifikation hatte. So entschieden sich am Ende die Sozialpartner vor allem wegen der unzureichend gegebenen Durchlässigkeit in den Arbeitsmarkt dafür, diesen Ausbildungsgang nicht in eine Regelverordnung zu überführen. Die Qualität der unterschiedlichen Methoden und Erhebungsinstrumente wird an Beispielen zu den Erhebungsergebnissen konkretisiert, die Grundlage für die bildungspolitischen Entscheidungen waren.

BWP (1974): Empfehlung betr. Kriterien und Verfahren für die Anerkennung und Aufhebung von Ausbildungsberufen. Bonn: Bundesausschuss für Berufsbildung. BWP 5/1974.

Becker, M.; Spöttl, G. (2008): Berufswissenschaftliche Forschung. Ein Arbeitsbuch für Studium und Praxis. Frankfurt a. M. u.a.: Peter Lang.

Becker, M.; Spöttl, G.; Karges, T.; Musekamp, F.; Bertram, B. (2012): Kfz-Service-Mechaniker/in auf dem Prüfstand. Chancen und Grenzen zielgruppenspezifischer Berufsausbildung. Reihe Berichte zur beruflichen Bildung. Bielefeld: W. Bertelsmann.

Kelle, U.; Kluge, S. (1999). Vom Einzelfall zum Typus: Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung. Opladen: Leske + Budrich.

Muskamp, F. (2010): Betrieblicher Einsatz von Absolventen zweijähriger Ausbildungsberufe. Eine empirische Untersuchung des Ausbildungsberufes Kfz Service-Mechaniker/in. Bielefeld: W. Bertelsmann Verlag.

Solga, H. (2005): Ohne Abschluss in die Bildungsgesellschaft. Die Erwerbchancen gering qualifizierter Personen aus soziologischer und ökonomischer Perspektive. Opladen: Barbara Budrich.



## **Zweijährige Ausbildung: Zeichen einer weiteren Segmentierung im System der dualen Ausbildung in Deutschland?**

Theoretischer Hintergrund: Die nichtakademische berufliche Ausbildung ist für die Qualifizierung eines großen Teils der SchulabgängerInnen in Deutschland bedeutsam, da weniger als die Hälfte der SchulabgängerInnen die allgemeinbildende Schule mit der Hochschulreife beendet. Die nichtakademische Berufsausbildung ist stark vom dualen System der Berufsausbildung geprägt: Rund 70% der Jugendlichen, die eine nichtakademische Ausbildung durchlaufen, erhalten sie im dualen System, ca. 30% in einer vollzeitschulischen Ausbildung (Schulberufssystem) (Statistisches Bundesamt, 2013). Die 350 Ausbildungsberufe im dualen System lassen sich nach der Dauer in 2-jährige bzw. 3 – 3,5 jährige Ausbildungsberufe unterscheiden (Uhly et al., 2011).

Wenngleich 2-jährige Berufe ein eingeschränkteres und theoriegemindertes Qualifikationsprofil aufweisen, gilt es für alle Ausbildungsberufe im dualen System unabhängig von der Dauer der Ausbildung Kenntnisse und Fähigkeiten einer Vollqualifizierung zu vermitteln. Zweijährige Berufe stehen seit Mitte der 90er Jahre im Fokus der (berufs)bildungspolitischen Diskussion, da sie aus Sicht der Politik angesichts steigender Anforderungen auf dem Arbeitsmarkt und eines schwankenden Ausbildungsangebots vorrangig Jugendlichen mit besonderem Förderbedarf Ausbildungsmöglichkeiten bieten sollen (Granato & Kroll, 2013).

Der Beitrag fokussiert auf die Frage inwieweit die unterschiedliche Dauer der Ausbildungsberufe eine (weitere) Segmentationslinie im System der dualen Ausbildung darstellt, entlang derer sich Chancen unterschiedlich verteilen. Diese Frage wird vor dem Hintergrund segmentationstheoretischer Ansätze (Doeringer & Piore, 1985; Sengenberger, 1987), der Arbeitskräftewarteschlange (Thurow, 1978) und der Verdrängungs- bzw. Diskreditierungsthese (Gerstenberger & Seltz, 1978; Solga 2011) untersucht.

Fragestellung: Der vorliegende Beitrag beleuchtet, welche Jugendlichen in zweijährigen Ausbildungsberufen ausgebildet werden und untersucht dabei die Hypothese, inwieweit die unterschiedliche Dauer der Ausbildungsberufe eine (weitere) Segmentationslinie im System der dualen Ausbildung darstellt, entlang derer sich Chancen unterschiedlich verteilen. Ein besonderes Augenmerk gilt der Frage, inwieweit die Dauer einer Ausbildung zu einer (weiteren) geschlechtsspezifischen Segmentierung des dualen Systems beiträgt. Junge Frauen, im allgemeinbildenden Schulsystem als „Bildungsgewinnerinnen“ angesehen (Hadjar & Becker, 2011), haben beim Zugang in eine duale Ausbildung unter Berücksichtigung ihrer Bildungsvoraussetzungen und anderer zentraler Einflussfaktoren unterproportionale Chancen auf einen betrieblichen Ausbildungsplatz (Beicht & Granato, 2011). Welche Bedeutung hat die Dauer der Ausbildung bei der Zuweisung junger Frauen und Männer auf duale Ausbildungsplätze? Unsere Hypothese ist, dass junge Frauen mit mittleren Abschlüssen überproportional häufig in 2-jährige Ausbildungsberufe „kanalisiert“ werden.

Methode: Empirische Grundlage sind bi- und multivariate Auswertungen aus zwei Vollerhebungen zur dualen Ausbildung in Deutschland: zum einen die BIBB-Erhebung über die neu abgeschlossenen Ausbildungsverträge zum 30.09. (kurz: BIBB-Erhebung zum 30.09.) und zum anderen die Berufsbildungsstatistik der statistischen Ämter des Bundes und der Länder (kurz: Berufsbildungsstatistik). Beide Erhebungen erfassen Daten zu den dualen Ausbildungsberufen nach BBiG/HwO.

Ergebnisse: Die empirischen Analysen weisen darauf hin, dass sich junge Frauen mit einem mittleren Schulabschluss häufiger als die männliche Vergleichsgruppe in einer 2-jährigen Ausbildung wiederfinden, d.h. eine Ausbildung erhalten, die (aus Sicht der Berufsbildungspolitik) vorrangig für Jugendliche mit besonderem Förderbedarf bzw. für schulisch schwächere Jugendliche in Frage kommt (Busemeyer, 2009). Dies bedeutet eine zusätzliche Segmentierung innerhalb des dualen Systems – aus der Gender Perspektive ebenso wie aus institutioneller Perspektive. Genderungleichheit wird im dualen System nicht nur insgesamt, über die Segmentierung in den Ausbildungsberufen – horizontal mit vertikalen Folgen (z.B. Ausbildungsvergütung, Entlohnung; Trappe, 2006; Granato & Schwerin, 2008), sondern auch über die überproportionale Beteiligung schulisch gut vorqualifizierter junger Frauen an 2-jähriger Ausbildung, d.h. am „unteren Rande“ des Ausbildungssystems „transportiert“. Für das institutionelle Gefüge des dualen Systems in Deutschland deutet sich durch den Ausbau zweijähriger Ausbildungsberufe eine Ausdifferenzierung bzw. Ausweitung des Ausbildungssystems nach unten hin an.

Beicht, Ursula; Granato, Mona 2011: Risiken am Übergang Schule – Ausbildung: Verlieren junge Frauen hier ihre Bildungsvorteile? Und verschärfen sich die Nachteile für Jugendliche mit Migrationshintergrund? in M. Icking, Hrsg., Die Zukunft der Beruflichen Bildung. Schriften zu Bildung und Kultur, Band 4, Berlin, Heinrich-Böll-Stiftung, S. 37 – 58.

Busemeyer Marius 2009: Wandel trotz Reformstau. Die Politik der beruflichen Bildung seit 1970. Schriften aus dem Max-Planck-Institut für Gesellschaftsforschung Köln, n° 65.

Doeringer, Peter B.; Piore, Michael J. 1985: Internal Labor Markets and Manpower Analysis, 2nd ed., Armonk/NY.

Gerstenberger; Friedrich; Seltz, Rüdiger 1978: Disproportionalitäten zwischen Ausbildungs- und Beschäftigungssystem», Sozialwissenschaftliche Informationen für Unterricht und Studium 4, S. 160-165.

Granato, Mona; Kroll, Stephan 2013: L'alternance en Allemagne: différenciation de la formation sans différenciation des diplômes? In: Cahiers de la recherche et des savoirs, H-S n 4, S. 109-131 (im Erscheinen).

Granato, Mona; Schwerin, Christine 2008: Potenziale erkennen – Kompetenzen nutzen. Ausbildung junger Frauen in technisch orientierten Berufen, in H. Schwitzer, C. Wilke & M. Kopel, (Hrsg.), aktiv – kompetent – mittendrin. Frauenbilder in der Welt der Arbeit, Hamburg, S. 75-90.

Hadjar, Andreas; Becker, Rolf 2011: Erwartete und unerwartete Folgen der Bildungsexpansion, in R. Becker, Hrsg., Lehrbuch der Bildungssoziologie, Wiesbaden, S. 195-213.

Sengenberger, Werner 1987: Struktur und Funktionsweisen von Arbeitsmärkten: Die Bundesrepublik Deutschland im internationalen Vergleich, Frankfurt a.M./New York.

Solga, Heike 2011: Bildungsarmut und Ausbildungslosigkeit in der Bildungs- und Wissensgesellschaft, in R. Becker, Hrsg., Lehrbuch der Bildungssoziologie, Wiesbaden, pp. 411-448.

Statistisches Bundesamt 2013: Bildung und Kultur. Integrierte Ausbildungsberichterstattung. Anfänger, Teilnehmer und Absolventen im Ausbildungsgeschehen nach Sektoren/Konten und Ländern, Wiesbaden.

Trappe, Heike 2006: Berufliche Segregation im Kontext. Über einige Folgen geschlechtstypischer Berufsentscheidungen in Ost- und Westdeutschland », Kölner Zeitschrift für Soziologie und Sozialpsychologie, S. 51-77.

Uhly, Alexandra; Kroll, Stephan; Krekel, Elisabeth M. 2011: Strukturen und Entwicklungen der zweijährigen Ausbildungsberufe des dualen Systems. Bundesinstitut für Berufsbildung, Wissenschaftliche Diskussionspapiere, n° 125, Bonn

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
IG-Farben-Haus  
Raum 1.411



## **Berufseinmündung von AbsolventInnen der Integrativen Berufsausbildung in Österreich - Eine Analyse der Beschäftigungsverläufe**

Theoretischer Hintergrund: Seit der Einrichtung der Integrativen Berufsausbildung (IBA) im Jahr 2003 in Österreich kann ein kontinuierlicher Anstieg der Zahl der teilnehmenden Jugendlichen verzeichnet werden (von 1.114 Lehrlingen zu Jahresende 2004 auf 5.741 Lehrlinge zu Jahresende 2012) (Dornmayr & Nowak, 2013). Die überwiegende Mehrheit der IBA-Lehrlinge (61%) befindet sich (2012) in Unternehmen, 39% in überbetrieblichen Ausbildungseinrichtungen.

Die Integrative Berufsausbildung verfolgt das Ziel einer Verbesserung der Eingliederung von benachteiligten Personen in das Berufsleben. Dafür sind zwei Varianten vorgesehen:

Verlängerung der Lehrzeit um bis zu einem Jahr, in Ausnahmefällen auch bis zu zwei Jahre

Teilqualifizierung Einschränkung auf bestimmte Teile des Berufsbildes eines Lehrberufs. Dauer dieser („reduzierten“) Ausbildung: 1- 3 Jahre.

Rund 74% der IBA-Lehrlinge absolvieren 2012 die Integrative Berufsausbildung in Form einer Verlängerung der Lehrzeit, rund 26% in Form einer Teilqualifizierung.

Fragestellung: Primäres Ziel der Untersuchung (Dornmayr, 2012) ist die Gewinnung von Erkenntnissen über Berufseinmündung und Beschäftigungsverläufe von AbsolventInnen der Integrativen Berufs-ausbildung (IBA) in Österreich, welche bisher nur für kurzzeitige Nachkarrieren von IBA-AbsolventInnen der überbetrieblichen Ausbildung vorlagen (Bergmann, Lechner, Matt, Riesenfelder, Schelepa, & Willsberger, 2011). Es soll somit eine unmittelbare Antwort auf die Frage erfolgen, inwieweit das Ziel der IBA (Verbesserung der Eingliederung von benachteiligten Personen in das Berufsleben) erreicht wurde. Vergleichend werden auch die Beschäftigungsverläufe jener Personen, welche die IBA vorzeitig beendet haben („Drop-Outs“), untersucht.

Methode: Insgesamt wurden als Grundgesamtheit alle Personen, die seit Einführung der Integrativen Berufsausbildung (IBA) in Österreich (2003) bis zum 13.12.2011 diese Ausbildung beendet oder abgebrochen haben, definiert. Diese Grundgesamtheit besteht aus 11.142 Personen.

Anzumerken bleibt, dass „Abbruch“ („Drop-Out“) in diesem Zusammenhang nicht immer einen Abbruch der Lehrausbildung bedeutet, sondern in manchen Fällen auch einen Übertritt in ein anderes (teilweise „reguläres“) Lehrverhältnis.

Als Datenquellen der Untersuchung dienen zum einen die über die Lehrverträge erfassten Daten der IBA-TeilnehmerInnen sowie zum anderen (zur Dokumentation der Beschäftigungsverläufe) die Sozialversicherungsdaten (inkl. AMS).

Ergebnisse: Die wesentlichen Ergebnisse der Analyse der Beschäftigungsverläufe der AbsolventInnen und „Drop-Outs“ nach Beendigung der IBA lassen sich folgendermaßen zusammenfassen:

- 1.) Die IBA „wirkt“, d.h. sie begünstigt die Arbeitsmarktintegration erheblich: Ein Monat nach Ende der Integrativen Berufsausbildung (IBA) sind 52% der AbsolventInnen in Beschäftigung. Weitere 10% befinden sich in einem neuen (teilweise „regulären“) Lehrverhältnis. 22% der AbsolventInnen sind arbeitslos gemeldet und 8% in einer Qualifizierungsmaßnahme des AMS. Unter den „Drop-Outs“ hingegen beträgt ein Monat nach Ausbildungsende der Anteil der Beschäftigten lediglich 8%, während 42% der „Drop-Outs“ arbeitslos gemeldet sind und weitere 11% sich in einer Qualifizierungsmaßnahme des AMS befinden.
- 2.) Auch vier bzw. fünf Jahre nach Beendigung der IBA liegt der Anteil an Beschäftigten unter den IBA-AbsolventInnen noch immer deutlich über jenem der „Drop-Outs“.
- 3.) In einer relevanten Zahl von Fällen stellt die IBA auch eine Art „Einstieg“ für einen Übertritt in ein anderes (teilweise „reguläres“) Lehrverhältnis dar. Insgesamt haben 38% der „Drop-Outs“ später wieder ein neues Lehrverhältnis aufgenommen. Die Hälfte davon (also 19% und somit fast ein Fünftel aller „Drop-Outs“) hat (irgendwann) nach Abbruch der IBA eine „reguläre“ Lehrausbildung begonnen.
- 3.) Die mittels IBA bewirkten Effekte unterscheiden sich erheblich dahingehend, ob die IBA in einem Betrieb oder in einer überbetrieblichen Ausbildungseinrichtung stattfindet: Ein Monat nach Ende der IBA sind 76% der AbsolventInnen einer betrieblichen Ausbildung in Beschäftigung aber lediglich 20% der AbsolventInnen einer IBA in einer überbetrieblichen Ausbildungseinrichtung. Von erstgenannter Gruppe sind lediglich 9% arbeitslos und 1% in einer Qualifizierungsmaßnahme des AMS, von letztgenannter 38% arbeitslos und weitere 18% in einer Qualifizierungsmaßnahme des AMS. Sogar fünf Jahre nach Ausbildungsende ist der Unterschied noch erheblich: Der Beschäftigtenanteil unter den AbsolventInnen einer betrieblichen IBA beträgt hier 69%, unter den AbsolventInnen einer überbetrieblichen IBA 43%.

Bergmann, Nadja / Lechner, Ferdinand / Matt, Ina / Riesenfelder, Andreas / Schelepa, Susanne / Willsberger, Barbara (2011): Evaluierung der überbetrieblichen Lehrausbildung (ÜBA) in Österreich, Wien

Dornmayr, Helmut (2012): Berufseinmündung von AbsolventInnen der Integrativen Berufs-ausbildung – Eine Analyse der Beschäftigungsverläufe, ibw-Forschungsbericht Nr. 167, Wien (Download: <http://www.ibw.at/de/ibw-studien>)

Dornmayr, Helmut / Nowak, Sabine (2013): Lehrlingsausbildung im Überblick 2013 – Strukturdaten, Trends und Perspektiven, ibw-Forschungsbericht Nr. 176, Wien (Download: <http://www.ibw.at/de/ibw-studien>)

Marlise Kammermann (Eidgenössisches Hochschulinstitut für Berufsbildung),  
Achim Hättich (Interkantonale Hochschule für Heilpädagogik)

## ***Berufsbiografien nach einer zweijährigen beruflichen Grundbildung mit eidgenössischem Berufsattest (EBA) in der Schweiz***

Theoretischer Hintergrund: Berufliche Ausbildungen in dualen Berufsbildungssystemen fördern erfolgreiche Übergänge von der Schule ins Berufsleben und garantieren in der Regel eine gute Arbeitsmarktintegration (Raffe, 2008; Stalder, 2012). Mit der Inkraftsetzung eines neuen Berufsbildungsgesetzes in der Schweiz im Jahr 2004 wurde angestrebt, berufliche Ausbildungsmöglichkeiten im gesamten Begabungsspektrum anzubieten (Schweizerische Eidgenossenschaft, 2002). Zweijährige Grundbildungen, die zu einem eidgenössischen Berufsattest (EBA) führen, sind das niederschwelligste, auf eidgenössischer Ebene reglementierte berufliche Ausbildungsangebot. Sie unterscheiden sich von drei- oder vierjährigen Grundbildungen mit eidgenössischem Fähigkeitszeugnis (EFZ) durch eine stärkere Orientierung an praktischen Arbeitstätigkeiten und durch individuelle Unterstützungsmassnahmen während der Ausbildung (Bundesamt für Berufsbildung und Technologie, 2005 & 2007). Sie richten sich an Jugendliche des unteren schulischen Leistungsspektrums und sollen somit auch jungen Menschen mit Lern- und Leistungsbeeinträchtigungen einen standardisierten, qualifizierenden Berufsabschluss und eine anschliessende Integration in den Arbeitsmarkt ermöglichen (Kammermann, Stalder, & Hättich, 2013). Eine Abstimmung der Ausbildungsinhalte zwischen den zwei- und drei- oder vierjährigen Ausbildungen stellt die Durchlässigkeit in ein höheres Ausbildungsniveau innerhalb eines Berufes sicher (Barmettler, 2008). Die ersten Absolventinnen und Absolventen einer zweijährigen Grundbildung schlossen ihre Ausbildung im Jahr 2007 ab. Es sind dies Detailhandelsassistentinnen und -assistenten sowie Küchen-, Restaurations- und Hotellerieangestellte mit eidgenössischem Berufsattest (EBA). Die vier Berufe sind in den Berufsfeldern des Detailhandels sowie des Gastgewerbes angesiedelt.

Fragestellung: Der vorliegende Beitrag fokussiert die folgenden drei Fragestellungen:

- 1) Wie gestalten sich mittelfristige Arbeitsmarktintegration (horizontale Durchlässigkeit) und Zugang zu weiterführenden Ausbildungen (vertikale Durchlässigkeit) für Absolventinnen und Absolventen einer zweijährigen beruflichen Grundbildung?
- 2) Welche Faktoren bestimmen Stabilität respektive Prekarität der Berufsverläufe?
- 3) Wie zufrieden sind junge Berufsleute mit EBA mit ihrer Ausbildung und ihrer bisherigen Berufskarriere?

Methode: Qualitative und quantitative Daten aus einer Schweizer Längsschnittstudie liefern die Grundlage zur Erörterung der obenerwähnten Fragestellungen. 169 junge Berufsleute mit eidgenössischem Berufsattest im Detailhandel und in der Gastronomie werden seit ihrem Berufsabschluss im Jahr 2007 wissenschaftlich begleitet. Im Rahmen von bisher vier Messzeitpunkten – kurz vor Ausbildungsabschluss sowie ein Jahr, zweieinhalb Jahre und fünf Jahre nach Abschluss der Ausbildung – wurden sie sowohl mittels Fragebogen als auch telefonischen Interviews zu ihrer Ausbildung, zu ihrer Integration in den Arbeitsmarkt sowie zu ihrer Berufslaufbahn befragt.

Die quantitativen Analysen des Datensatzes fokussieren Vergleiche über die verschiedenen Messzeitpunkte, die qualitativen inhaltsanalytischen Auswertungen erlauben differenzierte Einblicke in individuelle Berufsbiografien.

Ergebnisse: Die Ergebnisse der Längsschnittstudie zeigen, dass die zweijährige Grundbildung eine gute Basis sowohl für eine qualifizierte Arbeitstätigkeit im gelernten Beruf als auch für einen Einstieg in eine weiterführende Ausbildung bildet (Kammermann, Balzer & Hättich, im Druck): rund zwei Drittel der Befragten arbeiten fünf Jahre nach ihrer ersten beruflichen Qualifizierung im erlernten Beruf, meist in einem unbefristeten Vollzeit-Arbeitsverhältnis; ungefähr jede/r Dritte hat in der Zwischenzeit zusätzlich zum Berufsattest noch ein Fähigkeitszeugnis (meist) im erlernten Beruf erworben.

Die Berufskarrieren der meisten jungen Berufsleute während den ersten zweieinhalb Jahren nach Ausbildungsabschluss weisen jedoch keine gradlinigen Verläufe auf, sie sind geprägt durch Stellenwechsel und Unterbrüche (Kammermann & Hättich, 2010). Ergebnisse zu Berufsverläufen bis fünf Jahre nach der beruflichen Erstqualifizierung sowie zu Bestimmungsfaktoren stabiler respektive prekärer Verläufe werden gegenwärtig ausgewertet und am Kongress präsentiert werden.

Die befragten Personen selber zeigen sich fünf Jahre nach Ausbildungsabschluss mehrheitlich zufrieden mit ihrer momentanen Erwerbssituation sowie mit ihrer bisherigen Berufslaufbahn. Mehr als zwei Drittel sind der Überzeugung, dass ihre berufliche Situation ihren schulischen und beruflichen Fähigkeiten sowie ihren persönlichen Interessen entspricht.

Barmettler, H. (2008). Zusammenarbeit, Durchlässigkeit und Transparenz – Grundzüge der Schweizerischen Berufsbildungsreform. In: *Berufsbildung in Wissenschaft und Praxis* 37(4), 31-34.

Bundesamt für Berufsbildung und Technologie (2005). *Berufliche Grundbildung mit eidgenössischem Berufsattest. Leitfaden*. Bern: Bundesamt für Berufsbildung und Technologie.

Bundesamt für Berufsbildung und Technologie (2007). *Individuelle Begleitung von Lernenden in der beruflichen Grundbildung*. Bern: Bundesamt für Berufsbildung und Technologie.

Kammermann, M., Balzer, L. & Hättich, A. (im Druck). *Eidgenössisches Berufsattest. Mittelfristige Erwerbssituation der jungen Berufsleute*. *Panorama* 27(6).

Kammermann, M., Stalder, B. E. & Hättich, A. (2013). Two-year apprenticeships – a successful model of training? In A. Fuller & L. Unwin (Hrsg.). *Contemporary Apprenticeship. International Perspectives on an Evolving Model of Learning*. Abingdon: Routledge, 140-159.

Kammermann, M. & Hättich, A. (2010). Mit Berufsattest in den Arbeitsmarkt. Ergebnisse einer schweizerischen Längsschnittstudie über die Berufsverläufe nach einer zweijährigen beruflichen Grundbildung. *Berufsbildung in Wissenschaft und Praxis* 39(5), 11-14.

Raffe, D. (2008). The concept of transition system. *Journal of Education and Work* 21(4), 277-96.

Schweizerische Eidgenossenschaft (2002). *Bundesgesetz über die Berufsbildung vom 13. Dezember 2002*. Bern: Bundeskanzlei.

Stalder, B. E. (2012). School-to-Work Transitions in Apprenticeship-Based VET Systems: The Swiss Approach. In S. Billett, G. Johnson, S. Thomas, C. Sim, S. Hay & J. Ryan (Hrsg.). *Experience of School Transitions: Policies, Practice and Participants*. Berlin: Springer, 123-139.



## **Motivation für das Studieren: Theorie - Messung - Intervention**

Chair(s): Thomas Martens (Deutsches Institut für Internationale Pädagogische Forschung),  
Christiane Metzger (Fachhochschule Kiel), Rolf Schulmeister (Universität Hamburg)

Diskutant/in: Katrin Rakoczy (Deutsches Institut für Internationale Pädagogische Forschung)

In diesem Symposium sollen die Prozesse der Motivation und der Motivationsregulation beim Lernen in der Hochschule aus den Blickwinkeln von Psychologie, Erziehungswissenschaft und Hochschuldidaktik betrachtet werden. Die Analyse der motivationalen Prozesse ermöglicht wichtige Erkenntnisse über die Heterogenität des Lernens aus denen dann die Bedingungen für effektives und nachhaltiges Lernen abgeleitet werden können.

Die Beiträge für dieses Symposium sind passgenau aufeinander abgestimmt und können so das Spektrum von Theorie (Beitrag 1 von Martens & Kuhl), Messung (Beitrag 2 von Schulmeister, Metzger & Martens und Beitrag 4 von Jenert & Brahm) sowie praktische Intervention (Beitrag 3 von Metzger & Haag) aufzeigen. Insbesondere die Beiträge 1, 2 und 3 bauen direkt aufeinander auf: die Messungen in Beitrag 2 berücksichtigen den theoretischen Stand in Beitrag 1 und die Interventionen in Beitrag 3 nehmen direkten Bezug auf die Theorie und die Ergebnisse von Beitrag 1 und 2.

Im ersten Beitrag wird mit dem „Integrierten Lern- und Handlungsmodell“ eine theoretische Fundierung für die Motivation des Lernens an der Hochschule vorgestellt. Insbesondere werden neuere Erkenntnisse aus der Motivationsregulation, der Handlungsregulation und den Neurowissenschaften in einem gemeinsamen Theorierahmen vereinigt. Es ist so möglich, sowohl kognitive und affektive sowie bewusste und unbewusste Prozesse der Lernmotivation zu modellieren. So ist es mit diesem Modell u.a. möglich, so verschiedene Lernphänomene wie den Stereotype Threat, den Korruptionseffekt oder die Lernprokastination gleichermaßen zu erklären.

Im zweiten Beitrag werden verschiedenen Studien zur Messung der Lernmotivation und der Lernzeit von Studierenden berichtet. Dabei nimmt das hier verwendete Messmodell für die Motivation direkten Bezug auf die theoretischen Modellannahmen von Beitrag 1. Es zeigt sich, dass die Heterogenität von Lernprozessen durch die Messung der Komponenten des Integrierten Lern- und Handlungsmodells gut abgebildet werden kann. Die resultierenden Typen der Lernmotivation zeigen korrespondierende Workload und Studienerfolg.

Im dritten Beitrag werden hochschuldidaktische Konsequenzen diskutiert, angewendet und evaluiert, die aus dem Integrierten Lern- und Handlungsmodell (Beitrag 1) und der empirisch beschriebenen Heterogenität von Lernmotivation (Beitrag 2) gezogen werden können. Als zentrale Maßnahmen wurden die Veranstaltungen im einem Bachelor-Studiengang geblockt und die Selbstlernphasen und Präsenzphasen eng miteinander verzahnt. Hierdurch haben sich die Noten der Studierenden verbessert und die Durchfallquote ist gesunken.

Im vierten Beitrag wurde mit einer Mixed-Methods-Studie untersucht, welche kritischen Ereignisse sich auf die Motivationsentwicklung im ersten Studienjahr auswirken können. Der quantitativ belegte Abfall der intrinsischen Motivation korrespondiert mit qualitativ berichteten Widersprüchen zwischen Absichten und Erwartungen der Lehrenden und der kompetitiven Situation zwischen den Studierenden.

Mittwoch,

05.03.2014,

11:15 - 13:00,

IG-Farben-Haus

Raum 251

## Zur theoretischen Fundierung der motivationalen Regulation: das Integrierte Lern- und Handlungsmodell

Das „Integrierte Lern- und Handlungsmodell“ ist eine theoretische Fundierung für die Motivation des individuellen Lernens an der Hochschule. Hierfür werden neuere Erkenntnisse aus der Motivationsregulation, der Handlungsregulation und den Neurowissenschaften in einem gemeinsamen Theorierahmen vereinigt. Bereits das Integrierte Handlungsmodell von Rost und Martens (1998) verbindet die motivational-emotionale und die kognitive Dynamik der Handlungsregulation miteinander und kann damit auch die Prozesse des Lernens erklären, wenn Lernen als spezielle Handlungsform aufgefasst wird (Dulisch, 1986). In der erweiterten Konzeption von Martens (2012) wird insbesondere gezeigt, wie die Phasen und Konstrukte des Modells mit den neurowissenschaftlich fundierten Erkenntnissen von Kuhl (2000, 2001) korrespondieren.

Im Integrierten Lern- und Handlungsmodell werden drei Prozessphasen einer vollständigen Lernhandlung unterschieden, die für tiefgreifende Lernprozesse notwendig sind:

- die Motivierungsphase,
- die Intentionsphase und
- die Volitionsphase.

In der Motivierungsphase wird eine Lernmotivation ausgebildet, d.h. es entsteht ein Bedürfnis, eine lernbezogene Soll-Ist-Diskrepanz zu reduzieren.

In der Intentionsphase wird eine Lernintention gebildet, die die Lernmotivation erfüllen kann.

In der Volitionsphase wird schließlich eine Lernintention in eine tatsächliche Lernhandlung umgesetzt.

Die Phasen können durch die folgenden Konstrukte charakterisiert werden:

Motivierungsphase: Soll-Ist-Diskrepanz (Bedrohungswahrnehmung), Sensitives Coping, Verantwortungsübernahme

Intentionsphase: Handlungssuche, Handlungs-Ergebnis-Erwartung, Kompetenzerwartung

Volitionsphase: Persistente Zielverfolgung, Selbstkongruente Zielverfolgung, Emotions- und Motivationsregulation, Planen & Problemlösen, Implementation

So ist es mit diesem Modell möglich, so verschiedene Lernphänomene wie den Stereotype Threat, den Korruptionseffekt oder die Lernprokrastination zu erklären. Es kann etwa angenommen werden, dass die Korruptionseffekt dann besonders groß ist, wenn vorher keine starke Lernmotivation entwickelt wurde. In diesem Fall wird die durchgeführte Lernhandlung der Belohnung zugeschrieben. Diese nachträgliche Assoziation einer Lernhandlung zu einer externalen Motivation führt dann zur „Korruption“: die Lernhandlung wird bei einer Wiederholung instrumentell eingesetzt um Motive zu erfüllen, die mit der eigentlichen Lernhandlung nichts zu tun haben.

Im Falle einer empirischen Überprüfung des Integrierten Lern- und Handlungsmodells muss der Prozesscharakter des Modells in besonderer Weise berücksichtigt werden. Bei einer Verwendung von Querschnittsdaten müssen etwa verschiedene Prozesszustände durch das Messmodell abgebildet werden können. Hierfür können Mischverteilungsmodellen verwendet werden, die in der Lage sind, alle multivariaten Interaktionen höherer Ordnung simultan zu prüfen (Rost, 2004).

Es werden zwei Beispiele für die empirische Anwendung des Integrierten Lern- und Handlungsmodells in den Studiengängen Erziehungswissenschaften und Wirtschaftswissenschaften gezeigt, die jeweils in der Identifikation von 5 Typen der Lernmotivation resultieren (vgl. Metzger, Schulmeister & Martens, 2012).

Dulisch, Frank (1986). Lernen als Form menschlichen Handelns. Eine handlungstheoretisch orientierte Analyse von Lernprozessen unter besonderer Berücksichtigung des Selbststeuerungsaspektes. Bergisch-Gladbach: Thomas Hobein.

Kuhl, Julius (2000). The volitional basis of Personality Systems Interaction Theory: applications in learning and treatment contexts. International Journal of Educational Research, 33, S. 665-703.

Kuhl, Julius (2001). Motivation und Persönlichkeit - Interaktionen psychischer Systeme. Göttingen, Bern, Toronto, Seattle: Hogrefe.

Martens, Thomas & Rost, Jürgen (1998). Der Zusammenhang von wahrgenommener Bedrohung durch Umweltgefahren und der Ausbildung von Handlungsintentionen. Zeitschrift für Experimentelle Psychologie, 45(4), S. 345 - 364.

Metzger, Christiane, Schulmeister, Rolf & Martens, Thomas (2012). Motivation und Lehrorganisation als Elemente von Lernkultur. Zeitschrift für Hochschulentwicklung, Jg.7 / Nr.3

Rost, Jürgen (2004). Testtheorie - Testkonstruktion. Bern: Huber.





## **Die Workload im Bachelor und andere Prädiktoren für Studienerfolg**

Studien, die die Workload der Bachelor per Zeitbudget-Methode erforscht haben (wie das ZEITLast-Projekt; Schulmeister & Metzger, 2011), führen zu der überraschenden Einsicht, dass Studierende mit einem geringen zeitlichen Einsatz keine/einige/alle Prüfungen bestanden, während Studierende mit einem enorm hohen zeitlichen Einsatz etliche Prüfungen nicht bestanden. Die minutiöse Erhebung und Analyse der Workload über fünf Monate eines Semesters in 27 Bachelor-Studiengängen aus unterschiedlichen Fächern offenbart nicht nur das Versagen des Selbststudiums, sondern öffnet den Blick zugleich für die enorme interindividuelle Diversität im Lernverhalten. Die große interindividuelle Varianz selbst des Präsenzstudiums, aber besonders des Selbststudiums führt zur Einsicht, dass es weder die bzw. den „Normalstudierenden“ gibt noch einen gleichförmigen Studienverlauf (Schulmeister, Metzger & Martens, 2012). Die Zeit, die Studierende im Studium aufwenden, korreliert nicht mit den Prüfungsnoten. Differenziert man jedoch die Studierenden nach ihrer Motivation auf der Grundlage des Integrierten Lern- und Handlungsmodells (ILHM) von Martens (2012), so werden Typen von Studierenden erkennbar, die sich doch nach Workload und Studienerfolg unterscheiden (Metzger, Schulmeister & Martens, 2012): In einer großen Stichprobe im Studiengang BSc BWL wurden Studierende mit einer Erhebung auf der Basis des ILHM getestet, die auch an der ZEITLast-Studie teilgenommen hatten, so dass man die motivationalen Profile von fünf Studierendengruppen mit den Zeitbudgetwerten vergleichen konnte.

Um diese Erkenntnisse abzusichern, wurden in einer neuen Metastudie 300 empirische Studien zu Workload, Employment, Prokrastination, demographischen Variablen und Studienerfolg gesichtet (Schulmeister, im Druck). Die meisten Studien betrachten die Höhe der Workload nicht als Prädiktor für den Studienerfolg, und ebenfalls erweisen sich die üblichen demographischen Variablen (Intelligenz, familiärer Hintergrund, Vorbildung, gender) nicht als entscheidend für den Studienerfolg. Die Effekte der Workload und der demographischen Variablen verlieren an Stärke im Vergleich zu Variablen der Motivation und des Lernverhaltens wie Gewissenhaftigkeit, Aufmerksamkeit, Konzentration und persistente Zielverfolgung als wirksame Determinanten oder Prädiktoren für Studienerfolg. Auf diese Weise wird das Integrierte Lern- und Handlungsmodell mit der Anfügung der Volitionsphase eindrucksvoll bestätigt.

Martens, Th.: Was ist aus dem Integrierten Handlungsmodell geworden? In Kempf & Langeheine, Item-Response-Modelle in der sozialwissenschaftlichen Forschung (S. 210 -229). Berlin: Regener 2012.

Metzger, Ch., Schulmeister, R. & Martens, Th.: Motivation und Lehrorganisation als Elemente von Lernkultur. ZfHE 7/3, 2012.

Schulmeister, R. & Metzger, Ch.: Die Workload im Bachelor. Zeitbudget und Studierverhalten. Waxmann: Münster 2011.

Schulmeister, R., Metzger, Ch. & Martens, Th.: Heterogenität und Studienerfolg. Lehrmethoden für Lerner mit unterschiedlichem Lernverhalten. Paderborner Universitätsreden 123, 2011.

Schulmeister, R.: Auf der Suche nach Determinanten des Studienerfolgs. In: Brockmann & Pilniok, Studieneingangsphase in der Rechtswissenschaft. Nomos 2013: (im Erscheinen)

**„Ist das freiwillig, oder muss ich dahin?!“.**

**Geblockte Module als Reaktion auf eine heterogene Lernmotivation**

Studien zeigen, dass viele Studierende zu Prokrastination neigen – einige Untersuchungen ergeben, dass weit über die Hälfte der Studierenden davon betroffen ist (z.B. Thielsch & Bechler 2012, Güntert & Schleider 2011, Schraw et al. 2007, Ackerman & Gross 2005). Zugleich scheinen Variablen wie Gewissenhaftigkeit, Fleiß, Entschlossenheit und Durchhaltevermögen (conscientiousness, diligence, determination, effort) einen großen Einfluss auf den Studienerfolg zu haben (z.B. Kaufman et al., 2008; Robbins et al., 2006; Brint & Cantwell, 2008); allerdings wird zunehmend beobachtet, dass viele Studierende nicht in ausreichendem Maße über diese Eigenschaften, die unter den „Bologna-Bedingungen“ verschärft wichtig werden, verfügen. Der Vortrag schließt an die Beiträge von Martens sowie von Schulmeister, Metzger & Martens in diesem Symposium an und thematisiert didaktisch-methodische Konsequenzen, die aufbauen auf einer eigenen Untersuchung zur Lernmotivation von Studierenden basierend auf dem Integrierten Lern- und Handlungsmodell von Martens (2012) (Metzger, Schulmeister, Martens 2012). Im Studiengang BSc IT Security an der Fachhochschule St. Pölten wurde im Wintersemester 2011/12 versuchsweise die Lehrorganisation im ersten Fachsemester verändert: Die Module bzw. Teilmodule wurden nicht mehr alle parallel zueinander unterrichtet, sondern jedes Modul wurde exklusiv in geblockter Form über einige Wochen abgehalten. Die Selbststudienphasen wurden eng mit den Präsenzphasen verzahnt. Die jeweilige(n) Prüfung(en) fanden innerhalb des für das Modul vorgesehenen Zeitraums statt.

Die Interventionen wurden durch verschiedene Maßnahmen begleitend evaluiert: Die subjektiven Eindrücke wurden mittels einer Befragung und Fokusgruppengesprächen erhoben. Auf der Basis von Zeitbudget-Erhebungen wurde die Workload der Studierenden analysiert. Zudem liegen die Prüfungsergebnisse vor. Die Resultate zeigen, dass sich die Veränderung der Lehrorganisation positiv auf die Lehre und das Lernen auswirkt: Die Noten haben sich verbessert, die Zahl der Studierenden, die durch Prüfungen durchfallen, ist gesunken. Die Zufriedenheit der Studierenden und Lehrenden ist so groß, dass diese Form der Lehrorganisation im gesamten Bachelor- und auch im Masterstudiengang übernommen wurde.

Aus einer organisationalen Perspektive der Lehrentwicklung zeigt sich, dass die Reorganisation viel zur Weiterentwicklung der Module beigetragen hat: Die Umstrukturierung erforderte es, Lehrinhalte und Methoden zu überdenken und an die neue Art der Lehrorganisation anzupassen. Der gemeinsame Beschluss für die Reorganisation gab den Lehrenden Anlass, unterstützt durch hochschuldidaktische Workshops die Präsenzsitzungen, Selbststudienaufgaben und Prüfungsformate neu zu konzipieren.

Im Vortrag werden die motivationspsychologischen Wirkungsfaktoren im Modell der geblockten Module erläutert sowie die Ergebnisse der Erhebungen und die Erfahrungen bis zum WS 2013/14 mit der Reorganisation der Lehre präsentiert.

Ackerman, D. S & B. L. Gross (2005): My Instructor Made Me Do It: Task Characteristics of Procrastination. In: Journal of Marketing Education; 27, 1, S. 5-13.

Brint, S. & Cantwell, A. M. (2008): Undergraduate Time Use and Academic Outcomes: Results from UCUES 2006. Research & Occasional Paper Series: CSHE.14.08. University of California, Berkeley Oct. 2008, 22 S.

Güntert, M. & Schleider, K. (2011). Studienbezogene Lern- und Arbeitsstörungen. Hamburg: Kovac.

Kaufman, J. C. (2008): The role of personality and motivation in predicting early college academic success in non-traditional students at a Hispanic-serving institution. Learning and Individual Differences 18, S. 492-496.

Martens, T. (2012). Was ist aus dem Integrierten Handlungsmodell geworden? In W. Kempf & R. Langeheine (Hrsg.): Item-Response-Modelle in der sozialwissenschaftlichen Forschung. Berlin: Regener, S. 210-229.

Metzger, Ch., R. Schulmeister & T. Martens (2012): Motivation und Lehrorganisation als Elemente von Lernkultur. In: Euler, Dieter & Taiga Brahm (Hrsg.): Zeitschrift für Hochschulentwicklung Jg. 7, Nr. 3, S. 36-50.

Robbins, S. B. et al. (2006): Unraveling the Differential Effects of Motivational and Skills, Social, and Self-Management Measures From Traditional Predictors of College Outcomes. Journal of Educational Psychology, Vol. 98, No. 3, S. 598-616.

Schraw, G., T. Wadkins & L. Olafson (2007): Doing the Things We Do: A Grounded Theory of Academic Procrastination. In: Journal of Educational Psychology, Vol. 99; 1, S. 12-25.

Thielsch, M. T. & O. Bechler (2012): Probleme von Psychologiestudierenden in der Klausurvorbereitung. In Krämer, M., S. Dutke & J. Barenberg (Hrsg.): Psychologiedidaktik und Evaluation IX. Aachen: Shaker, S. 323-330.

## **Entwicklung der Motivation von Studierenden der Wirtschaftswissenschaften im ersten Studienjahr – eine Mixed-Methods-Studie**

**Forschungsgegenstand und Zielsetzungen:** Die Mixed-Method-Studie untersucht die Entwicklung der Studienmotivation Wirtschaftsstudenten im Verlauf des ersten Studienjahres. Im Rahmen einer quantitativen Längsschnittuntersuchung mit vier Messzeitpunkten wird dabei die Motivationsentwicklung über die gesamte Studierendenpopulation hinweg erfasst. Im Mittelpunkt steht die Frage, wie sich die intrinsische und extrinsische Studienmotivation sowie die studienbezogene Selbstwirksamkeit über das erste Studienjahr verändern. Zudem wird untersucht, inwiefern verschiedene studienbezogene Variablen (z. B. Angst, Qualität der Lehrenden) sich auf die Motivationsentwicklung auswirken. Um vertiefte Erkenntnisse über die Bedeutsamkeit verschiedener Gestaltungsmerkmale des ersten Studienjahres auf die Motivationsentwicklung der Studierenden zu gewinnen, wurde parallel zur quantitativen Untersuchung eine qualitative Längsschnittstudie durchgeführt mit dem Ziel (1) verschiedene Verläufe der Motivationsentwicklung und (b) kritische Ereignisse im Eingangsjahr zu identifizieren.

**Theoretischer Rahmen:** Die Studie befasst sich mit den Konstrukten der extrinsischen und intrinsischen Studienmotivation (Deci & Ryan, 1996) sowie der studienbezogenen Selbstwirksamkeit (Bandura, 1997). Eine Reihe empirischer Untersuchungen bestätigt einen positiven Zusammenhang zwischen verschiedenen motivationalen Konzepten und der Studienleistung (auch Busato, Prins, Elshouta, & Hamaker, 2000; ein Überblick findet sich bei Pintrich, 1999); einige Studien können aber auch keine derartigen Effekte nachweisen (z.B. Bruinsma, 2004). Wenig untersucht ist bisher die Verbindung zwischen motivationalen Konstrukten und anderen Einflussfaktoren auf die Studienleistung (z.B. Schiefele, Streblov, Ermgassen, & Moschner, 2003) sowie die Entwicklung von Studienmotivation im Zeitverlauf (vgl. Fazey & Fazey, 1998; Jacobs & Newstead, 2000). An dieser Stelle zeigt sich eine Forschungslücke betreffend (a) Längsschnittuntersuchungen zu Verläufen der Studienmotivation und (b) der Verknüpfung kontextueller Gestaltungsmerkmale der Studienumgebung mit der Motivationsentwicklung.

**Untersuchungsdesign:** Die Längsschnittstudie wurde an der Universität St. Gallen durchgeführt. Die quantitative Erhebung umfasste insgesamt vier Befragungen im Verlauf des ersten Studienjahres, wobei die erste Erhebung noch vor Studienbeginn lag. Die Verteilungen von Alter, Geschlecht und Nationalität entsprechen zu allen Messzeitpunkten annähernd den Werten der Grundgesamtheit und lassen den Schluss der Repräsentativität des Samples für die Studienanfänger des Jahrgangs 2011 an der Universität St. Gallen zu. Die Rücklaufquote beträgt mit 820 verwertbaren Fragebögen zum ersten Messzeitpunkt rund 63%, wobei infolge der Panel-Mortalität nach dem Messzeitpunkt drei eine bereinigte Stichprobe von 285 Studierenden ausgewertet werden konnte. Die eingesetzten Skalen erfüllen die psychometrischen Anforderungen (Cronbach's Alpha.702 - .787) und die angenommene Faktorstruktur des Instruments konnte bestätigt werden.

Für die qualitative Untersuchung wurden 14 Studierende im Zuge eines theoretischen Samplings kriterienbasiert ausgewählt. Insgesamt wurden mit jedem Studierenden fünf Interviews durchgeführt, wobei das erste Interview entlang der adressierten Konstrukte strukturiert war, die folgenden das erste Gespräch narrativ fortsetzten.

**Ergebnisse: Quantitative Studie.** Die deskriptive Analyse der untersuchten motivationalen Konstrukte zeigt für das erste Studiensemester eine deutliche Negativentwicklung. Alle Konstrukte, insbesondere jedoch die intrinsische Motivation sowie die Selbstwirksamkeit nehmen von t1 zu t2 signifikant ab, wobei die Selbstwirksamkeit im zweiten Semester (d.h. nach dem Bestehen der ersten Prüfungen) wieder etwas steigt. Anzumerken ist, dass intrinsische und extrinsische Motivation sowie Selbstwirksamkeit auf einem sehr hohen Niveau starten. Korrelationen zwischen der Motivationsentwicklung und anderen studienrelevanten Konstrukten sind theoretisch konsistent und bestätigen die Konstruktvalidität des Instruments (z.B. negativer Zusammenhang zwischen Angst und Selbstwirksamkeit).

**Qualitative Studie.** Die Interviewstudie zeigt, dass sich die Studienmotivation in den untersuchten 14 Fällen nicht kontinuierlich abwärts bewegt. Vielmehr berichten die Studierenden von konkreten Ereignissen im Studienverlauf, welche die Motivation stark beeinflussen. Neben Prüfungssituationen gehören dazu vor allem als unklar bzw. widersprüchlich wahrgenommene Absichten und Erwartungen seitens der Universität bzw. der Lehrenden sowie kompetitive Situationen zwischen den Studierenden (z.B. sozialer Vergleich und Wettbewerb bei der Prüfungsvorbereitung).

Die Verknüpfung von qualitativer und quantitativer Studie bietet Ansatzpunkte zur weiteren Gestaltung des Studiums, beispielsweise erscheint es notwendig, das Prüfungswesen so zu verändern, dass ein Motivationseinbruch in der Studieneingangsphase verhindert wird.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and Instruction*, 14(6), 549-568.

Busato, V. V., Prins, F. J., Elshouta, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, 29(6), 1057-1068.

Deci, E. L., & Ryan, R. M. (1996). Need Satisfaction and the Self-Regulation of Learning. *Learning and Individual Differences*, 8(3), 165-183.

Fazey, D., & Fazey, J. (1998). Perspectives on motivation: The implications of effective learning in higher education. In S. Brown, S. Armstrong & G. Thomson (Eds.), *Motivating students* (pp. 59-72). London: Kogan Page.

Jacobs, P. A., & Newstead, S. E. (2000). The nature and development of student motivation. *British Journal of Educational Psychology*, 70(2), 243-254.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459-470. doi: doi:10.1016/S0883-0355(99)00015-4

Schiefele, U., Streblov, L., Ermgassen, U., & Moschner, B. (2003). Lernmotivation und Lernstrategien als Bedingungen der Studienleistung. Ergebnisse einer Längsschnittstudie. *Zeitschrift für Pädagogische Psychologie*, 17(3/4), 185-198.

## Referenzgruppeneffekte der mittleren Intelligenz einer Klasse auf Fähigkeitsbeurteilungen durch Lehrkräfte

Chair(s): Rachel Wollschläger (Universität Trier)

Theoretischer Hintergrund: Individuelle Fähigkeitsbeurteilungen durch Lehrkräfte sind häufig von Kontextfaktoren beeinflusst. Beispiele für beeinflussende Kontextfaktoren sind die mittlere Intelligenz oder Leistungsfähigkeit einer Schulklasse, welche sowohl die Lehrerbeurteilung (Marsh, 1984, 1987; Marsh, Trautwein, Lüdtke & Köller, 2008) als auch die Fähigkeitswahrnehmung der einzelnen Schüler/innen beeinflusst. Der Einfluss der mittleren Klassenleistung auf das individuelle akademische Selbstkonzept ist als Big-Fish-Little-Pond Effekt (BFLPE; Marsh, 1987, 2005) bekannt. Der BFLPE bildet die theoretische Grundlage dieser Arbeit, da neben dem akademischen Selbstkonzept auch Schullaufbahneempfehlungen (z.B. Trautwein & Baeriswyl, 2007) durch den Referenzrahmen beeinflusst werden.

Fragestellung: Wir untersuchen für Grundschul Kinder Referenzgruppeneffekte auf die Lehrereinschätzung von kognitiven Fähigkeiten, auf die Empfehlungswahrscheinlichkeit eine Begabtenfördermaßnahme (Studie 1), sowie auf die Schullaufbahneempfehlung (Studie 2). Wir erwarten einen positiven Einfluss der individuellen Intelligenz sowie – nach Kontrolle dieses Zusammenhangs – einen negativen Einfluss der mittleren Intelligenz der Schulklasse auf die Lehrereinschätzung der kognitiven Fähigkeiten sowie die Empfehlungswahrscheinlichkeit bzw. die Schullaufbahneempfehlung.

Methode: Stichprobe Wir untersuchten unsere Fragestellung in zwei Stichproben. Neben den Kindern wurden auch die Eltern und Lehrer befragt.

In Studie 1 wurden 1201 Schüler/innen (51,9 % Schülerinnen; M = 8,5 Jahre, SD = 8,3 Monate) der 2. bis 4. Klasse von 20 Schulen in Rheinland-Pfalz getestet. In der noch laufenden Studie 2 wurden bisher 398 Viertklässler/innen (48,6 % Schülerinnen; M = 10,9 Jahre, SD = 7,6 Monate) aus 26 Schulklassen erhoben. Die Daten stammen aus drei Bundesländern. Unser Ziel sind 40 Schulklassen à mindestens 15 Schüler/innen, um den Anforderungen der Mehrebenenanalyse (Maas & Hox, 2005) zu entsprechen. Die Datenerhebung ist Ende Januar 2014 abgeschlossen.

Material Die kognitive Fähigkeit der Kinder wurde mittels des Tests für (hoch) intelligente Kinder – T(H)INK (Baudson & Preckel, 2013) erfasst. Der speziell für Grundschul Kinder und Gruppentestungen entwickelte T(H)INK erfasst fluide Intelligenz (Gf) in der verbalen, numerischen und figuralen Domäne. Die psychometrischen Eigenschaften des Verfahrens sind gut (Baudson & Preckel, 2013). Als Lehrerfragebögen wurden bestehende Verfahren eingesetzt und gegebenenfalls modifiziert. Die Lehrereinschätzung der kognitiven Fähigkeit der Kinder wurde anhand von 6 Items (7stufige Likert-Skala; Cronbach's  $\alpha = .97$ ) erfasst. Die Eltern wurden hinsichtlich sozioökonomischer und demographischer Variablen und der letzten Zeugnisnoten der Kinder befragt.

Ergebnisse: Studie 1 Aufgrund der genesteten Datenstruktur wurden die Daten mittels Mehrebenenanalyse ausgewertet. Die Ergebnisse zeigen einen negativen Einfluss der mittleren Intelligenz der Klasse auf die Einschätzungen der individuellen kognitiven Fähigkeiten ( $\beta = -.93$ ,  $SE = .12$ ,  $p < .001$ ). Für die Empfehlung für ein Begabtenförderprojekt scheinen individuelle Variablen wie Intelligenz ( $\beta = .38$ ,  $SE = .03$ ), Sozioökonomischer Status der Eltern ( $\beta = .13$ ,  $SE = .03$ ), Geschlecht zu Gunsten der Jungen ( $\beta = .15$ ,  $SE = .05$ ) und Schulnoten ( $\beta = -.26$ ,  $SE = .04$ , alle  $p < .001$ ) ausschlaggebender zu sein.

Studie 2 Die Daten werden analog zu Studie 1 ausgewertet. Erste Befunde replizieren den negativen Einfluss der mittleren Intelligenz der Klasse auf die Einschätzungen der individuellen kognitiven Fähigkeiten ( $\beta = -.53$ ,  $SE = .06$ ,  $p < .001$ ). Für die Schullaufbahneempfehlung lässt sich dieser Einfluss bisher nicht zeigen.

Diskussion und Implikationen: Während bei der Einschätzung der kognitiven Fähigkeiten der Schüler/Innen Referenzgruppeneffekte auftreten, scheinen für die Nominierung für ein Begabtenförderprojekt dagegen individuelle Schülermerkmale ausschlaggebender zu sein. Mögliche Mediatoren wurden bislang noch nicht untersucht, sind aber Bestandteil folgender Analysen. Auch wenn die diagnostischen Fähigkeiten der Lehrkräfte besser sind als ihr Ruf, sind die Urteile dennoch durch implizite Annahmen, Stereotype und Kontextvariablen beeinflusst (Südkamp, Kaiser & Möller, 2012). Da Entscheidungen in der Schulzeit weitreichenden Einfluss auf Bildungs- und Berufskarrieren haben (Baumert, Trautwein & Artelt, 2003), sollten sie möglichst frei von Verzerrungen sein. Je mehr über verzerrende Einflüsse bekannt ist, desto genauer können Beurteilungen dahingehen überprüft und gegebenenfalls korrigiert werden.

Baudson, T. G. & Preckel, F. (2013). Development and validation of the German Test for (Highly) Intelligent Kids – T(H)INK. *European Journal of Psychological Assessment*, 29, 171–181.

Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In *Deutsches PISA-Konsortium (Hrsg.), Pisa 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (p. 261–331). Opladen: Leske + Budrich.

Maas, C. J. M. & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1, 86–92.

Marsh, H. W. (1984). Relations Among Dimensions of Self-Attribution, Dimensions of Self-Concept, and Academic Achievements. *Journal of Educational Psychology*, 76, 1291–1308.

Marsh, H. W. (1987). The Big-Fish-Little-Pond Effect on Academic Self-Concept. *Journal of Educational Psychology*, 79, 280–295.

Marsh, H. W. (2005). Big-Fish-Little-Pond Effect on Academic Self-Concept. *Zeitschrift für Pädagogische Psychologie*, 19, 119–127.

Rauer, W. & Schuck, K.-D. (2003). Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen – FEES 3–4. Göttingen, Germany: Hogrefe.

Rauer, W. & Schuck, K.-D. (2004). Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen – FEES 1–2. Göttingen, Germany: Hogrefe.

Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of Teachers' Judgments of Students' Academic Achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104, 743–762.

Trautwein, U. & Baeriswyl, F. (2007). Wenn leistungsstarke Klassenkameraden ein Nachteil sind: Referenzgruppeneffekte bei Übertrittsentscheidungen. *Zeitschrift für pädagogische Psychologie*, 21, 119–133.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G092



Thomas Lösch (Eberhard Karls Universität Tübingen), Ingo Zettler (Eberhard Karls Universität Tübingen), Augustin Kelava (Eberhard Karls Universität Tübingen), Ulrich Trautwein (Eberhard Karls Universität Tübingen), Marcus Hasselhorn (DIPF)

## **Welche Kinder werden von ihren Lehrkräften für Enrichment-Programme für begabte und hochbegabte Kinder nominiert? Eine Multilevelanalyse zu den Hector-Kinderakademien**

Theoretischer Hintergrund: Zur Identifizierung besonders begabter und hochbegabter Schüler werden häufig standardisierte Leistungstests (z.B. IQ  $\geq$  130) oder Nominierungen durch Lehrkräfte eingesetzt. Ein Vergleich beider Varianten zeigt, dass Lehrkräfte in nennenswerter Zahl auch Schüler mit „nur“ durchschnittlichem IQ als begabt oder hochbegabt identifizieren (Barber & Torney-Purta, 2008; Neber, 2004). Insofern beeinflussen (auch) andere Variablen – wie z.B. der sozioökonomische Status der Eltern – die Entscheidungen von Lehrkräften. Neben kognitiven Fähigkeiten nennen Lehrkräfte selbst insbesondere Kreativität und motivationale Faktoren als wichtige Kriterien (z.B. Endepohls-Ulpe & Ruf, 2005). Bisher wurde jedoch kaum untersucht, welche Prädiktoren mit der Nominierung durch Lehrkräfte in Zusammenhang stehen, wenn kognitive, motivationale und sozioökonomische Variablen gemeinsam betrachtet werden.

Fragestellung: In der vorliegenden Studie untersuchen wir, welche Merkmale die Nominierung von Kindern als besonders begabt und hochbegabt durch Lehrkräfte vorhersagen. Neben Facetten der Intelligenz werden dabei das forschende Interesse, der Einfluss des sozioökonomischen Status der Eltern und das Geschlecht der Schüler untersucht. Insgesamt liefert unsere Studie Erkenntnisse auf zwei Ebenen: Erstens können relevante Prädiktoren der Nominierung erfasst und zweitens kann auf implizite Modelle der Lehrkräfte über Begabung und Hochbegabung geschlossen werden (vgl. z.B. Sternberg & Zhang, 1995).

Methode: Im Rahmen der wissenschaftlichen Begleitung der Hector Kinderakademien wurde von insgesamt N = 2619 Schülern erhoben, ob sie von ihren Lehrkräften als besonders begabt oder hochbegabt eingestuft wurden. Hector-Kinderakademien bieten außerunterrichtliche Förderangebote für besonders begabte und hochbegabte Grundschul Kinder an. Lehrkräfte nominieren Schüler dabei für die Teilnahme an diesen Angeboten. Die Erhebung fand in 187 dritten Klassen aus 111 Schulen in Baden-Württemberg statt. Folgende Skalen wurden verwendet:

- Der Berliner Test zur Erfassung fluider und kristalliner Intelligenz (BEFKI), adaptiert für die 3. und 4. Klasse (Schipolowski, Schroeders & Wilhelm, in Vorbereitung).
- Die Dimension Investigative des RIASEC Modells (vgl. z.B. Holland, 1973), erfasst mit einer erprobten Eigenentwicklung als Selbstberichtsbogen für Grundschul Kinder, sowie Epistemic Curiosity (vgl. z.B. Litman & Mussel, 2003) als Variablen des forschenden Interesses.
- Die von Eltern beurteilten HEXACO-Skalen Offenheit für neue Erfahrungen und Gewissenhaftigkeit (Allgaier, Zettler, Göllner, Hilbig, & Trautwein, 2013).
- Die Vorjahresnoten in Mathematik und Deutsch, das Geschlecht der Grundschul Kinder und den sozioökonomischen Status der Eltern.

Statistische Auswertungen wurden mit R (R Core Development Team, 2013) durchgeführt. Zur Bestimmung relevanter Prädiktoren wurde eine Reihe generalisierter Mehrebenenmodelle mit Logit-Linkfunktion mit dem Package lme4 (Bates, Maechler, & Bolker, 2013) geschätzt. Die Modelle wurden unter schrittweiser Hinzunahme weiterer Prädiktoren in folgender Reihenfolge aufgebaut:

1. Intelligenz
2. Noten
3. Forscherisches Interesse und Persönlichkeit
4. Demografie und sozioökonomischer Hintergrund

Der Modellfit wurde durch Vergleiche zwischen Fit-Indices bewertet.

Ergebnisse: Das Modell mit der besten Anpassung enthielt alle oben genannten Prädiktoren. Sowohl fluide als auch kristalline Intelligenz waren signifikante Prädiktoren für die Nominierung. Allerdings wurde der Einfluss von fluider Intelligenz durch weitere Prädiktoren substantiell abgeschwächt. Als Grund ist die Multikollinearität zwischen kristalliner und fluider Intelligenz sowie den Noten zu sehen. Gute Prädiktoren für die Nominierung in allen geschätzten Modellen waren die Vorjahresnoten in Mathematik und Deutsch. Auch Offenheit für neue Erfahrungen, der sozioökonomische Status der Eltern und das Geschlecht waren signifikante Prädiktoren. Mädchen hatten eine geringere Wahrscheinlichkeit nominiert zu werden. Entgegen der Erwartungen war Gewissenhaftigkeit kein signifikanter Prädiktor. Auch hier ist die Ursache in der Multikollinearität zwischen Gewissenhaftigkeit und den Schulnoten zu sehen. Das forschende Interesse leistete keinen eigenständigen Beitrag. Insgesamt zeigte sich, dass Lehrer zwar die schulische Leistungsfähigkeit bei Nominierungen berücksichtigen, dass aber auch Variablen wie der sozioökonomische Status und das Geschlecht die Nominierungswahrscheinlichkeit beeinflussen.

Allgaier, K., Zettler, I., Göllner, R., Hilbig, Benjamin, E., & Trautwein, U. (2013, September). Entwicklung und erste Validierung des HEXACO-Persönlichkeitsinventars für die Grundschule (HEXACO-Elementary-School-Inventory, HEXACO-ESI). Präsentation auf der 12. Arbeitstagung der DGPS Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Greifswald.

Barber, C. & Torney-Purta, J. (2008). The relation of high-achieving adolescents' social perceptions and motivation to teachers' nominations for advanced programs. *Journal of Advanced Academics*, 19, 412-443.

Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using Eigen and Eigen. Retrieved from <http://cran.r-project.org/package=lme4>

Endepohls-Ulpe, M., & Ruf, H. (2005). Primary school teachers' criteria for the identification of gifted pupils. *High Ability Studies*, 16, 219-228.

Litman, J. A., & Mussel, P. (2013). Validity of the Interest-and Deprivation-Type Epistemic Curiosity Model in Germany. *Journal of Individual Differences*, 34(2), 59-68.

Neber, H. (2004). Teacher identification of students for gifted programs: nominations to a summer school for highly-gifted students. *Psychology Science*, 46, 348-362.

R Core Development Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>

Sternberg, R. J. & Zhang, L.-f. (1995). What do we mean by giftedness? A pentagonal implicit theory. *Gifted Child Quarterly*, 39, 88-94.

Schipolowski, S., Schroeders, U., & Wilhelm, O. (in Vorbereitung). Berliner Test zur Erfassung fluider und kristalliner Intelligenz für Grundschul Kinder.





## **Wer kommt weiter? Vorhersage der Qualifikation zur Internationalen Biologie- und Chemie-Olympiade auf Grundlage des Leistungsmotivations-Modells von Eccles**

Die internationale Biologie- (IBO) und Chemieolympiade (IChO) sind weltweit jährlich stattfindende Schülerwettbewerbe. Wettbewerbsziel ist es, interessierte wie talentierte Sekundarstufenschülerinnen und -schüler zu fördern. Anhand eines vierstufigen Qualifikationsprozesses werden die vier deutschen IBO- und IChO-Teammitglieder ausgewählt. Um in diesem Auswahlverfahren erfolgreich zu sein, sind neben einem hohen fachspezifischen Wissen motivationale Faktoren bedeutend. Auf Grundlage des Leistungsmotivations-Modells von Eccles et al. (1983) sollte daher die Qualifikation der Teilnehmerinnen und Teilnehmer für die vierte Auswahlrunde der IBO und IChO vorhergesagt werden. Studien mit ehemaligen Olympiateilnehmern ergaben, dass deren Familien sich durch ein hohes Bildungsniveau und einen geringen Leistungsdruck auszeichnen, viele Teilnehmer Erstgeborene und Mädchen unterrepräsentiert sind (z.B. Campbell, 1996). Zudem weisen Olympioniken gute mathematische und naturwissenschaftliche Noten auf, nehmen bereits während der Schulzeit erfolgreich an Wettbewerben teil (u.a. Campbell, 1996, Urhahne, Ho, Parchmann & Nick, 2012) und erleben ihre Schulklasse oftmals als langweilig und wettbewerbsarm (Verna & Feng, 2002). In einer ersten Studie zur Qualifikationsvorhersage konnten Urhahne et al. (2012) zeigen, dass zwischen männlichen und weiblichen IChO-Teilnehmern sowie zwischen qualifizierten und nicht qualifizierten Teilnehmern signifikante Unterschiede bestehen. Qualifizierte gaben einen höheren intrinsischen Wert in der Familie an, schnitten im Intelligenztest besser ab, hatten bessere Schulnoten, nahmen öfter an Wettbewerben teil, hatten eine höhere Erfolgserwartung, schrieben der Teilnahme einen höheren Nutzen zu und erlebten weniger Furcht vor Misserfolg. Die vorherige Teilnahme erwies sich als signifikanter Prädiktor der Qualifikation (Urhahne et al., 2012). Aufbauend auf der Studie von Urhahne et al. (2012), wurde hier der Fokus durch den Einbezug einer weiteren Olympiade erweitert. Anhand des Erwartungs-Wert-Modells (Eccles et al., 1983) wurde getestet, welche Modellvariablen die Qualifikation zur vierten Auswahlrunde der beiden Olympiaden vorhersagen. Es wurde geprüft, ob Qualifikantinnen und Qualifikanten auf Grundlage relevanter Modellvariablen richtig klassifiziert werden können und ob Unterschiede in der Ausprägung der Modellvariablen zwischen hochbegabten und hochleistenden Olympioniken bestehen. Im Fokus standen 87 Teilnehmerinnen und Teilnehmer der dritten Runde des Selektionsprozesses. IBO-Teilnehmende (n = 43, 46.5% weiblich) waren im Schnitt 17.16 (SD = 1.02) und IChO-Teilnehmende (n = 44, 25% weiblich) 17.08 (SD = 1.06) Jahre alt. Für die vierte Runde qualifizierten sich 12 IBO- und 13 IChO-Teilnehmer.

Zur Qualifikation mussten fachspezifische Prüfungen erfolgreich absolviert werden. Die Variablen des Eccles-Modells wurden anhand von sechs Blöcken erfasst: elterlicher Einfluss; Personenmerkmale; vorherige Leistung; Selbstschemata, Ziele, Motivation und Emotion; Erwartungen und Werte sowie Testleistung. Zur Messung der Modellvariablen füllten die Teilnehmer einen neunseitigen Fragebogen aus und nahmen an einer Intelligenztestung teil. Die Teilnahme erfolgte freiwillig und wurde vom Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik der Universität Kiel organisiert. Bei IBO und IChO korrelierten Auszeichnungen, Langeweile im Fachunterricht, Erfolgserwartung und relative Kosten signifikant positiv mit der Qualifikation. In der IBO-Stichprobe korrelierte zudem die vorherige Teilnahme und in der IChO-Stichprobe korrelierten kurz- und langfristige Ziele, fachliches Interesse, Zielerreichungswert und die mathematisch-naturwissenschaftliche Note signifikant positiv mit dem Wettbewerbserfolg. Die bei beiden Gruppen signifikant korrelierenden Modellvariablen gingen als Prädiktoren in binär-logistische Regressionen zur Vorhersage der Qualifikation ein. IBO-Teilnehmende mit mehr erlebter Langeweile im Fachunterricht und IChO-Teilnehmende mit einer höheren Erfolgserwartung qualifizierten sich eher für die vierte Runde. Das IBO-Regressionsmodell sagt die Qualifikation in 83.7% und das der IChO in 90.9% der Fälle richtig vorher. Univariate Varianzanalysen ergaben signifikante Unterschiede zwischen hochbegabten und hochleistenden Olympioniken. Bei der IBO erhielten Hochbegabte mehr Auszeichnungen für vorherige Leistungen und hatten eine höhere Testleistung. Bei der IChO ergaben sich Unterschiede in den Schulnoten, Hoffnung auf Erfolg und den relativen Kosten. Hochbegabte hatten eine bessere mathematisch-naturwissenschaftliche Note. Bei Hochleistenden fielen Hoffnung auf Erfolg und relative Kosten höher aus, sie waren motivierter.

Campbell, J. R. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *International Journal of Educational Research*, 25, 497–522.

Eccles (Parsons), J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motivation* (S. 75–146). San Francisco: Freeman.

Urhahne, D., Ho, L. H., Parchmann, I. & Nick, S. (2012). Attempting to predict success in the qualifying round of the International Chemistry Olympiad. *High Ability Studies*, 23, 167–182.

doi:10.1080/13598139.2012.738324

Verna, M. A. & Feng, A. X. (2002). American Chemistry Olympians achieve the highest level of equity. *Journal of Research in Education*, 12, 101–105.

## **Entwicklung und erste Validierung des Beliefs about Giftedness Questionnaires (BGQ)**

Lehrkräfte nehmen eine wichtige Rolle ein, wenn es um das Erkennen und die akademische Entwicklung von hochbegabten Kindern geht. Dementsprechend viele Studien befassen sich mit den Einstellungen von Lehrkräften zu Hochbegabung. Dabei stehen jedoch meist Einstellungen zum Vorhandensein bestimmter Eigenschaften bei hochbegabten Personen im Vordergrund (z.B. Moon & Brighton, 2008), jedoch nicht die generelle Einstellung zu Aspekten des Konstrukts Hochbegabung. Den wahrscheinlich bekanntesten Ansatz zur Erfassung der generellen Einstellung zur Hochbegabung, der über mehrere Studien verfolgt wurde, stellt die Pentagonal Implicit Theory von Sternberg und Zhang (1995; Zhang & Hui, 2002; Zhang & Sternberg, 1998) dar. Dabei werden verschiedene Einstellungen zu Hochbegabung über die Einschätzung von 60 Personenbeschreibungen erfasst. Es liegt jedoch noch kein systematischer Ansatz vor, der die Einstellung von Lehrkräften zu Hochbegabung mittels eines ökonomischeren Fragebogens erfragt. Der hier vorgestellte Beliefs about Giftedness Questionnaire (BGQ) soll diese Forschungslücke schließen.

In diesem Beitrag werden zunächst die Fragebogenentwicklung sowie die Überprüfung der faktoriellen Validität des BGQ vorgestellt. Anschließend werden Zusammenhänge zwischen Faktoren des Fragebogens mit weiteren Konstrukten aufgeführt, die Hinweise auf die konvergente und diskriminante Validität des Fragebogens geben. Dazu wurden Konstrukte aus den vier Bereichen (i) Hochbegabung, (ii) Lehrerkompetenz (nach dem Modell der professionellen Lehrerkompetenz; Baumert & Kunter, 2011), (iii) Unterrichtsqualität und (iv) Persönlichkeit herangezogen.

Methoden: Zunächst wurden bestehende Hochbegabungsmodelle analysiert und miteinander verglichen, was zu zehn möglichen Faktoren der Einstellung zu Hochbegabung führte. Zu diesen Faktoren wurden dann 86 Items generiert, die als Pilotversion 126 Lehramtsstudierenden vorgelegt wurden. Unter Berücksichtigung der Ergebnisse dieser Pilotstudie wurden im Anschluss einige Items umformuliert sowie der Fragebogen auf 50 Items reduziert. Der überarbeitete Fragebogen wurde zusammen mit Konstrukten der vier oben genannten Bereiche von 603 Lehramtsstudierenden, 175 Grundschullehrkräften und 560 Kursleiter/inne/n eines außerunterrichtlichen Hochbegabtenförderprogramms für Grundschulkindern (N = 135 Kursleiter/innen, die ebenfalls Grundschullehrer/innen sind, und N = 425 Kursleiter/innen, die keine Grundschullehrer/innen sind) bearbeitet.

Ergebnisse und kurze Diskussion: Nach einer Analyse der Items und Faktoren konnte der Fragebogen auf 19 Items gekürzt werden, die sich sechs Faktoren zuordnen lassen: 1. Globalität einer Hochbegabung, 2. Bedeutung der Intelligenz, 3. Veränderbarkeit einer Hochbegabung, 4. Einfluss von Training, 5. Definitionsunterschiede zwischen hochbegabten Kindern und Erwachsenen, 6. Homogenität innerhalb der Gruppe hochbegabter Kinder. Die – über explorative und konfirmatorische Faktorenanalysen (u.a. explorative Strukturgleichungsmodelle) gewonnene – Faktorenstruktur zeigt in allen Teilstichproben (Studierende, Lehrkräften und Kursleiter/innen) einen guten bis sehr guten Fit (z.B. CFI = .91 – .99). Die Korrelationen zwischen den sechs Faktoren des BGQ liegen im niedrigen bis mittleren Bereich ( $r = -.24 - .34$ ). Zur Ableitung von Hinweisen zur konvergenten und diskriminanten Validität wurden die Zusammenhänge zwischen den Faktoren des Fragebogens und den Konstrukten der vier oben genannten Bereiche überprüft. Es zeigen sich beispielsweise signifikante Korrelationen zur Impliziten Intelligenz-Skala von Dweck (1999; Bereich (i) Hochbegabung), zur Leistungsorientierung der Lehrkraft (PISA 03; vgl. Ramm et al., 2006; Bereich (ii) Lehrerkompetenz), zum selbsteingeschätzten Teacher-Support der Lehrkräfte (unveröffentlichte Skala der TRAIN-Studie; Bereich (iii) Unterrichtsqualität) und zur Offenheit für Erfahrung (Skala des HEXACO-60; vgl. Moshagen, Hilbig & Zettler, in Druck; Bereich (iv) Persönlichkeit). Im Vortrag werden die Faktoren des Fragebogens sowie die Validitätsbefunde detailliert vorgestellt und diskutiert.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften (S. 29-53). Münster: Waxmann.

Dweck, C. S. (1999). Self-theories: Their role in motivation, personality, and development. Philadelphia, PA: Psychology Press.

Moon, T. R. & Brighton, C. M. (2008). Primary Teachers' Conception of Giftedness. *Journal for the Education of the Gifted*, 31, 447-480.

Moshagen, M., Hilbig, B. E. & Zettler, I. (in Druck). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars. *Diagnostica*.

Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D. et al. (2006): PISA 2003. Dokumentation der Erhebungsinstrumente. Münster: Waxmann.

Sternberg, R. J. & Zhang, L. (1995). What Do We Mean by Giftedness? A Pentagonal Implicit Theory. *Gifted Child Quarterly*, 39, 88-94.

Zhang, L. & Hui, S. K. (2002). From Pentagon to Triangle: A Cross-Cultural Investigation of an Implicit Theory of Giftedness. *Roeper Review*, 25, 78-92.

Zhang, L. & Sternberg, R. J. (1998). The pentagonal implicit theory of giftedness revisited: A cross-validation in Hong Kong. *Roeper Review*, 21, 149-163.

## ***Lassen sich Unterschiede in der Unterrichtsqualität und in den Lesekompetenzen und der Lesefreude von Schülerinnen und Schülern durch Unterschiede in der Lehrerkooperation erklären?***

Chair(s): Brigitte Steinert (DIPF)

Ziel des Beitrags ist die systematische Analyse der Zusammenhänge zwischen der Lehrerkooperation im Fachkollegium Deutsch und den Grunddimensionen der Unterrichtsqualität im Fach Deutsch sowie deren Beziehung zu den Lesekompetenzen und der Lesemotivation der Schülerinnen und Schüler. Als Schulqualitätsmerkmal, das durch Lehrkräfte gestaltbar ist, ist die unterrichts- und fachbezogene Lehrerkooperation nicht nur ein Indikator für schulische Arbeitsbedingungen, sondern auf Schulebene auch ein Prädiktor für Unterrichtsqualität und – über diese vermittelt – für die Lernergebnisse der Schülerinnen und Schüler. Allerdings liegen zu den Zusammenhängen zwischen der Lehrerkooperation und der Unterrichtsqualität bisher nur wenige und uneinheitliche Befunde vor (Hochweber et al., 2012; Klieme et al., 2010). Deshalb wird diesem Beitrag gefragt, ob Unterschiede in der Unterrichtsqualität durch Unterschiede in der Lehrerkooperation erklärt werden können und, ob diese Zusammenhänge globaler oder differentiell Art sind. Ferner wurde geprüft, welche Zusammenhänge zwischen der Unterrichtsqualität und den Lernergebnissen der Schülerinnen und Schüler bestehen und inwiefern diese über die Lehrerkooperation vermittelt sind.

Theoretischer Hintergrund für die Entwicklung eines kriterienorientierten Messinstruments der unterrichtsbezogenen Lehrerkooperation (Steinert et al., 2008) sind Konzepte zu Lehrerkooperation und Lehrerkompetenzen aus der Organisationsforschung und aus der Lehrerprofessionsforschung, auf deren Basis sich formelle und inhaltliche Kooperationsanforderungen formulieren lassen. Dieses Instrument wurde anhand mehrerer Stichproben und mithilfe von Mehrebenenfaktorenanalysen validiert. Die Unterrichtsqualität wurde in Anlehnung an Klieme et al. (2009) und Baumert et al. (2010) über drei auf das Lesen bzw. das Fach Deutsch bezogene Dimensionen Kognitive Stimulierung, Strukturiertheit, Unterstützung erfasst und angenommen, dass sie jeweils eine spezifische Bedeutung für die Schülerkompetenzen oder die Schülermotivation haben.

Grundlage der Analysen waren die Daten der um die Neuntklässler und das Fachkollegium Deutsch erweiterten Stichprobe von PISA 2009 für Deutschland mit 9246 Schülerinnen und Schülern aus 384 Klassen und 198 Schulen. Folgende Merkmale gingen in die Analysen ein: a) Prädiktoren auf Schulebene: Lehrerkooperation, Schularten, b) Prädiktoren auf Klassen-/Unterrichtsebene: Kognitive Aktivierung, Strukturiertheit, Unterstützung sowie als Kompositionsmerkmale die soziale und kognitive Klassenzusammensetzung, der Mädchen- und Migrantenteil, c) Prädiktoren auf Schülerebene: Sozialstatus, Kognitive Grundfähigkeiten, Geschlecht, Migrationshintergrund sowie Lesekompetenz und Lesefreude als Kriterien. In Mehrebenenanalysen mit der erweiterten PISA-2009-Stichprobe zeigten sich wie in früheren Untersuchungen keine globalen Zusammenhänge zwischen Lehrerkooperation und Unterrichtsqualitätsmerkmalen im Fach Deutsch, wenn die Schularten sowie individuelle und kompositionelle Hintergrundmerkmale kontrolliert wurden. Allerdings konnte ein durch die Schulart moderierter Zusammenhang zwischen Lehrerkooperation und Unterrichtsqualität festgestellt werden.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180.

Hochweber, J., Steinert, B., & Klieme, E. (2012). Lehrerkooperation, Unterrichtsqualität und Lernergebnisse im Fach Englisch. *Unterrichtswissenschaft*, 40 (4), 351-370.

Klieme, E., Steinert, B., & Hochweber, J. (2010). Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse. In W. Bos, E. Klieme, & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 231-255). Münster: Waxmann.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137-160). Münster: Waxmann.

Steinert, B., Hartig, J., & Klieme, E. (2008). Institutionelle Bedingungen sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 411-450). Weinheim: Beltz.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G102

**Beeinflussen Hausaufgaben im (Deutsch-)Unterricht die erreichte Leseleistung?**

Die Vertiefung des Unterrichtstoffes durch außerschulische Hausaufgabenzeit ist in der pädagogischen Praxis und theoretischen Lehr-Lern-Modellen seit jeher gängig. Gewisse Reformansätze forderten in den letzten Jahren, die Hausaufgabenlast der Schülerinnen und Schüler deutlich zu reduzieren, da die Hausaufgaben (HA) in Kritik geraten waren, die Schere zwischen bildungsfernen und bildungsnahen Schichten deutlich zu befördern (Stichworte: reiche Eltern bezahlen kompetenten Nachhilfeunterricht; nicht-deutschsprachige Elternteile könnten bei den HA nicht mithelfen etc.). Die Belastung der Eltern-Kind-Beziehung durch die HA wird zusätzlich als Konfliktpotenzial gesehen. Die Forschungen zur HA-Nützlichkeit sind kontrovers geblieben; auch empirische Resultate bezüglich der Frage, ob HA leistungssteigernd wirken, sind divergent, dabei werden methodische Probleme teilweise als Hintergrund der Nicht-Vergleichbarkeit der Ergebnisse gesehen. Große Längsschnitt-Analysen zu den kausalen Effekten von HA-Zeit, resp. HA-Vergabe durch die Lehrpersonen, auf die Leistungsentwicklung sind noch selten und liegen nicht für alle Kernfächer vor.

In Deutschland sieht die Forschungslage grob skizziert so aus, dass die meisten älteren Studien keine oder nur geringen Einfluss der HA-Zeit auf die Schulleistung (SL) fanden; einzig Roßbach (1995) konnte in den Achtzigerjahren in einer kleineren Studie (N = 900) an der Grundschule (2. & 4.Klassen) in NRW positive Effekte von HA auf den Lernerfolg nachweisen. Die angloamerikanische Forschung präsentierte hingegen positivere Resultate: Auf dem Hintergrund von insgesamt weniger HA-Zeit gegenüber von Deutschland fand sich in Klassen mit HA ein leistungssteigernder Einfluss der HA gegenüber von Klassen, in denen keine HA gegeben wurden; dies traf jedoch vor allem in höheren Klassen zu, wie die Metaanalyse von Cooper Ende der Achtziger zeigen konnte (vgl. Wild, 2004).

Bei neueren Studien finden sich erste differenzierte Ergebnisse bezüglich der Hausaufgabeneffekte: Beispielsweise für Französisch (Klasse 8) zeigt ein neuerer Längsschnitt in der Schweiz betreffend der HA-Zeit negative Korrelationen von langer HA-Zeit mit der Leistung, resp. der -entwicklung, jedoch positive Zusammenhänge mit der von den Schülern gezeigten und berichteten Sorgfalt & Engagement beim Hausaufgabenerledigen (Schnyder, Niggli, Cathomas, Trautwein & Lüdtko, 2006). Beim deutsch-schweizerischen Projekt zu Unterrichtsqualität (Mathematik in Klassen 8 und 9) finden Lipowsky und Kollegen (2004), dass sich die Häufigkeit von HA-Vergabe positiv auf die Leistungsentwicklung auswirkt, währendem die HA-Zeit, respektive die HA-Länge, einen eher negativen Effekt auf die Leistung hat. Auch zeigt der didaktisch-prozessorientierte HA-Umgang der Lehrpersonen im Unterricht leistungssteigernden Effekt. Ebenfalls für Mathematik finden Dettmers und Kollegen (2010) in der deutschen PISA-2003-Erweiterung einen positiven längsschnittlichen Effekt von qualitativ hochwertigen Hausaufgaben mit hohen kognitiven Anforderungen (2 Messzeitpunkte: Klassen 9 und 10). Für Latein identifizieren Haag und Mischo (2002) in der 7. Klassenstufe in Abhängigkeit der Frequenz des HA-Erledigens den "Saisonarbeiter" als erfolglosen Lerner im Unterschied zu den konstanten erfolgreichen Lerner/innen. Für das Kernfach Deutsch stehen ähnlich umfangreiche und differenzierte Analysen auf Klassen- und Individualebene noch aus.

Der empirische Teil dieses Beitrag im Rahmen der Forschungsdaten des Nationalen Bildungspanels (NEPS) stützt sich auf Befragungsdaten der Deutschlehrer/innen und Eltern, Angaben aus den Schüler-Fragebogen zu ihrem Hausaufgabenverhalten, sowie die Schüler-Kompetenzmaße der Klassen 5 und 7. Die Schülerleistungen im Lesen wurden erfasst mittels der NEPS-Lesekompetenztests (vgl. Gehrer, Zimmermann, Artelt, & Weinert, 2012; Gehrer & Artelt, 2013). Untersucht werden in den Klassen 5 und 7 mittels Mehrebenenanalysen mögliche Zusammenhänge von HA-Dauer der Schüler/innen, HA-Vergabeverhalten der Lehrpersonen und der Funktion der HA in ihrem Unterricht (Benotung, Kontrolle, Planung, Fehleranalyse; sowie Extraaufgaben für leistungsstarke Schüler/innen), sowie ggf. das unterstützende Verhalten des Elternhauses.

Dettmers, S., Trautwein, U., Lüdtko, O., Kunter, M., & Baumert, J. (2010). Homework works if homework quality is high: Using multilevel modeling to predict the development of achievement in mathematics. *Journal of Educational Psychology*, 102, 467-482.

Gehrer, K., & Artelt, C. (2013). Literalität und Bildungslaufbahn: Das Bildungspanel NEPS. In Bertschi-Kaufmann, Rosebrock (Hg.) 2013 – Literalität erfassen (pp. 168–187).

Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2012). The assessment of reading competence (including sample items for grade 5 and 9). Scientific Use File 2012, Version 1.0.0. Bamberg: University of Bamberg, National Educational Panel Study.

Haag, L. & Mischo, C. (2002). Saisonarbeiter in der Schule - einem Phänomen auf der Spur. *Zeitschrift für Pädagogische Psychologie*, 16, 109-115.

Lipowsky, F., Rakoczy, K., Klieme, E., Reusser, K., & Pauli, C. (2004). Hausaufgabenpraxis im Mathematikunterricht - ein Thema für die Unterrichtsqualitätsforschung? In J. Doll & M. Prenzel (Eds.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (pp. 250–266). Münster: Waxmann.

Roßbach, Hans-Günther (1995): Hausaufgaben in der Grundschule. Ergebnisse einer empirischen Untersuchung. In: *Die Deutsche Schule* 87 (1), S. 103–112. Online verfügbar unter <http://www.digizeitschriften.de/dms/gcs-wrapper>, zuletzt geprüft am 04.01.2013.

Schnyder, I., Niggli, A., Cathomas, R., Trautwein, U. & Lüdtko, O. (2006): Wer lange lernt, lernt noch lange nicht viel mehr: Korrelate der Hausaufgabenzeit im Fach Französisch und Effekte auf die Leistungsentwicklung. In: *Psychologie in Erziehung und Unterricht* 53, S. 107–121.

Wild, E. (2004): Häusliches Lernen. Forschungsdesiderate und Forschungsperspektiven. In: D. Lenzen, J. Baumert, R. Watermann & U. Trautwein (Hg.): *PISA und die Konsequenzen für die erziehungswissenschaftliche Forschung*. Wiesbaden: VS, Verlag für Sozialwiss (Zeitschrift für Erziehungswissenschaft: Beiheft, 3), S. 37–64.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G102





## ***Eine neue Perspektive auf Lesegewohnheiten und deren Messung: Der Self-Report Habit Index for Reading***

Die Ergebnisse der PISA-Studie (Programme for International Student Assessment) aus dem Jahr 2000 zeigten, dass etwa 42% der 15-jährigen deutschen Schülerinnen und Schüler nicht zum Vergnügen lesen. Dabei hat Lesemotivation die zentrale Funktion, Leseaktivitäten zu initiieren und damit Lesegewohnheiten zu bilden, die schließlich die Lesekompetenz fördern (Artelt, Naumann, & Schneider, 2010). Lesegewohnheiten werden also als vermittelnde Variable zwischen Lesemotivation und Lesekompetenz angesehen (Schiefele, Schaffner, Möller, & Wigfield, 2012).

Gewohnheit wurde in der vergangenen Forschung meist mit Verhaltensfrequenz gleichgesetzt. Elaboriertere Techniken zur Erhebung von Lesegewohnheiten, die nicht lediglich über die Messung durch nur ein bis drei Items geschieht, sind in vielen Fällen aufwändig, wie es beispielsweise bei Tagebuchstudien oder Leselögbüchern der Fall ist (Wigfield & Guthrie, 1997). In moderneren Ansätzen jenseits der Domäne Lesen wird die Konzeptualisierung der Gewohnheit durch den Aspekt der Automtizität ergänzt (Verplanken, Myrbakk, & Rudi, 2005). In der aktuellen Leseforschung mangelt es an dieser weitreichenden Konzeption von Lesegewohnheiten. Eine Methode, die auf einer elaborierten theoretischen Konzeption basiert und die zudem ökonomisch in der Anwendung ist, stellt der Self Report Habit Index for Reading (SRHI-R) dar, der hier vorgestellt werden soll. Die dieser Skala zugrunde liegende Konzeptualisierung von Gewohnheiten umfasst neben der Wiederholung in der Vergangenheit auch das Fehlen von Bewusstheit, mentale Effizienz, das Fehlen von Kontrolle und den Ausdruck der eigenen Identität (Verplanken & Orbell, 2003). Diese ursprünglich aus der Gesundheitspsychologie stammende Gewohnheitskonzeption lässt sich theoretisch plausibel auch auf das Leseverhalten beziehen. So stellt Kintsch (1988) in seinem construction-integration Modell ähnliche Bezüge zwischen Leseverhalten und Leseverstehen her, wie es Verplanken et al. (2005) für die allmähliche Automatisierung von Verhaltensweisen formulieren.

Der SRHI-R stellt eine Adaption des Self-Report Habit Index (SRHI) nach Verplanken und Orbell (2003) dar. Letztere Skala umfasst zwölf Items und verfügt über eine gute interne Konsistenz von  $\alpha > .90$  und gute Test-Retest-Reliabilitäten (Verplanken et al., 2005). Ziel dieser Studie ist es erste Hinweise auf die Validität des SRHI-R zu gewinnen.

Methode: Bei der Adaption des wurde die Skala zunächst an den Lesebegriff angepasst und die Items ins Deutsche übersetzt. Aufgrund von theoretischen Überlegungen wurden zwei Items entfernt, sodass die Skala noch 10 Items umfasst. Die Stichprobe umfasste  $N = 1439$  Schülerinnen und Schülern der elften Klasse (54% weiblich; Alter:  $M = 17.24$ ,  $SD = 0.67$ ) aus 21 Schulen in Schleswig-Holstein.

Ergebnisse: Die korrigierten Trennschärfen und Itemschwierigkeiten liegen in einem akzeptablen bis sehr guten Bereich. Für den SRHI-R ergab sich ein sehr hohes Cronbachs Alpha von  $\alpha = 0.96$ . Die konfirmatorische Faktorenanalyse erbrachte für ein modifiziertes einfaktorielles Modell (zwei Residualkorrelationen zugelassen) mit den verbleibenden zehn Items eine gute Anpassung an die Daten:  $\chi^2(33) = 340.58$ ; CFI = .97; TLI = .95; RMSEA = .08; SRMR = .02. Alle Items luden signifikant ( $p < .001$ ) und substantiell ( $\lambda \geq .72$ ) auf dem latenten Faktor. Zudem stützen die Ergebnisse die Annahme strikter Messinvarianz über die Geschlechter und über die Profilerbestufen gemäß den von Chen (2007) vorgeschlagenen Kriterien.

Um erste Hinweise auf die externe Validität des SRHI-R zu erhalten, wurden Zusammenhangsmaße zu relevanten Konstrukten berechnet die alle in die erwartete Richtung weisen. Auch für die Gruppenvergleiche für Geschlecht und Profilerbestufenwahl waren die Ergebnisse erwartungsgemäß.

Fazit: Die Ergebnisse dieser Studie geben damit erste Hinweise auf die Validität des SRHI-R. Gerade in Large-Scale Untersuchungen bietet sich der Einsatz dieses Testverfahrens an, da es besonders ökonomisch in seiner Anwendung ist und dennoch eine elaboriertere theoretische Fundierung aufweist als die üblicherweise erfragte Verhaltensfrequenz. Auch ist von einer erhöhten Reliabilität im Gegensatz zu Messungen anhand nur eines oder nur weniger Items auszugehen.

Artelt, C., Naumann, J., & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Eds.), PISA 2009. Bilanz nach einem Jahrzehnt (pp. 73–112). Münster, New York, München, Berlin: Waxmann.

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. doi:10.1080/10705510701301834

Kintsch, W. (1988). The Role of Knowledge in Discourse Comprehension : A Construction-Integration Model, 95(2), 163–182.

Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of Reading Motivation and Their Relation to Reading Behavior and Competence. *Reading Research Quarterly*, 4(47), 427–463.

Verplanken, B., Myrbakk, V., & Rudi, E. (2005). The measurement of habit. In T. Betsch & S. Haberstroth (Eds.), *The routines of decision making* (pp. 231–247). Mahwah, NJ: Erlbaum.

Verplanken, B., & Orbell, S. (2003). Reflections on Past Behavior : A Self-Report Index of Habit Strength. *Journal of Applied Social Psychology*, 33(1), 1313–1330. doi:10.1111/j.1559-1816.2003.tb01951.x

Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth or their reading. *Journal of Educational Psychology*, 89(3), 420–432. doi:10.1037//0022-0663.89.3.420





Eine empirische Untersuchung integrativer Deutschlehrwerke der Sekundarstufe I  
(Ergebnisse einer Dissertation, PH Karlsruhe 2013)

In Anbetracht weltweiter Globalisierung und einer verstärkten Heterogenität der deutschen Gesellschaft besitzt interkulturelles Lernen im Rahmen des Deutschunterrichts als allgemeines Bildungsziel eine hohe politische und pädagogische Relevanz und dient der Persönlichkeitsbildung. Durch die Ähnlichkeit zwischen interkulturellen und literarischen Lernprozessen besitzt das interkulturelle Lernen zudem eine spezifische deutschdidaktische Relevanz.

Wie ein Blick in die Schulbuchforschung zeigt, gibt es bislang nur wenige Untersuchungen zum interkulturellen Lernen in und mit aktuellen Deutschlehrwerken. Dabei spiegeln Lehrwerke aktuelle fachdidaktische Konzeptionen wider, besitzen als Informationsquelle und Lernmedium eine wesentliche Sozialisationsfunktion und beeinflussen Unterricht und Lernprozesse in hohem Maße.

Um diese Forschungslücke zu schließen, analysiert die vorgestellte Untersuchung Textauswahl und Aufgabenstellungen in 15 integrativen Sprachlesebüchern der Klassenstufe 7 (Erscheinungsjahr ab 2005) für alle drei Schulzweige nach folgenden Fragestellungen:

1. Wie hoch ist der Anteil an Texten mit einem möglichen interkulturellen Potenzial?
2. Welches Bild der Fremde und welches Bild heterogener Lebenswelten vermittelt die in den Lehrwerken getroffene Textauswahl?
3. Welche dieser Texte und der sie begleitenden Aufgabenstellungen erscheinen für das interkulturelle Lernen geeignet?

Dabei geht es in doppelter Hinsicht um die Analyse einer symbolischen Konstruktion von Wirklichkeit: Sie zeigt sich sowohl auf der Ebene der Texte selbst als auch in der Textauswahl, die eine „mediale Praktik“ darstellt.

Die methodische Wahl fällt auf die qualitative Inhaltsanalyse nach Philipp Mayring, der ein systematisches, regelgeleitetes und theorieorientiertes Vorgehen zu qualitativen Fragestellungen aufzeigt, da die Inhaltsanalyse eine Verbindung von quantitativem und qualitativem Vorgehen darstellt.

Der Untersuchung liegt ein weiter, bedeutungsorientierter Kulturbegriff im Sinne Alois Wierlachers zugrunde; aus der Gesamtzahl aller 2133 in den Lehrwerken vorhandenen Texte wurden 550 Texte aufgrund des Handlungsschauplatzes, der ethnischen Herkunft der Autor(inn)en, der Sprache oder der Darstellung heterogener Lebenswelten als potenziell relevant ausgewählt. Zur qualitativen Untersuchung dieser Texte und der dazugehörigen Aufgabenstellungen wurde unter Berücksichtigung des aktuellen Forschungsstands in der Fachdidaktik Deutsch und der interkulturellen Germanistik ein theoriegeleitetes Kategoriensystem aufgestellt, das durch Explikationen, Beispiele und Erläuterung der Kategoriengrenzen operationalisiert und intersubjektiv überprüfbar ist; die erhobenen Daten wurden mithilfe einer Datenbank verwaltet und ausgewertet.

Die Auswertung zeigt: Die Textauswahl der Lehrwerke hat sich im Sinne einer kulturwissenschaftlichen Ausrichtung geöffnet und allen Lehrwerken liegt ihnen ein offener, weiter Kulturbegriff zugrunde, der Alltagskulturen mit einschließt. Dennoch zeigt sich an vielen Stellen die Gefahr, dass die Lesebücher kulturelle Stereotypen herausbilden und festigen:

- Der Blick in fremde Welten beschränkt sich überwiegend auf die großen europäischen Bildungskulturen und Nordamerika als Handlungsorte; diese weitgehende Beschränkung auf westliche Hochkulturen zeigt sich ebenfalls thematisch, stilistisch und im Bereich der Normen und Werte.
- Die eigene Gesellschaft präsentiert sich in den Lehrwerken überwiegend monokulturell: innergesellschaftliche Heterogenität nimmt sowohl in der Darstellung der eigenen als auch in der Darstellung fremder Gesellschaften einen geringen Raum ein und wird überwiegend problem- und nicht ressourcenorientiert behandelt. Auch in länderübergreifenden interkulturellen Begegnungen überwiegen Problemorientierung und klar differenzorientierte Sichten.

Die Textauswahl in ihrer Gesamtheit gibt somit durchaus Anlass darüber nachzudenken, inwiefern sowohl der Blick in fremde als auch in eigene Welten kritisch überdacht werden sollte; die Untersuchung könnte ein kleiner Anstoß dazu sein, eine Auseinandersetzung mit Konzepten der interkulturellen Bildung im Rahmen des Deutschunterrichts anzuregen.

Mittwoch,  
05.03.2014,  
11:15 - 13:00,  
PEG-Gebäude  
Raum 1.G102

Pamela Sammons (University of Oxford)

***Using Mixed Methods to Investigate the Influence of Family, Pre-school and School Experiences on Children's Development and Progress: Results from the EPPSE 3-16 Research in England***

Chair(s): Eckhard Klieme (DIPF)

The Effective Provision of Pre-school Primary and Secondary Education Project (EPPSE 3-16) is a major longitudinal study funded by the Department of Education in England (1996-2014). It adopted a mixed methods educational effectiveness design to follow a sample of around 3000 children from age 3 to 16 across different phases of education from pre-school to the end of compulsory secondary education. The presentation will outline the rationale for the study's design and illustrate how the research sought to combine both quantitative and qualitative approaches to identify and analyse the influences of various individual child, family (including the home learning environment), pre-school and school influences on children's cognitive and social behavioural development. The research used a variety of statistical approaches particularly multilevel modelling to study children's attainment, progress and social behavioural development. This strategy was used to take account of the effects of clustering in the data (children nested in pre-school settings and later in primary or secondary schools) and identify the strength of various predictors. The study shows that pre-school effects can be identified in the short, mid and longer term and also points to the importance of parenting as measured by the early years home learning environment (HLE), as well as of other predictors including parents' qualifications, SES and income in shaping children's educational outcomes. In addition, in depth case studies of individual pre-schools illuminate features of good practice in early years settings, and case studies of children and families who succeeded against the odds provide further qualitative evidence to complement the quantitative findings. EPPSE findings show that pre-school and school effects can be of particular importance in reducing the likelihood of a child being identified as having SEN in primary school and in ameliorating the adverse effects of multiple disadvantages. The research has had a major impact on policy and practice in England.

Mittwoch,

05.03.2014,

14:00 - 15:00,

Casino-Gebäude

Raum 823

Karin Zimmer (DIPF)

## ***Bildungspotenziale: Investitionen und Erträge im Lebenslauf***

Chair(s): Karin Zimmer (Leibniz-Forschungsverbund Bildungspotenziale, DIPF), Marcus Hasselhorn (DIPF, Leibniz-Forschungsverbund Bildungspotenziale)

Diskutant/in: Jürgen Baumert (Max-Planck-Institut für Bildungsforschung)

Die Ausgestaltung von Bildungsangeboten und die Leistungsfähigkeit des Bildungswesens sind für die wirtschaftliche Entwicklung, den kulturellen Reichtum und den sozialen Zusammenhalt in modernen Gesellschaften von hoher Bedeutung. Aus individueller Perspektive stärkt Bildung die Möglichkeit zu gesellschaftlicher Teilhabe und eigenbestimmter Lebensgestaltung. Der Leibniz-Forschungsverbund Bildungspotenziale versucht, Potenziale von und für Bildung zu identifizieren und zu klären, wie sie besser nutzbar gemacht werden können. Auf individueller, institutioneller und gesellschaftlicher Ebene werden Ansatzpunkte für tragfähige Konzepte und erfolgversprechende Reformen gesucht.

In der bildungspolitischen Diskussion wird dabei häufig die These vertreten, dass Bildungsinvestitionen vor allem dann wirksam sind, wenn sie früh erfolgen; zudem seien Investitionen in familiäre Bildung ertragreicher als Investitionen in Einrichtungen des Bildungssystems. Solche Thesen bergen - unabhängig von ihrer empirischen Evidenz - das Risiko verkürzter Rezeption und falscher Schlüsse. Das Symposium präsentiert empirische Forschungsperspektiven zu Bildungsinvestitionen und Bildungserträgen über den Lebenslauf hinweg aus der Sicht verschiedener Fachdisziplinen.

Im Bereich der frühen Bildung untersuchen C. Katharina Spieß und Pia S. Schober (DIW Berlin) aus ökonomischer Perspektive die Wirksamkeit von Bildungsinvestitionen. Olaf Köller (IPN) stellt aus psychologischer Blickrichtung die schulische Bildung in den Mittelpunkt. Vom Standpunkt der Soziologie aus richten Heike Solga und Paula Protsch (WZB) die Aufmerksamkeit auf die berufliche Bildung. Aus der Sichtweise der Erziehungswissenschaft lenken schließlich Josef Schrader (Tübingen und DIE) und Annika Goeze (DIE) den Fokus auf die Erwachsenenbildung. Zusammengenommen demonstrieren die vier Einzelbeiträge des Symposiums, wie im Leibniz-Forschungsverbund Bildungspotenziale durch eine multidisziplinäre Herangehensweise die Frage nach Investitionen und Erträgen im Lebenslauf angemessen bearbeitet und tragfähige Ansatzpunkte für wirksame Interventionen identifiziert werden können.

Eine Einschätzung des dadurch entstehenden Potenzials für die empirische Bildungsforschung durch Jürgen Baumert (MPI für Bildungsforschung) rundet das Symposium ab.

## **Investitionen in frühkindliche Bildung und Betreuung – Indirekte Effekte über Wohlbefinden und Erwerbstätigkeit von Müttern**

Insbesondere in der pädagogischen Literatur existieren eine Vielzahl von Studien, welche die Bedeutung frühkindlicher Bildung und Betreuung auf die kindliche Entwicklung untersuchen (statt vieler Anders, 2013). Dabei wird sehr häufig von direkten Effekten der Qualität auf die kindliche Entwicklung ausgegangen. Daneben stehen vor allem zahlreiche ökonomische Studien, welche zeigen, dass ein Angebot außerfamiliärer Bildung und Betreuung im frühkindlichen Bereich die Vereinbarkeit von Familie und Beruf insbesondere für Mütter erleichtern kann (vgl. Anderson et al., 2002). Diese Studien fokussieren üblicherweise lediglich darauf, dass die Erwerbstätigkeit von Müttern mit einem zunehmenden Bildungs- und Betreuungsangebot steigt. In diesem Beitrag soll es darum gehen mögliche direkte und indirekte Effekte von Investitionen in eine außerfamiliäre frühe Bildung und Betreuung auf die Kinder und ihre Familien zu diskutieren. In diesem Kontext werden zwei Studien beispielhaft vorgestellt, die auf indirekte Wirkungszusammenhänge hinweisen. Zum einen wird untersucht, inwiefern frühkindliche Bildungs- und Betreuungsangebote das mütterliche Wohlbefinden signifikant beeinflussen können (Schober & Schmitt, im Erscheinen). Hier zeigen sich für Ost- und Westdeutschland deutliche Zusammenhänge. Dies ist für die Bildungsforschung auch insofern von besonderer Bedeutung, da andere Studien gezeigt haben, dass die Zufriedenheit der Mutter, einen Einfluss auf die kindliche Entwicklung hat (z.B. Berger & Spiess, 2011). Zum anderen geht es darum, wie neben dem bloßen Angebot an Bildungs- und Betreuungseinrichtungen deren pädagogische Qualität mit Faktoren der mütterlichen Erwerbstätigkeit zusammenhängen - basierend auf der Hypothese, dass Eltern an einer guten Qualität interessiert sind (Schober & Spiess, 2013). Es stellt sich demnach die Frage, inwiefern bessere Bildungsqualitäten mit Veränderungen in der mütterlichen Erwerbstätigkeit zusammenhängen, die dann z.B. über Veränderungen im Erwerbseinkommen und damit verbunden höheren materiellen Ressourcen wiederum auf die kindliche Entwicklung wirken können (z.B. Ruhm, 2009). Es zeigt sich, dass insbesondere bei Müttern von Kindern unter drei Jahren Erwerbsmerkmale und Merkmale der Strukturqualität positiv zusammenhängen, sofern es sich von den Eltern beobachtbare Indikatoren handelt. Die Analysen basieren auf den Daten des Sozio-oekonomischen Panel (SOEP) und Familien in Deutschland (FiD), die mit Daten der amtlichen Kinder- und Jugendhilfestatistik auf der Ebene der Jugendamtsbezirke gematcht wurden.

Anders, Y. (2013): Stichwort: Auswirkungen frühkindlicher, institutioneller Bildung und Betreuung, in: Zeitschrift für Erziehungswissenschaft (16, 2: 237-275).

Anderson, P. M. und Levine, P. B. (2002): Childcare and mothers' employment decisions, in: D. Card and R. M. Blank (Eds.), Finding jobs. (S. 420-462), New York: Russel Sage Foundation.

Berger, Eva und C. Katharina Spiess (2011): Maternal life satisfaction and child outcomes: are they related?, in: Journal of Economic Psychology (32: 142-158).

Ruhm, Christopher J. (2009): Maternal employment and child development, in: Handbook of Families and Work, University Press of America, 331-408.

Schober, Pia und C. Katharina Spiess (2013): Does local day-care quality matter for maternal employment? Evidence for West and East Germany, Berlin, mimeo.

Schober, Pia S. und Christian Schmitt (2013): Day-care expansion and parental subjective wellbeing: Evidence from Germany, Berlin, mimeo.

Trotz aller Strukturformen des deutschen Bildungssystems in den letzten 15 bis 20 Jahren werden Schülerinnen und Schüler am Ende der 4. bzw. 6. Jahrgangsstufe auf unterschiedliche Schulformen der Sekundarstufe I verteilt. Bei aller Modernisierung des Systems bleibt diese Praxis von der Idee getragen, dass eine Aufteilung nach der Primarstufe zu günstigeren Effekten auf der Schülerseite führt als ein längeres oder gar dauerhaftes gemeinsames Lernen. Dabei ist die empirische Lage hinsichtlich der positiven Effekte der Leistungshomogenisierung international (z. B. Hattie, 2009) und national (Köller & Baumert, 2012) nicht eindeutig. Mit dem vorliegenden Beitrag wird zum einen intendiert, einen Überblick über die Forschung zu institutionellen Effekten auf Schüleroutcomes zu geben. Weiterhin soll auf der Basis eigener Forschungsarbeiten die empirische Befundlage zu Effekten der Differenzierung auf schulisches Lernen erweitert werden. Dabei werden die Mathematik- Orthografie- und Leseleistungen sowie die Intelligenzwerte zweier Schülerstichproben, die jeweils in der 5. und 7. Jahrgangsstufe erhoben wurden, berücksichtigt. Bei der Analyse der gewonnenen Daten wurde versucht, Effekte der Klassenzusammensetzung (Kompositionseffekt: Leistungsstärke, Intelligenz, Anteile der Mädchen, sozioökonomischer Status, Migrationshintergrund) und der Schulform (Institutioneller Effekt: Gymnasium vs. Nichtgymnasium) zu trennen. Basierend auf früheren Befunden auf früheren Befunden (Köller et al., 2013) erwarteten wir stärkere institutionelle als Kompositionseffekte.

Für die vorliegende Arbeit wurden Daten aus dem Projekt „Entwicklung und Implementierung eines neuen Konzeptes zur Eingliederung Jugendlicher in die Berufs- und Arbeitswelt in Schulen mit erhöhtem Förderbedarf“ (EIKA) verwendet. Zur Beantwortung unserer Fragestellung wurden zwei Kohorten aus dem EIKA-Projekt berücksichtigt. Bei der Kohorte 1 handelt es sich um N = 1001 Schülerinnen (50.1%) und Schüler (49.9%) aus 57 Klassen der 5. Jahrgangsstufe (T1; 31% Gymnasiasten), die zu Beginn des Schuljahres 2004/2005 untersucht wurden. Eine zweite Erhebung (T2) fand in dieser Stichprobe zu Beginn der 7. Jahrgangsstufe im Schuljahr 2006/2007 statt. Die zweite Kohorte besteht aus N = 842 Schülerinnen (49.7%) und Schülern (50.3%) aus 39 Klassen der 7. Jahrgangsstufe (30.6% Gymnasiasten), die parallel zu Kohorte 1 zu Beginn des Schuljahres 2006/2007 (T1) getestet wurden. Die zweite Datenerhebung (T2) in dieser Stichprobe fand zu Beginn der 9. Jahrgangsstufe im Schuljahr 2008/2009 statt. Die Schülerinnen und Schüler beider Kohorten stammen aus denselben Schulen (insgesamt 11). In beiden Kohorten ist der Anteil der Schülerinnen und Schüler mit Migrationshintergrund vergleichsweise hoch (um 50%). In beiden Kohorten wurden zu beiden Messzeitpunkten hinreichend reliable standardisierte Mathematik-, Lesetests aus der Hamburger Lehrausgangslagenuntersuchung (LAU; vgl. u.a. Lehmann, Peek, Gänsfuß, & Husfeldt, 2001) und Intelligenztests (CFT 20; Weiß, 1997) sowie Schüler- und Elternfragebögen eingesetzt. Die Testung erfolgte durch vorher geschulte Lehrkräfte an jeweils zwei Tagen in je zwei Schulstunden.

Für die statistischen Analysen wurden zunächst fehlende Werte mittels multipler Imputation (50 Datensätze) im Softwarepaket Mplus (Version 6.12, Muthén & Muthén, 2010) geschätzt. Entsprechend wurden alle späteren Analysen fünfzimal durchgeführt und gemittelt. Zur Trennung von individuellen, institutionellen und Kompositionseffekten wurden längsschnittliche Mehrgruppen-Mehrebenenanalysen in Mplus durchgeführt. Die Analysen zeigen durchgängig, dass alle geschätzten Regressionsparameter invariant über die beiden Stichproben sind. Auf Individualebene hat das Vorwissen zu T1 den stärksten Effekt auf die Leistung zu T2. Robuste Effekte auf Individualebene zeigen sich zusätzlich für die Intelligenz (zu T1) auf Leistungen zu T2. Im Falle sprachlicher Leistungen ist weiterhin das Geschlecht (Mädchen weisen höhere Leistungen auf) prädiktiv. Hinsichtlich aller Kompositionsmerkmale ergibt sich, dass ihre Effekte nach Kontrolle von Schulform verschwinden. Durchgängig lassen sich signifikante institutionelle Effekte auf die T2-Maße zugunsten des Gymnasiums nachweisen. Die Ergebnisse werden mit Blick auf die Rolle differenzieller Entwicklungsmilieus diskutiert.

Hattie, J. A. C. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. Oxon: Routledge.

Köller, O. & Baumert, J. (2012). Schulische Leistungen und ihre Messung. In W. Schneider & U. Lindenberger (Hrsg.), Entwicklungspsychologie (7. Auflage, S. 645-661). Weinheim: Beltz/PVU.

Köller, O., Schütte, K., Zimmermann, F., Retelsdorf, J. & Leucht, M. (2013). Starke Klasse, hohe Leistungen? Die Rolle der Leistungsstärke der Klasse für die individuellen Mathematik- und Leseleistungen in der Sekundarstufe I. Psychologie in Erziehung und Unterricht, 60, 184-197.

Lehmann, R. H., Peek, R., Gänsfuß, R. & Husfeldt, V. (2001). LAU 9: Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9 – Ergebnisse einer längsschnittlichen Untersuchung in Hamburg. Hamburger Schulbehörde.

Muthén, L. K. & Muthén, B. (2010). Mplus. Statistical analysis with latent variables. User's guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Weiß, H. (1997). Grundintelligenztest CFT 20. Göttingen: Hogrefe.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
Casino-Gebäude  
Raum 823



## ***Hindernisse für eine zweite Chance für leistungsschwache Jugendliche auf dem Ausbildungsmarkt – und wie diese überwunden werden können***

Es ist aus vielfältigen Untersuchungen und Statistiken bekannt, dass Hauptschüler/innen heute kaum noch eine Chance haben, einen Ausbildungsplatz zu bekommen (vgl. Solga & Menze, 2013). Zumeist wird dies mit der geringeren oder fehlenden Ausbildungsreife dieser Schulabgänger/innen begründet. Unklar ist jedoch, wie Betriebe die „Ausbildungsreife“ und wie sie rekrutieren bestimmen. Welche Diskriminierungs- und Diskreditierungsprozesse spielen für hierbei eine Rolle (vgl. Solga, 2002, 2005). In dem Beitrag gehen wir daher mit Blick auf die Rekrutierungsseite drei Fragen nach: 1. Wie rekrutieren Betriebe, welche Kriterien nutzen sie? 2. Wie haben sich die Anforderungen an Ausbildungen in den letzten Jahrzehnten verändert, in welchen Berufen gibt es tatsächlich Verschiebungen der kognitiven Anforderungen nach oben? 3. Wie gelingt es Hauptschüler/innen dennoch, in eine Ausbildung direkt nach der Schule überzugehen – wie überwinden sie die mit den Antworten auf Frage 1 und 2 spezifizierten Hindernisse?

Mit den gängigen Daten, die zumeist die Übergangsbioographien der Jugendlichen abbilden und dann als abhängige Variable deren Ausbildungserfolg (Übergang in eine Ausbildung) untersuchen, ist die Beantwortung dieser drei Fragen nicht möglich. In den Befunden dieses Vorgehens gibt es eine Konfundierung des Suchverhaltens der Jugendlichen und des Rekrutierungsverhaltens der Betriebe. Dies wird auch als „revealed outcome approach“ bezeichnet (Sen, 1986). In dem Beitrag werden daher Befunde aus drei Untersuchungen vorgestellt, die diesem Problem gerecht werden und die Rolle der Betriebe zu spezifizieren vermögen.

Der Beitrag ist eine systematische Zusammenfassung von drei unterschiedlichen Untersuchungen. Für die Frage 1 wurden Feldexperimente, so genannte Correspondence Test, durchgeführt (siehe Riach & Rich, 2002). Hier wurden fiktive Bewerbungen auf reale Ausschreibungen und an reale Firmen versendet. Hinsichtlich der 2. Frage wurden Ausbildungsordnungen im mittleren und unteren Segment des Ausbildungsmarktes untersucht (Protsch, 2012, 2013). An diesen Ausbildungsordnungen sind Betriebe über den Hauptausschuss am Bundesinstitut für Berufsbildung (BIBB) maßgeblich mitbeteiligt. Für die 3. Frage werden Befunde einer Längsschnittanalyse Niedersächsischer Hauptschüler/innen (aus ca. 90 Klassen von 46 Schulen) verwendet (vgl. Kohlrausch & Solga, 2012; Solga, Baas, & Kohlrausch, 2011; Solga & Kohlrausch, 2013). Dabei handelt es sich um eine Interventionsstudie (mit einem Kontrollgruppendesign), in der eine frühzeitige und regelmäßige Einbindung von Hauptschüler/innen seit der 8. Klasse in betriebliche Kontexte erfolgte.

Kohlrausch, B./H. Solga (2012). Übergänge in die Ausbildung: Welche Rolle spielt die „Ausbildungsreife“? Zeitschrift für Erziehungswissenschaft 15: 753-773.

Protsch, P. (2013): Höhere Anforderungen in der beruflichen Erstausbildung? WSI-Mitteilungen 66: 15-22.

Protsch, P. (2012). Segmentierte Ausbildungsmärkte. Berufliche Chancen von Hauptschülerinnen und Hauptschülern im Wandel. Dissertation. Freie Universität Berlin.

Riach, P.A./J. Rich (2002). Field experiments of discrimination in the market place. The Economic Journal 112: F480-F518.

Sen, A. (1986). Behavior and the concept of preference. In: J. Elster (Ed.), Rational choice. Oxford, UK: Blackwell, S. 60–81.

Solga, H. (2002). „Stigmatisierung by negative selection“. European Sociological Review 18: 159-178.

Solga, H. (2005). Ohne Abschluss in die Bildungsgesellschaft. Opladen: Verlag Barbara Budrich.

Solga, H./M. Baas/B. Kohlrausch (2011). Übergangschancen benachteiligter Hauptschüler und Hauptschülerinnen. Evaluation der Projekte „Abschlussquote erhöhen – Berufsfähigkeit steigern 2“ und „Vertiefte Berufsorientierung“. IAB-Forschungsbericht 6/2011. Nürnberg: IAB. (<http://doku.iab.de/forschungsbericht/2011/fb0611.pdf>).

Solga, H./B. Kohlrausch (2013). How low-achieving German youth beat the odds and gain access to vocational training – Insights from within-group variation. European Sociological Review 29 (5): 1068-1082.

Solga, H./L. Menze (2013). Der Zugang zur Ausbildung: Wie integrationsfähig ist das deutsche Berufsbildungssystem? WSI-Mitteilungen 66 (1): 5-14.



## **Implementation pädagogischer Innovationen in der Aus- und Fortbildung von Lehrkräften – Effekte unterschiedlicher Strategien der Implementierung**

Um die Potenziale des Bildungssystems besser als bisher nutzen zu können, ist es eine fundamentale Frage, wie wissenschaftliche Erkenntnisse und darauf beruhende Innovationen in die Praxis implementiert werden können. Insbesondere gilt dies für die Aus- und Fortbildung von Lehrkräften, der allgemein eine Schlüsselfunktion bei der Weiterentwicklung des Bildungssystems zugeschrieben wird (z.B. Hattie, 2009). Angesichts der Notwendigkeit des Aufbaus professioneller Handlungskompetenzen stellt sich die Frage nach wirksamen (Aus)Bildungskonzepten für ihre zielgerichtete Unterstützung. Die hier vorgestellte Studie geht der Frage nach, welchen Einfluss bestimmte Formen videofallbasierten Lernens, die sich in experimentellen Studien als wirksam und nachhaltig erwiesen haben, noch haben, wenn sie mit unterschiedlichen Graden pädagogischer Gestaltungsautonomie in die alltägliche Bildungspraxis implementiert werden. Das in mehreren Forschungs- und Entwicklungsprojekten entwickelte Konzept videofallbasierten Lernens zielt auf die Förderung der Kompetenz zur Diagnose von Lehr-Lernsituationen. Unter dieser Kompetenz wird die Fähigkeit verstanden, Unterricht u.a. aus a) verschiedenen Perspektiven der handelnden Akteure deuten und b) mithilfe allgemein-pädagogischer, fachdidaktischer oder pädagogisch-psychologischer Konzepte, Modelle bzw. Theorien analysieren zu können (vgl. Goeze et al., 2013, S.82).

Um die Frage nach den Effekten unterschiedlicher Strategien der Implementierung zu untersuchen, wurde eine quasi-experimentelle Interventionsstudie im Feld der zweiten Phase der Lehrerbildung mit Fachleitern (Unterrichtsfach Englisch; N=21) und ihren Studienreferendaren (N=261) realisiert. Systematisch variiert wurde der Grad pädagogischer Gestaltungsautonomie (gering/mittel/hoch) (vgl. Penuel & Gallagher, 2009; Raudenbush, 2007), der Fachleitern gewährt wurde für die Umsetzung in insgesamt je zweitägigen, videofallbasierten Lernangeboten in Referendarskursen. Die Fachleiter mit geringer Autonomie adoptierten das Konzept gleichsam 1:1, diejenigen mit mittlerer Autonomie konnten es bei Wahrung von Kernelementen adaptieren, Fachleiter mit hoher Autonomie konnten es eigenständig designen. Zur Erfassung der Kompetenzentwicklung bei den Studienreferendaren dienten schriftliche Fallanalysen authentischer videographierter Lehr-Lernsituationen, die vor und nach der Intervention geschrieben, inhaltsanalytisch ausgewertet und quantifiziert werden.

Erste Befunde aus varianzanalytischen Auswertungen zeigen, dass Studienreferendare statistisch und praktisch signifikant positivere Kompetenzentwicklungen erzielen unter Angeboten videofallbasierten Lernens, die dem evidenzbasierten Originalkonzept am ähnlichsten sind (geringe pädagogische Gestaltungsautonomie von Fachleitern bei der Konzeptumsetzung im Vergleich zu Studienreferendaren, deren Fachleiter höhere pädagogische Gestaltungsautonomie in der Umsetzung (mittel/hoch) gewährt wurden). Entwicklungsunterschiede von Studienreferendaren zeigen sich in den Teilkompetenzen (a) Perspektivenübernahme (z.B. unter geringer vs. hoher pädagogischer Gestaltungsautonomie: MDif=1.0; SEboot=.476; p=.024; Cohen's d=.335) und (b) Theorieverwendung (z.B. unter geringer vs. hoher pädagogischer Gestaltungsautonomie: MDif=3.2; SEboot=.676; p=.001; Cohen's d=.806).

Die vorgestellten Ergebnisse der Implementationsstudie werden abschließend unter der allgemeinen Fragestellung diskutiert, unter welchen Bedingungen der Nutzung empirischer Forschungsbefunde die Potenziale des Bildungssystems sinnvoll ausgeschöpft werden können.

Goeze, A., Hetfleisch, P., & Schrader, J. (2013). Wirkungen des Lernens mit Videofällen bei Lehrkräften: Welche Rolle spielen instruktionale Unterstützung, Personen- und Prozessmerkmale? Zeitschrift für Erziehungswissenschaft, 16, 79-113.

Hattie, J. A. C. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. Oxon: Routledge.

Raudenbush, S. W. (2007). Designing field trials of educational innovations. In B. L. Schneider & S.-K. McDonald (Eds.), Scale-up in Education: Ideas in Practice (Vol. 2, pp. 23-40). Lanham, MD: Rowman & Littlefield.

Penuel, W. R., & Gallagher, L. P. (2009). Preparing teachers to design instruction for deep understanding in middle-school earth science. Journal of the Learning Sciences, 18, 461-508.

Thomas Bäumer (Otto-Friedrich-Universität Bamberg)

## ***Das Nationale Bildungspanel als interdisziplinäre Forschungsinfrastruktur – Ergebnisse aus der Startkohorte 3 „Wege durch die Sekundarstufe I“***

Chair(s): Thomas Bäumer (Universität Bamberg)

Diskutant/in: Jutta von Maurice (Universität Bamberg)

Das Nationale Bildungspanel (NEPS) untersucht Bildungsprozesse von der frühen Kindheit bis ins hohe Erwachsenenalter. Dafür stellt es umfangreiche Daten für die Bildungsforschung bereit, die über Befragungen und Testungen an sechs sogenannten Startkohorten erhoben werden, die sich in je spezifischen Abschnitten des Bildungsverlaufs befinden. NEPS wurde mit der expliziten Absicht gegründet, multiperspektivische und interdisziplinäre Ansätze zu ermöglichen. In dem Symposium sollen am Beispiel der Startkohorte 3, die im Jahr 2010 mit Schülerinnen und Schülern der Jahrgangsstufe 5 startete, die vielfältigen Analysemöglichkeiten, die der Datensatz bietet, exemplarisch dargestellt werden. Dementsprechend folgen die einzelnen Beiträge der Kolleginnen und Kollegen aus verschiedenen Standorten des NEPS-Konsortiums mindestens einer jeweils eigenen disziplinären Perspektive. Es werden Daten aus der Klasse 5 und der Klasse 6 analysiert, d. h. aus den Wellen 1 und 2 der Startkohorte 3, die der wissenschaftlichen Öffentlichkeit als Scientific Use File aktuell bzw. in Kürze zur Verfügung stehen.

Im ersten Beitrag wird zunächst die Startkohorte 3 vorgestellt, die im Fokus der Analysen der weiteren Beiträge steht. Als spezifische Fragestellung wird die Rechtschreibkompetenz mit Hilfe sowohl quantitativer als auch qualitativer Analysen aus einer didaktischen Perspektive ins Blickfeld gerückt. Der zweite Beitrag befasst sich mit außerschulischen Bildungsangeboten, die im NEPS neben der schulischen Bildung ebenfalls erfasst werden. Dabei kommen soziologische wie auch pädagogische Perspektiven zum Tragen. Im dritten Beitrag werden aus Perspektive der soziologisch geprägten Integrationsforschung die Bildungsmuster von Kindern aus Zuwandererfamilien untersucht. Der vierte Beitrag schließlich beschäftigt sich mit dem Thema der „Health Literacy“, also gesundheitsbezogener Kompetenzen und deren Einfluss auf das Gesundheitsverhalten.

Die vier Beiträge verdeutlichen – obwohl nur ein kleiner Ausschnitt aus einer der sechs Startkohorten des NEPS – die vielfältigen theoretischen und methodischen Perspektiven, die sich mit diesem Datensatz verbinden.

Mittwoch,

05.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 311



Sarah Frahm (Universität Hamburg), Martin Goy (TU Dortmund, IFS), Stephan Jarsinski (TU Dortmund, IFS), Anna Prosch (Universität Hamburg), Inge Blatt (Universität Hamburg), Wilfried Bos (TU Dortmund, IFS), Michael Kanders (TU Dortmund, IFS)

### **NEPS Startkohorte 3: Anlage, Erhebungsprogramm und erste etappenspezifische Ergebnisse zur Rechtschreibkompetenz aus didaktischer Perspektive**

In diesem Beitrag werden zunächst in knapper Übersicht Anlage und Erhebungsprogramm der Startkohorte 3 im Rahmen der Etappe 4 des Nationalen Bildungspanels (NEPS) präsentiert. Mit Fokus auf die etappenspezifischen Forschungsfragen werden anschließend Ergebnisse zur Kompetenzdomäne Orthographie vorgestellt.

Etappe 4 des NEPS verfolgt die Wege der Zielpersonen durch die Sekundarstufe I und die Übergänge in die Sekundarstufe II, die berufliche Bildung und den Arbeitsmarkt. Im Rahmen des Mehrkohorten-Designs des NEPS begleitet die Etappe 4 den ersten Erhebungsabschnitt der Startkohorte 3 von Klassenstufe 5 bis 10 (Blossfeld, Schneider, & Doll, 2009; Blossfeld, Roßbach, & von Maurice, 2011). Die erste Haupterhebung dieser Startkohorte wurde 2010 mit einer Stichprobe von rund 6800 Fünftklässlerinnen und Fünftklässlern durchgeführt. Im Herbst 2013 findet die Haupterhebung in Klassenstufe 8 statt.

In den Erhebungen der Etappe 4 stehen Kompetenztests und Befragungen der Schülerinnen und Schüler im Mittelpunkt, die durch die Erhebung von Kontextinformationen bei deren Eltern (mittels telefonischer Interviews) sowie bei den Klassen-, Deutsch- und Mathematiklehrkräften und Schulleitungen (mittels Fragebögen) flankiert werden (Frahm et al., 2011). Die Inhalte dieser Erhebungen werden in erster Linie durch die Säulen des NEPS bereitgestellt. In Etappe 4 liegt darüber hinaus ein etappenspezifischer Forschungsschwerpunkt auf der Entwicklung von Schriftsprachkompetenz in der Sekundarstufe, wobei hier im Einzelnen die Entwicklung der Rechtschreibkompetenz, die Unterrichtsqualität im Fach Deutsch sowie das Engagement der Schülerinnen und Schüler in den Bereichen Lesen und Rechtschreibung in den Blick genommen werden (Frahm et al., 2011; Strietholt et al., in Druck).

In der NEPS-Studie wird die Rechtschreibkompetenz mit dem Sprachsystematischen Rechtschreibtest (SRT) erhoben, der diese auf der Ebene des ganzen Wortes und von fünf Teilkompetenzen erfasst (Blatt et al., 2011, Hinney, 1997). Dem Test liegt ein validiertes Kompetenzmodell zugrunde. Ergänzend wurden als Hintergrundvariablen im allgemeinen Schülerfragebogen u.a. Geschlecht, SES und Schulform erfragt (Frahm et al., 2011). Die Fragen im spezifischen Rechtschreibfragebogen beruhen auf dem Konstrukt „Rechtschreib-Engagement“, das auf der Wert-Erwartungstheorie fußt und in Analogie zum „Reading Engagement“ konzipiert wurde (Guthrie & Wigfield, 2000).

In den Leistungsdaten steckt aus didaktischer Sicht weitaus mehr, als in der Regel in großen empirischen Leistungsstudien ausgewertet und berichtet wird. Um das Potential der Studien für erziehungswissenschaftliche und insbesondere didaktische Fragestellungen auszuschöpfen, muss jedoch die quantitative Auswertung im Sinne der Methodenkombination durch eine vertiefende qualitative ergänzt werden.

Den Anstoß für die vertiefenden Analysen geben die Ergebnisse zur mehrdimensionalen Kompetenzmodellierung sowie die Ermittlung der Personenfähigkeit und Itemschwierigkeit mit den Daten der Haupterhebung K5. Daran anschließend erfolgten weiterführende quantitative Analysen, und zwar die Berechnung der Leistung in Prozent korrekt nach Quartilen sowie die Ermittlung von Rechtschreibprofilen mit LCA-Analysen (Frahm & Blatt, in Vorbereitung). Weiterhin wurde die Zahl der Variantenschreibungen für die Gesamtgruppe, nach Quartilen und Profilen errechnet (Blatt & Frahm, 2013). Für die anschließenden weiterführenden explorativen Analysen wurden auch die Daten aus den Entwicklungs- bzw. Pilotstudien, die den Haupterhebungen vorausgehen, herangezogen. Sie eignen sich in besonderer Weise für eine tiefergehende Betrachtung. Unter didaktischer Zielsetzung erfolgten Analysen von Fehlerschwerpunkten für die Gesamtgruppe, nach Quartilen und für unterschiedliche Profile. In diesem Beitrag werden die Ergebnisse dieses Forschungsprozesses zusammenfassend vorgestellt. Auf die ebenfalls untersuchten Zusammenhänge zu Hintergrundvariablen wird aus Zeitgründen nicht eingegangen. Im Ausblick wird aufgezeigt, welche Chancen die Analyseergebnisse bieten, um die Entwicklung didaktischer Konzepte auf eine durch Empirie gestützte Grundlage zu stellen. Das kann zu mehr Bildungsgerechtigkeit beitragen, da die Rechtschreibkompetenz immer noch ein ausschlaggebender Faktor für die Schullaufbahn und die beruflichen Chancen ist.

Blossfeld, H.-P. Roßbach, H.-G. & von Maurice J. (Hrsg.). (2011). Education as a lifelong process. The German National Educational Panel Study (NEPS). Wiesbaden: VS. (Zeitschrift für Erziehungswissenschaft, Sonderheft 14).

Blossfeld, H.-P., Schneider, T., & Doll, J. (2009). Methodological advantages of panel studies: Designing the new National Educational Panel Study (NEPS) in Germany. *Journal for Educational Research Online*, 1(1), 10–32.

Blatt, I., Voss, A., Kowalski, K. & Jarsinski, S. (2011). Messung von Rechtschreibleistung und empirische Kompetenzmodellierung. In U. Bredel & T. Reißig (Hrsg.), *Weiterführender Orthographieunterricht*. DTP 5 (S. 226–256). Baltmannsweiler: Schneider Verlag Hohengehren.

Blatt, I. & Frahm, S. (2013). Explorative Analysen zur Entwicklung der Rechtschreibkompetenz im Rahmen der NEPS-Studie (Klassenstufe 5–7). *Didaktik Deutsch*, 34, 12–36.

Frahm, S., Goy, M., Kowalski, K., Sixt, M., Strietholt, R., Blatt, I., Bos, W. & Kanders, M. (2011). Transition and development from lower secondary to upper secondary school. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (S. 217–232). (Zeitschrift für Erziehungswissenschaft, Sonderheft 14).

Frahm, S. & Blatt, I. (in Vorbereitung). *Von der Grundlagen- zur Anwendungsforschung: Der Weg zur evidenzbasierten Nutzung eines Rechtschreibtests aus dem Large-Scale-Bereich im Unterricht*.

Guthrie, J. T. & Wigfield, A. (2000). Engagement and motivation in reading. In M. Kamil, P. Mosenthal, D. Pearson and R. Barr (Hrsg.), *Handbook of reading research (Volume III)* (S. 403–422). Mahwah, NJ: Lawrence Erlbaum Associates.

Hinney, G. (1997). Neubestimmung von Lerninhalten für den Rechtschreibunterricht. Ein fachdidaktischer Beitrag zur Schriftaneignung als Problemlöseprozess. Frankfurt/M.: Lang.

Strietholt, R., Naujokat, K., Mai, T., Kretschmer, S., Jarsinski, S., Goy, M., Frahm, S., Kanders, M., Bos, W., Blatt, I. (in Druck). The National Educational Panel Study (NEPS) in Germany. An overview of design, research options, and access with a focus on lower secondary school. *European Educational Research Journal*.

Andrea Hopf (Universität Gießen), Ludwig Stecher (Universität Gießen)

## ***Extended Education – Bildungsangebote und ihre Wirkung außerhalb formeller Lernkontexte. Empirische Ergebnisse auf Basis der Daten aus der Startkohorte 3 im Nationalen Bildungspanel (NEPS)***

Das Lernen als aktive Form des Wissenserwerbes beschränkt sich nicht ausschließlich auf formale Bildungsinstitutionen, wie etwa die Schule, sondern findet (zunehmend) auch außerhalb dieser Bildungskontexte in non-formalen und informellen Lernumgebungen statt. Für diese erweiterten Bildungsmöglichkeiten hat sich in der internationalen Bildungsforschung der Begriff „Extended Education“ etabliert. In den letzten Jahren ist eine deutliche Steigerung des öffentlichen wie wissenschaftlichen Interesses an diesem Forschungsgebiet erkennbar. Dabei rücken vor allem Fragen von Qualität, Nutzung und nicht zuletzt der Wirkung solcher neuen Lernformen mehr und mehr in den Fokus wissenschaftlicher Untersuchungen – so etwa auch im Rahmen des Nationalen Bildungspanels (NEPS). Die bis dato veröffentlichte Befunde (s. dazu z.B. Ecarius et al., 2013) belegen die wachsende Bedeutsamkeit von non-formalen und informellen Bildungsangeboten auch auf internationaler Ebene. Gleichzeitig zeigt sich, dass die Entwicklung und Erprobung von Instrumenten, Methoden und Verfahren zur Erforschung, Beschreibung und Messung von Kompetenzzuwächsen und Bildungswirkungen in nicht formalisierten Lernkontexten bisher ein Forschungsdesiderat darstellt. Aufgabe des Beitrags ist es, diese Forschungsdesiderate mithilfe der Daten des NEPS zu erfüllen.

Dabei geht der Beitrag konkret den Fragen nach, ob und inwiefern Angebote im Bereich der Extended Education tatsächlich genutzt werden, und von wem (das heißt, von welchen SchülerInnen-Gruppen), sowie wie die pädagogische Qualität seitens der Schülerinnen und Schüler eingeschätzt wird. Darüber hinaus steht die Wirkung der Teilnahme an diesen Angeboten auf die Schülerinnen und Schüler im Vordergrund der Untersuchung.

Als Datengrundlage dient die Startkohorte 3 (Klasse 5) des Nationalen Bildungspanels (NEPS). Auf der Datenbasis der ersten und zweiten Erhebungswelle wird mithilfe multivariater Analysen untersucht, ob sich gruppenspezifische Unterschiede im Hinblick auf die Nutzung und die Wirkung von Angeboten der Extended Education auf die Teilnehmerinnen und die Teilnehmer feststellen lassen und wie sich ideale Lernkontexte bzw. optimale Rahmenbedingungen für solche Bildungsarrangements beschreiben lassen.

Ecarius, J./Klieme, E. /Stecher, L./Woods, J.(eds.) (2013): Extended Education - an International Perspective. Proceedings of the International Conference on Extracurricular and Out-of-School Time Educational Research. Leverkusen: Barbara Budrich.

**GEBF** TAGUNG  
2014

G02 b

Symposiums-  
beitrag

Mittwoch,

05.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 311

2

Das Nationale Bildungspanel als interdisziplinäre Forschungsinfrastruktur



## **Bildungserwerb von Kindern mit Zuwanderungshintergrund im Generationenverlauf**

Im Mittelpunkt des Beitrags stehen die Bildungserwerbsmuster von Zuwanderern und deren Entwicklung im Generationenverlauf. In bisherigen Studien zu in Deutschland lebenden Zuwanderergruppen konnte gezeigt werden, dass sich Unterschiede im Bildungserwerb, wie sie für die erste Zuwanderergeneration im Vergleich zur Mehrheitsbevölkerung beobachtet werden können, im Generationenverlauf für einige Herkunftsgruppen verringern und auflösen, während sie für andere Gruppen bestehen bleiben (z.B. Kristen & Dollmann, 2010; Kristen & Granato, 2007; Segeritz, Walter, & Stanat, 2010; Stanat, Rauch, & Segeritz, 2010). In der Mehrheit der Studien wird die 1. bzw. 1,5., 2. und/oder 2,5. Generation betrachtet, weshalb bislang offen ist, wie sich der Bildungserwerb im weiteren Generationenverlauf gestaltet.

Anhand der Daten der Startkohorte 3 des Nationalen Bildungspanels (NEPS; Schülerinnen und Schülern der fünften Jahrgangsstufe) wird im Beitrag zunächst der Frage nachgegangen, inwieweit Unterschiede in der Bildungsbeteiligung sowohl zwischen Kindern ohne und mit Zuwanderungshintergrund als auch zwischen Kindern einzelner Zuwanderergenerationen bestehen. Dabei ist es mit den Daten des NEPS möglich, die 3. Generation über die erfassten Geburtsländer der Großeltern zu identifizieren und folglich auch Aussagen zu den Bildungserwerbsmustern in den höheren Zuwanderergenerationen zu treffen. Zusätzlich wird die intergenerationale Entwicklung des Bildungserfolgs für spezifische Herkunftsgruppen betrachtet, und zwar für türkischstämmige Kinder und für Kinder, deren Familien aus der ehemaligen Sowjetunion stammen.

Schließlich werden die Determinanten untersucht, die im Generationenverlauf zu einer Angleichung von Bildungsmustern bzw. zum Fortbestand oder der Vergrößerung von Bildungsunterschieden führen. Ausgangspunkt bildet die analytische Trennung zwischen sogenannten primären und sekundären Effekten der sozialen Herkunft (Boudon, 1974), die vermehrt auch auf die ethnische Herkunft übertragen wird (z.B. Heath & Brinbaum, 2007; Kristen & Dollmann, 2010; Kristen et al., 2011). Primäre Effekte der sozialen bzw. ethnischen Herkunft verweisen dabei auf Bedingungen, die das schulische Lernen beeinflussen, während sekundäre Effekte der sozialen bzw. ethnischen Herkunft Bedingungen beschreiben, die den Bildungsentscheidungen zugrunde liegen. Die Trennung zwischen sozialen und ethnischen Herkunftseffekten kann auf verschiedene Erklärungsfaktoren angewendet werden (z.B. Kristen & Granato, 2007) – der Fokus im Beitrag liegt auf der Ressourcenausstattung der Kinder und Familien und den vorherrschenden Bildungsorientierungen.

Es wird mittels multivariater Analyseverfahren zunächst untersucht, inwieweit es sich bei den ungleichen Bildungsmustern tatsächlich um das Resultat von sekundären Effekten der ethnischen Herkunft handelt. In diesem Zusammenhang wird neben der erbrachten Leistung in der Grundschule auch die soziale Herkunft der Familien berücksichtigt. Im nächsten Schritt wird geprüft, inwieweit sich die verbleibenden Unterschiede über migrationsspezifische Bedingungen erklären lassen, und zwar der Kompetenz in der deutschen Sprache und den Bildungsaspirationen der Kinder.

Boudon, R. 1974. Education, opportunity, and social inequality. New York: Wiley.

Heath, A. F. & Y. Brinbaum. 2007. Guest Editorial: Explaining ethnic inequalities in educational attainment. *Ethnicities* 7: 291–304.

Kristen, C. & J. Dollmann. 2010. Sekundäre Effekte der ethnischen Herkunft: Kinder aus türkischen Familien am ersten Bildungsübergang. IN *Vom Kindergarten bis zur Hochschule*, Hrsg. B. Becker & D. Reimer, 117–144. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kristen, C., A. Edele, F. Kalter, I. Kogan, B. Schulz, P. Stanat & G. Will. 2011. The education of migrants and their children across the life course. IN *Education as a lifelong process: The German National Educational Panel Study (NEPS)*. Sonderheft 14 der Zeitschrift für Erziehungswissenschaft, Hrsg. H.-P. Blossfeld, H.-G. Roßbach & J. v. Maurice, 121–137. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kristen, C. & N. Granato. 2007. The educational attainment of the second generation in Germany. *Ethnicities* 7: 343–366.

Segeritz, M., O. Walter & P. Stanat. 2010. Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 62: 113–138.

Stanat, P., D. Rauch & M. Segeritz. 2010. Schülerinnen und Schüler mit Migrationshintergrund. IN *PISA 2009. Bilanz nach einem Jahrzehnt*, Hrsg. E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat, 200–230. Münster: Waxmann.



## Soziale Herkunft, „Health Literacy“ und Gesundheit

Viele empirische Studien zeigen, dass mit steigendem sozioökonomischem Status eine geringere Morbidität einhergeht (vgl. z.B. Mielck, 2008). Dies trifft nicht nur für Erwachsene, sondern auch für Kinder und Jugendliche zu, bei denen analoge Muster der Gesundheitschancen nach sozialer Herkunft gefunden werden (Lampert & Richter, 2009). Seit einiger Zeit beschäftigt sich die Forschung mit der Frage, über welche sozialen Mechanismen vertikale und horizontale Ungleichheiten in Gesundheitschancen transformiert werden. Die Spannbreite reicht von Belastungen aus der physischen und (psycho-)sozialen Umwelt über Ressourcen aus sozialer Unterstützung und sozialem Kapital und materiellen Ressourcen bis zu gesundheitlichem Risikoverhalten, gesundheitsbezogenen Lebensstilen und der Inanspruchnahme von gesundheitlicher Versorgung. Weitgehend ausgeklammert wurden bisher Studien, die die Bedeutung von Wissen und Kompetenzen für die Prägung von Gesundheitsverhalten oder die Möglichkeiten, gesundheitliche Belastungen zu bewältigen (vgl. Modell in Jungbauer-Gans & Gross, 2009), untersucht haben, was nicht zuletzt aufgrund der Datenlage kaum möglich war. Health Literacy bezeichnet die Fähigkeit von Personen, gesundheitsbezogene Informationen verstehen und verarbeiten zu können, um daraus angemessene Entscheidungen abzuleiten. In der Gesundheitsförderung bezeichnet der Begriff Health Literacy darüber hinaus die Gesamtheit des alltagspraktischen Wissens und der Fähigkeiten im Umgang mit Gesundheit und Krankheit, mit dem eigenen Körper ebenso wie mit den gesundheitsprägenden sozialen Lebensbedingungen (Abel et al., 2011). Empirische Studien konnten Effekte der Gesundheitskompetenz auf den Gesundheitszustand zeigen (Nielsen-Bohlman et al., 2004; Baker et al., 2002; zusammenfassend WHO, 2013). Für den vorliegenden Beitrag und die untersuchte Zielpopulation von 5.- und 6.-Klässlern nehmen wir an, dass Health Literacy im Wesentlichen einen Einfluss auf das Gesundheitsverhalten ausübt, das sich wiederum im Gesundheitszustand widerspiegelt.

Mit den NEPS-Daten liegen nun Daten vor, die es erlauben, die Bedeutung von gesundheitsrelevantem Wissen als Erklärung gesundheitlicher Ungleichheit zu untersuchen. Aus diesem Grund analysieren wir in diesem Beitrag die Bedeutung von Kompetenzen, d.h. insbesondere von Health Literacy als einer Teildimension der Science Kompetenz zur Erklärung sozial ungleicher subjektiver Gesundheit. Die theoretischen Modelle legen nahe, dass die soziale Ungleichheit von Gesundheit durch den Erwerb von gesundheitsrelevanten Kompetenzen abgeschwächt, wenn nicht sogar ausgeglichen werden kann. In diesem Beitrag untersuchen wir daher die Wirkung von gesundheitsrelevanten Kompetenzen als Mediator des Effekts der sozialen Herkunft auf den individuellen Gesundheitszustand.

Die Analysen werden mit der ersten und zweiten Welle der Startkohorte 3 des NEPS durchgeführt. Hier liegen unterschiedliche Kompetenzmaße zur Operationalisierung von Health Literacy vor. Zunächst stehen die selbstberichteten Schulnoten in Biologie bzw. Naturwissenschaften zur Verfügung, des Weiteren wurde in Welle 1 ein Maß für die allgemeinen kognitiven Fähigkeiten erhoben, Welle 2 enthält einen Score zur Scientific Literacy, der auch Gesundheitsitems enthält. Darüber hinaus wird geprüft, ob die Gesundheitsitems als Teildimension des Scientific Literacy Scores verwendet werden können. Als Maß für den Gesundheitszustand bildet das Item Subjektiver Gesundheitszustand (Self-Rated Health) die abhängige Variable. Zur Anwendung kommen querschnittsanalytische Regressionsverfahren.

Abel, Thomas; Sommerhalder, Kathrin; Bruhin, Eva (2011): Health Literacy/Gesundheitskompetenzen. In: Stephan Blümel, Peter Franzkowiak, Lotte Kaba-Schönstein, Guido Nöcker und Alf Trojan (Hg.): Leitbegriffe der Gesundheitsförderung und Prävention. Glossar zu Konzeptionen, Strategien und Methoden. Gamburg: Verlag für Gesundheitsförderung, S. 337–340.

Baker, David W.; Gazmararian, Julie A.; Williams, Mark V.; Scott, Tracy; Parker, Ruth M.; Green, Diane et al. (2002): Functional Health Literacy and the Risk of Hospital Admission Among Medi-care Managed Care Enrollees. In: American Journal of Public Health 92 (8), S. 1278–1283.

Jungbauer-Gans, Monika; Gross, Christiane (2009): Erklärungsansätze sozial differenzierter Gesundheitschancen. In: Matthias Richter und Klaus Hurrelmann (Hg.): Gesundheitliche Ungleichheit: VS Verlag für Sozialwissenschaften, S. 77-98.

Lampert, Thomas; Richter, Matthias (2009): Gesundheitliche Ungleichheit bei Kindern und Jugendlichen. In: Matthias Richter und Klaus Hurrelmann (Hg.): Gesundheitliche Ungleichheit: VS Verlag für Sozialwissenschaften, S. 209-230.

Mielck, A. (2008): Soziale Ungleichheit und Gesundheit in Deutschland. In: Bundesgesundheitsbl. 51 (3), S. 345-352. DOI: 10.1007/s00103-008-0465-1.

Nielsen-Bohlman, Lynn; Panzer, Allison M.; Kindig, David A. (2004): Health Literacy: A Prescription to End Confusion. Washington, DC: National Academies Press.

WHO (2013): Health Literacy. The Solid Facts: World Health Organization.



Chair(s): Gloria Jahn (TUM School of Education), Tina Seidel (TUM School of Education)

Diskutant/in: Miriam Steffensky (IPN)

Die Beobachtung und Reflexion von Unterricht stellen gemäß der Bildungsstandards der Kultusministerkonferenz ein zentrales Merkmal für die Innovation und Weiterentwicklung von Lehrkräften dar (KMK, 2004). Aus diesem Grund wird im Rahmen der Professionellen Unterrichtswahrnehmung untersucht, welche Aspekte von Unterricht Lehrkräfte identifizieren (Noticing) und wie sie diese interpretieren (Knowledge Based Reasoning) (van Es & Sherin, 2002). Darüber hinaus stellt sich die Frage, wie diese Fähigkeiten im Sinne des lebenslangen Lernens in der Aus- und Weiterbildung gefördert werden können.

Die Professionelle Unterrichtswahrnehmung beruht auf pädagogisch-psychologischem Wissen und dessen Anwendung auf konkrete Unterrichtssituationen (Sherin & van Es, 2009; Seidel & Stürmer, 2013). Als methodischer Zugang werden typischerweise Videos als Stimulus zusammen mit Fragen im offenen Antwortformat verwendet, um Noticing und Reasoning bei Lehrkräften zu aktivieren.

Im Symposium werden in vier Beiträgen verschiedene methodische Ansätze in der Forschung zur Professionellen Unterrichtswahrnehmung mit erziehungswissenschaftlichem und psychologischen Hintergrund vorgestellt, die sich zum einen auf unterschiedliche Aspekte des Wissens über Lehren und Lernen (z.B. Klassenführung oder Zielorientierung, Lernbegleitung und Lernatmosphäre) konzentrieren und zum anderen unterschiedliche Verfahren zur qualitativen Analyse der Antworten auf offene Fragen entwickelt haben. Alle Beiträge bauen auf einem vergleichbaren Verständnis von Professioneller Unterrichtswahrnehmung auf und zielen darauf ab, unterschiedliche Kompetenzniveaus in der Analyse von Videos abzubilden.

Ziel des Symposiums ist es aus einem methodischen Blickwinkel Gemeinsamkeiten und Unterschiede aufzuzeigen und zu diskutieren, welche Aspekte der Auswertungen von Antworten offener Fragen sich als tragfähig und welche als weiterentwicklungsbedürftig herausgestellt haben. Im ersten Beitrag wird eine Studie zur Erfassung des Zusammenspiels zwischen Noticing und Reasoning unter Berücksichtigung einer Expertensicht auf Unterricht vorgestellt. Im Fokus des zweiten Beitrags steht die Analyse schriftlicher Kommentare zu Unterrichtsvideos im Kontext einer Interventionsstudie zur Förderung der Professionellen Unterrichtswahrnehmung bei Lehramtsstudierenden. Im Rahmen des dritten Beitrags werden in Bezug auf die Evaluation eines Trainings zur Klassenführung Kodierungsmöglichkeiten offener Antwortformate diskutiert und mit standardisierten Messinstrumenten verglichen. Der vierte Beitrag berücksichtigt im Kontext fallbasierten Lernens zum Thema Klassenführung Antworten von Lehramtsstudierenden in Bezug auf die Beschreibung von beobachteten Unterrichtssequenzen und die Entwicklung von Handlungsalternativen. Abschließend werden alle Beiträge aus einer fachdidaktischen Perspektive diskutiert.

KMK 2004 = Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der KMK vom 16.12.2004). Bonn: KMK.

Seidel, T. & Stürmer, K. (2013). Modeling and measuring the structure of professional vision in pre-service teachers. Manuskript zur Publikation überarbeitet.

Sherin, M. G. & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60, 20-37.

van Es, E., & Sherin, M. G. (2002). Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571-596.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G168



## ***Eine Mixed-Methods-Studie zur Erfassung des Zusammenspiels von Noticing and Reasoning bei der Professionellen Unterrichtswahrnehmung von Lehramtsstudierenden***

Als zentrale Voraussetzung für Handlungskompetenzen von Lehrpersonen ist die Professionelle Unterrichtswahrnehmung Gegenstand zahlreicher Studien, nicht minder auch aufgrund ihrer bedeutsamen Implikationen für die Lehrerbildung. Professionelle Unterrichtswahrnehmung beschreibt, ob Lehrpersonen in der Lage sind ihr professionelles Wissen über Lehren und Lernen zu nutzen, wenn sie Ereignisse im Unterricht identifizieren (Noticing) und interpretieren (Reasoning) (van Es & Sherin, 2002). Nach dem aktuellen Forschungsstand bietet der Einsatz videografiertes Unterrichtsaufzeichnungen einen geeigneten Zugang, um die Prozesse des Noticings und Reasonings von Lehramtsstudierenden zu erfassen (Blomberg, Renkl, Sherin, Borko, & Seidel, 2013). Bis heute fehlen jedoch Studien, die Prozesse des Zusammenspiels zwischen den beiden Fähigkeiten in der konkreten Anwendung auf authentische Klassensituationen untersuchen. Methodisch wird kritisch diskutiert, dass bisher vor allem qualitative Unterschiede in der Tiefe der angewandten Interpretationen (z.B. Nutzung von konzeptuellem Wissen zur Beschreibung, Erklärung und Vorhersage von Unterrichtsprozessen) beschrieben werden. Es fehlen aber Informationen dazu, ob die Inhalte der Interpretationen auch aus Expertensicht geteilt werden. Beispielsweise kann bei einer Videosequenz zur Einführung in eine Unterrichtsstunde von einem Studierenden unter Benennung wissenschaftlicher Begriffe eine positive Zielklärung detailliert beschrieben werden. Dies würde bei den bisherigen Anwendungen der Verfahren zu einer „hohen“ Einstufung im Niveau der Interpretationen führen. Gleichzeitig kann es aber sein, dass gerade Experten bei der gleichen Sequenz klar der Meinung sind, dass dort eben keine Ziele geklärt wurden. In der Bewertung der „Nutzung“ professionellen Wissens würde dies zu deutlich unterschiedlichen Resultaten führen. Es wird deshalb argumentiert, dass der Inhalt der Interpretationen (Studierendensicht: Ziel geklärt; Experte: Ziel nicht geklärt) bei klar vorliegendem Expertenwissen mit zum Gegenstand der Forschung gemacht werden sollte.

Vor diesem Hintergrund haben wir in der explorativ angelegten Studie bei Lehramtsstudierenden (N = 109) anhand eines speziell ausgewählten Videoclips deren Fähigkeiten im Noticing und Reasoning untersucht und dabei den Inhalt der Interpretationen mit berücksichtigt. In Bezug auf den Videoclip bestanden umfangreiche Informationen zu allgemein didaktischen Aspekten des Lehrens und Lernens (Lehr-Lernkomponenten Zielorientierung und Bereitstellung einer positiven Lernatmosphäre), die Experten identifizierten und interpretierten. Als methodischer Ansatz wurden offene Fragen zum Videoclip und qualitative Inhaltsanalysen genutzt.

Die Ergebnisse zeigen, dass Lehramtsstudierende bereits eine Reihe relevanter Ereignisse identifizieren, die repräsentativ für die Lehr-Lern-Komponenten sind (Zielorientierung: 23% aller Analyseeinheiten pro Lehramtsstudierenden; Lernatmosphäre: 30% aller Analyseeinheiten pro Lehramtsstudierenden). Allerdings haben sie Schwierigkeiten, zu Erklärungen bzw. Vorhersagen zu kommen, die denen von Experten entsprechen (Zielorientierung: 63% keine Expertenübereinstimmung; Lernatmosphäre: 74% keine Expertenübereinstimmung).

Die Studie stellt damit Wissen zur Bedeutung des Zusammenspiels zwischen den beiden Fähigkeiten des Identifizierens und des Interpretierens bereit. Darüber hinaus macht sie deutlich, dass man nicht nur die Qualität der Interpretationen von Lehramtsstudierenden beachten sollte, sondern auch ob die Inhalte der Interpretationen mit denen von Experten übereinstimmen. Zukünftige Studien sollten den gewählten Ansatz und die Ergebnisse der explorativ angelegten Studie nutzen, um die Qualität des Zusammenspiels von Noticing und Reasoning noch näher zu analysieren.

## **Unterrichtsvideos als Instrument der Erfassung unterrichtsbezogener Analysekompetenz**

Theorie: Die unterrichtsbezogene Analysekompetenz gilt als bedeutsam für das erfolgreiche Unterrichten. Die Fähigkeit zur differenzierten Analyse von Unterrichtssituationen umfasst die Fähigkeit, relevante Situationen in Bezug auf die Unterstützung der Lernprozesse der Schülerinnen und Schüler zu erkennen (noticing) und die Fähigkeit, diese hinsichtlich ihrer Bedeutung wissensbasiert zu begründen (knowledge based reasoning), (Bromme, 2008; Davis, 2006; Sherin, Jacobs, & Philipp, 2011; van Es & Sherin, 2002). Voraussetzung dazu ist ein fundiertes Wissen über Bedingungen eines lernwirksamen Unterrichts und die Anwendung dieses Wissens auf konkrete Unterrichtssituationen (Berliner, 1991; Rosaen, Lundeberg, Cooper, Fritzen, & Terpsta, 2008; Sherin et al., 2011).

Unterrichtsvideos tragen den zentralen Charakteristika des Lehrberufs Rechnung und stellen ein kontextnahes Medium für den Aufbau der Analysekompetenz sowie die valide und reliable Messung der Analysekompetenz dar (Oser, Heinzer, & Salzmann, 2010; Seidel, Blomberg, & Stürmer, 2010).

Der Aufbau der Analysekompetenz kann durch Ansätze des fallbasierten Lernens gefördert werden, bei denen die Auseinandersetzung mit möglichst realen Anwendungssituationen (z.B. Unterrichtsvideos) im Zentrum steht (u.a. Shulman, 1992). Ermutigende Ergebnisse zum fallbasierten Lernen mit Videos stammen aus Weiterbildungsprojekten (z.B. Borko, Jacobs, Eiteljorg, & Pittman, 2008; Sherin & van Es, 2009), auch für die Ausbildung von Lehrpersonen gibt es erste Untersuchungen (z.B. Goeze, Hetfleisch, & Schrader, 2013; Gold, Förster, & Holodyski, 2013; Santagata & Guarino, 2011; Seidel, Blomberg, & Renkl, 2013; Star & Strickland, 2008; Stürmer, Könings, & Seidel, 2013).

Die valide und reliable Messung der Analysekompetenz von (angehenden) Lehrpersonen stellt ein anspruchsvolles Unterfangen dar (Kunter & Klusmann, 2010). Im Vergleich zur Erfassung der Wirkungen auf der Ebene von subjektiven Einschätzungen bleibt die objektive Erfassung von Wirkungen auf den Kompetenzzuwachs häufig unberücksichtigt. Aufgrund des methodischen Aufwands zur Erfassung der Entwicklung der Analysekompetenz liegen insgesamt erst wenige Erkenntnisse vor und wurden häufig nur kleine Stichproben untersucht, zudem existieren erst wenige standardisierte Instrumente, die Kompetenzen von Lehrpersonen jenseits von Fragebögen testen (Blomberg, Seidel, & Prenzel, 2011; Kunter & Klusmann, 2010). OBSERVER ist ein standardisiertes Instrument, welches längsschnittliche Untersuchungen mit grösseren Stichproben ermöglicht (Seidel et al., 2010). Im VideA-Projekt wird OBSERVER eingesetzt. In Ergänzung zu diesem quantitativ-standardisierten Zugang wird ein qualitativer Zugang verfolgt, was die gegenseitige Validierung dieser beiden Instrumente erlaubt. Im Symposiums-Beitrag wird der qualitative Zugang zur Analyse der Videokommentare thematisiert.

Fragestellung: Ziel des Projekts VideA (Videos in der Ausbildung von Lehrpersonen) ist es, Erkenntnisse zur Wirksamkeit des Einsatzes von Videos in der Ausbildung von Lehrpersonen in Bezug auf die unterrichtsbezogene Analysekompetenz zu generieren. Es wird geprüft, ob und unter welchen Bedingungen (eigene vs. fremde Videos) die angehenden Lehrpersonen durch die Analyse von Unterrichtsvideos ihre Fähigkeit verbessern, Merkmale lernwirksamen Unterrichts zu erkennen und zu begründen. Während der einsemestrigen Intervention arbeiteten ca. 150 Studierende des ersten Ausbildungsjahres der Pädagogischen Hochschule Luzern wöchentlich in Kleingruppen von ca. 8 Studierenden mit Videos. Sie wurden dabei von einer Dozentin begleitet.

In der VideA-Interventionsstudie wird unter anderem geprüft, inwiefern Studierende theoretische Kenntnisse zu den drei behandelten Unterrichtsqualitätsmerkmalen Zielorientierung, unterstützende Lernbegleitung und positive Lernatmosphäre sowie diesbezügliches Begründungswissen auf konkrete Situationen anwenden können, wenn sie Unterrichtsausschnitte schriftlich kommentieren und Optimierungsvorschläge generieren.

Methode: Die Studierenden kommentierten insgesamt drei kurze Videoausschnitte von Unterricht aus unterschiedlichen Schulstufen und -fächern. Für die inhaltsanalytische Auswertung dieser schriftlichen Antworten wurde ein daten- sowie theoriegestütztes Kategoriensystem entwickelt. Die Analyse der Kommentare zu den Videosequenzen wird von sechs geschulten Codierer/innen vorgenommen. Die schriftlichen Antworten der Studierenden werden im Hinblick darauf analysiert, ob die Studierenden relevante Situationen bezüglich der fokussierten Unterrichtsqualitätsmerkmale sowie Optimierungsmöglichkeiten erkennen und deren Bedeutung begründen können. Durch das Pre-Post-Design der Interventionsstudie können anschliessend Veränderungen der Analysefähigkeit der Studierenden beschrieben werden. Im Symposium werden Vorgehen und Kategoriensystem zur Analyse der schriftlichen Kommentare zu den Unterrichtssituationen vorgestellt und Fragen bezüglich des Vorgehens diskutiert.



## **Offene und standardisierte Messinstrumente der Unterrichtswahrnehmung in Trainings zur Klassenführung**

In der vorliegenden Studie geht es um die Frage, wie sich standardisierte und offene Messinstrumente für die Evaluation von Trainingsmaßnahmen zur Klassenführung eignen.

Effektive Klassenführung ist eine Voraussetzung für kognitive Aktivierung (Emmer & Stough, 2001) und somit ein wichtiger Bedingungsfaktor für erfolgreiche Schulleistungen (Seidel & Shavelson, 2007). Grundlage für ein erfolgreiches Handeln seitens der Lehrperson ist die professionelle Wahrnehmung, d.h. die Fähigkeit lernrelevante Ereignisse zu erkennen und angemessen interpretieren zu können (Sherin, 2007). Die professionelle Wahrnehmung hilft nicht nur den eigenen Unterricht besser zu verstehen, sondern wirkt sich auch positiv auf die Leistung der Schülerinnen und Schüler aus (Roth et al., 2011). Es stellt sich die Frage, wie man die unterrichtliche Wahrnehmung schon in der ersten Phase der Lehrerbildung erfolgreich schulen kann. Immer häufiger werden dazu videobasierte Lernsettings eingesetzt (Blomberg et al., 2013), in der Hoffnung, dass diese, im Gegensatz zu schriftlichen Fallbeispielen, die Simultanität und Komplexität von Unterricht dokumentieren können.

Dabei stellt sich die Frage, wie man mögliche Veränderungen in der professionellen Wahrnehmung der Teilnehmer messen kann. In den letzten Jahren wurden einige Ansätze entwickelt, die entweder mit einer quantitativen, standardisierten oder einer qualitativen Vorgehensweise versuchen, die professionelle Wahrnehmung zu erfassen (z.B. Seidel, Blomberg, & Stürmer, 2010; Van Es & Sherin 2002). Häufig dienen die Messungen der Erfassung von Veränderungen der professionellen Wahrnehmung nach videobasierten Förderungsprogrammen. Sowohl die Erfassungs- als auch die Trainingsmethoden weichen dabei leicht voneinander ab. Das Ziel des vorliegenden Beitrags ist zunächst die Untersuchung des Mehrwertes von videobasierten vs. nicht-videobasierten Trainings zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht. Des Weiteren wird der Vergleich zwischen der standardisierten und der offenen Erfassung solcher Lerneffekte thematisiert.

An der Trainingsstudie nahmen 51 Lehramtsstudierende im 4. Semester im Rahmen Ihres bildungswissenschaftlichen Studiums teil. 20 Studierende erhielten eine videobasierte Förderung (EG-Video), 13 Studierende nahmen an einem parallelen Seminar ohne Videoeinsatz (behandelte KG) und 28 Studierende an einem thematisch unterschiedlichen bildungswissenschaftlichen Seminar teil (unbehandelte KG).

In den Seminaren im Umfang von ca. 60h analysierten die Studierenden der Trainingsseminare (mit und ohne Video) die Klassenführung in fremden und eigenen Unterrichtssequenzen. Die EG-Video erhielt eine videobasierte Intervention, in der die Teilnehmer\_innen zunächst in die Theorie der Klassenführung eingeführt wurden und diese dann in kurzen Videobeispielen (Clips 3-5min) sowie einer ganzen Stunde aus dem Unterricht fremder Lehrpersonen analysierten. Im zweiten Block analysieren die Studierenden die Klassenführung in ihrem eigenen Unterricht, der zuvor in Vierergruppen durchgeführt und videographiert wurde.

Die behandelte Kontrollgruppe besuchte eine parallele Intervention ohne den Einsatz von Videoelementen. Im ersten Block wurde die Klassenführung fremder Lehrpersonen anhand von Fallbeispielen diskutiert. Im zweiten Block analysieren die Studierenden die Klassenführung in ihrem eigenen Unterricht, der zuvor in Vierergruppen durchgeführt und protokolliert wurde.

Zur Überprüfung der Veränderung der professionellen Wahrnehmung wurden zwei unterschiedliche Ansätze herangezogen: ein standardisiertes videobasiertes Messinstrument, das die professionelle Wahrnehmung von Klassenführung im Grundschulunterricht anhand von kurzen Videoclips und geschlossenen Ratingitems erfasst (Gold & Holodynski, subm.) und eine qualitative Messung mit einem offenen Video, über das die Studierenden eine schriftliche Analyse schreiben sollten. Es wurde ein Kategoriensystem nach den Dimensionen Allgegenwärtigkeit, Strukturierung und Regeln / Routinen entwickelt. Wobei sich gute Interrater-Reliabilitäten ( $Kappa = .69$ ) ergaben.

Die standardisierte Messung wurde zu Beginn und am Ende des Seminars eingesetzt, wohingegen die offene Analyse zusätzlich vor dem eigenen Unterricht der Trainingsteilnehmer/innen (also dreimal) eingesetzt wurde.

Eine Kovarianzanalyse mit der standardisierten Messung der professionellen Wahrnehmung zeigte signifikante Unterschiede zwischen den drei Gruppen (partielles Eta-Quadrat = .16). Es ergab sich jedoch kein signifikanter Unterschied zwischen den beiden Trainingsgruppen. Die näheren Ergebnisse der qualitativen Auswertung werden im Vortrag berichtet. Dabei wird insbesondere die Entwicklung der Kodierungsmöglichkeiten solcher offener Analysen thematisiert und die Vor- und Nachteile gegenüber einer standardisierten Messung diskutiert.

## **Fallbasiertes Lernen in der Lehrerbildung.**

### **Erhebung der Analysekompetenz anhand eines offenen Itemformates**

Das fallbasierte Lernen anhand videographierter Unterrichtsstunden wird in den letzten Jahren zunehmend in der Lehrerbildung umgesetzt und empirisch untersucht. Aktuelle Konzepte beschreiben dabei die professionelle Wahrnehmung bzw. Analyse von Lehr- und Lernsituation als zentrales Lernziel videofallbasierter Lehrerbildung (Krull et al., 2007) und definieren diese Kompetenz als eine Facette der unterrichtlichen Reflexionskompetenz (van Es & Sherin, 2008). Der Blick in die forschungsbasierte Literatur zeigt allerdings auch, dass sich die Definition der professionellen Wahrnehmungs- und Analysekompetenz zwischen den Projekten deutlich unterscheidet (Lindmeier, 2013). Infolgedessen differieren die Arten der Datenerhebung, als auch die Verfahren der Datenanalyse.

Für unsere Studie definierten wir die professionelle Unterrichtswahrnehmung auf Basis von Literatur zur Reflexion (u.a. Dewey, 1910; Korthagen, 1985; Schön, 1983; Zeichner, 1981) anhand zweier Hauptkomponenten: Selektion und Analyse. Selektion beschreibt die Fähigkeit, relevante Situationen bezüglich bestimmter Thematiken (z.B. Klassenführung) zu erkennen. Die Analyse umfasst mehrere Teilfähigkeiten, wie die Beschreibung der Ausgangssituation, die Benennung von Optionen und Alternativen für das Lehrerhandeln, das Erkennen bzw. Antizipieren von Ursachen und Konsequenzen sowie das Entscheiden für bzw. wider eine Option. Als zentraler Aspekt der Analyse kann dabei die Abwägung unterschiedlicher Alternativen gesehen werden.

Zur Erhebung der professionellen Unterrichtswahrnehmung in einer Studie zum fallbasierten Lernen mit dem Thema Klassenführung, erstellten wir ein web- und videobasiertes Instrument. Die N=647 Studierenden wurden gebeten, videographierte Unterrichtssequenzen in einem offenen Itemformat schriftlich bezüglich Klassenführung zu analysieren. In einem Prä- und Posttest kommentierten die Studierenden drei Unterrichtssequenzen. Um einen Erinnerungseffekt zu vermeiden, wurden inhaltsähnliche Sequenzen bestimmt und auf Basis von Expertenurteilen entweder dem Prä- oder Posttest zugeordnet. Die gespeicherten Kommentare konnten anschließend anhand der definierten Teilfähigkeiten der Analyse ausgewertet werden. Eine Kodierung entlang der Teilfähigkeiten gibt die Struktur der gegebenen Analysen wider und lässt sie ordinal ordnen.

Die Ergebnisse werden entlang der wissenschaftlichen Gütekriterien Validität, Reliabilität und Objektivität diskutiert.

Inhaltsvalidität: Die Definition des zu erhebenden Konstrukts, betont die bewusste Abwägung unterschiedlicher Alternativen. Im Erhebungsinstrument nutzten wir ein offenes Format, in dem Studierende ihre Analyse schriftlich darlegen, um diese Denkprozesse sichtbar zu machen.

Kriteriumsvalidität: Da die Analyse ein bestimmtes Thema fokussiert (in unserem Fall Klassenführung), lässt sich vermuten, dass die professionelle Unterrichtswahrnehmung mit dem deklarativen Wissen zum jeweiligen Thema korreliert. Erste Berechnungen (auf Basis von 20% der ausgewerteten Daten) hierfür ergeben eine geringe aber signifikante Korrelation.

Konstruktvalidität: Typischerweise werden hierfür Experten-Novizen Vergleiche angestellt, die allerdings in diesem Projekt erst noch erhoben werden.

Reliabilität: Die jeweiligen Items unterschieden sich im Stimulus zur Analyse – dem Unterrichtsvideo. Aus diesem Grunde haben wir die unterschiedlichen Sequenzen auf mehreren Dimensionen (z.B. Komplexität, Interessantheit) auf Vergleichbarkeit untersucht. Andere Faktoren konnten konstant gehalten werden, da die Sequenzen aus derselben Unterrichtsstunde stammten. Die Reliabilitäten des Pre- bzw. Posttests betragen Cronbach's  $\alpha = .68 - .83$ .

Durchführungsobjektivität: Web-basierte Erhebungsinstrumente, wie sie in diesem Forschungsfeld eingesetzt werden, ermöglichen den Einfluss des durchführenden Testleiters zu minimieren.

Auswertungsobjektivität: Für die Auswertung wurde jeweils kodiert, ob eine Teilfähigkeit der Analyse realisiert wurde oder nicht. Hierbei konnten zufriedenstellende Werte in der Beurteilerübereinstimmung erzielt werden ( $K = .64 - .75$ ).

Interpretationsobjektivität: Anhand der kodierten Teilfähigkeiten konnte die Struktur der Analyse jedes Kommentars dargestellt werden. Zur Interpretation und Bewertung dieser Struktur wurden drei unterschiedliche Analyselevel definiert (Dewey, 1910): Kommentare, die (a) über eine beschreibende Auseinandersetzung nicht hinauskommen, (b) sich um eine Sinnggebung bemühen, (c) denen eine Abwägung unterschiedlicher Alternativen stattfindet.

Vorteile von offenen Erhebungsformaten für die professionelle Unterrichtswahrnehmung liegen in ihrer kompetenznahen Messung. Gegenüber geschlossenen Formaten fällt hierbei jedoch zusätzlicher Arbeitsaufwand bei der Auswertung an (Nebengütekriterium Ökologie). Die bisher ermutigenden Werte bei der Objektivität und Reliabilität weisen darauf hin, dass offene Erhebungsformate durchaus ernst genommen werden sollten. Ob die erwarteten Ergebnisse der Validierung anhand der Daten gestützt werden können, werden die weiteren Auswertungen zeigen.

Sabine Sommer (Humboldt-Universität zu Berlin)

## ***Auf dem Weg zu einer inklusiven Schule: Aktuelle Perspektiven auf verschiedene Akteure der Schul- und Unterrichtsentwicklung***

Chair(s): Sabine Sommer (Humboldt-Universität zu Berlin), Stefanie Czempiel (Friedrich-Schiller-Universität Jena)

Diskutant/in: Simone Börner (Friedrich-Schiller-Universität)

Spätestens seit der Ratifizierung der UN-Behindertenrechtskonvention durch Deutschland im Jahr 2008 lassen sich in der Praxis deutscher Schulen sowie in der Forschung zahlreiche Aktivitäten und Entwicklungen im Bereich der Inklusion beobachten. Schulen begeben sich freiwillig oder angestoßen durch Gesetzesänderungen auf den Weg, Schüler/innen mit und ohne sonderpädagogischem Förderbedarf gemeinsames Lernen zu ermöglichen. Seit den 1970er Jahren dokumentierten Begleitforschungen die ersten Schulversuche mit Integrationsklassen und arbeiteten Gelingensbedingungen für Gemeinsamen Unterricht heraus. Nach punktuellen Maßnahmen über vier Jahrzehnte stellt sich aktuell die Frage, wie jede Schule in Deutschland eine inklusive Schulkultur entwickeln kann, also wie Inklusion in der Breite etabliert werden soll. Dazu muss die Forschung die einzelnen Akteure im System Schule in den Blick nehmen.

Im Mittelpunkt des Symposiums steht die Frage, wie die schulischen Akteure Lehrer/innen, Eltern und Schüler/innen den Prozess der Umsetzung der Behindertenrechtskonvention wahrnehmen. Welche Modelle der bedarfsgerechten Unterstützung von Fachkräften sind sinnvoll und wirksam? Erkenntnisse vier aktueller empirischer Studien aus den Blickwinkeln verschiedener Fachrichtungen sollen dabei helfen, diese Frage zu beantworten. Der erste Beitrag untersucht die Einstellung von Lehrkräften und Eltern zu Inklusion. Bisher in der Forschung zu Inklusion kaum beachtet und aus diesem Grund von besonderem Interesse sind dabei die Einflussfaktoren, die eine höhere Bereitschaft zur Inklusion hervorrufen. Der zweite Beitrag fokussiert die Veränderung von Lehrerhaltungen, basierend auf concern-theoretischen Ansätzen, innerhalb eines innovativen Schulentwicklungsprozesses, wie der der Umsetzung schulischer Inklusion. Dieser Beitrag basiert auf den Ergebnissen einer Begleitstudie des Pilotprojektes „Inklusive Grundschule“ des Landes Brandenburg, innerhalb derer 900 Lehrer/innen befragt wurden. Der dritte Beitrag untersucht den Umgang von Lehrer/innen mit Heterogenität im Unterricht. Im Zentrum der Untersuchung steht die Frage, wie multiprofessionelle Teamarbeit im Gemeinsamen Unterricht gestaltet wird. Es zeigt sich, dass gelingende Kooperationsprozesse zwischen Pädagog/innen unterschiedlicher Qualifikationen von entscheidender Bedeutung für eine hohe Unterrichtsqualität sind. Der vierte Beitrag thematisiert den Response-to-Intervention-Ansatz aus den USA, der im Rahmen des Rügener Inklusionsmodells erstmals in der Praxis in Deutschland umgesetzt wurde. In einer Vergleichsstudie wurde dessen Wirksamkeit durch die Beobachtung der Entwicklung der Schulleistungen von Schüler/innen mit und ohne sonderpädagogischem Förderbedarf untersucht. Von den Einstellungen und Bedürfnissen der verschiedenen Akteure hin zu einem handlungsorientierten Konzept werden im Symposium zusammenfassend aktuelle Erkenntnisse sowohl quantitativer als auch qualitativer Studien aus der Perspektive der Erziehungswissenschaft, der Psychologie und der Sonderpädagogik zum Thema der schulischen Inklusion dargestellt.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G191

## **Auf dem Weg zur inklusiven Schule?**

### **Einflussfaktoren auf die Einstellungen von Lehrkräften und Eltern zu Inklusion**

In der Studie „Einstellungsforschung zu Inklusion“ (EFI) der Pädagogischen Hochschule Heidelberg wurden Lehrkräfte und Eltern an verschiedenen Schularten in öffentlicher Trägerschaft zu ihrer Einstellung zu Inklusion befragt. Als forschungsleitend wird innerhalb des Projekts die These angesehen, dass die Umsetzung von Inklusion in der Praxis maßgeblich von den Einstellungen der beteiligten Personen abhängt.

Direkte Vergleiche der Einstellungen zu Inklusion von Lehrkräften und Eltern fehlen bislang. Die meisten Studien im Themenschwerpunkt „Einstellungen zu Inklusion“ beschäftigen sich mit den Einstellungen von Lehrkräften oder bestimmten Gruppen von Eltern (z.B. Eltern von Kindern mit sonderpädagogischem Förderbedarf). Ferner wird in der einschlägigen Literatur zwar darauf verwiesen, dass eine positive Einstellung zu Inklusion nicht unbedingt mit der entsprechenden Bereitschaft einhergeht, eine inklusive Klasse zu übernehmen (vgl. z.B. Eberl, 2000; Bless, 2007; de Boer et al., 2011). Wovon diese Bereitschaft aber abhängt, bleibt unklar. Die vorliegende Studie möchte zur Klärung dieser offenen Fragen beitragen. Im Vortrag wird folgenden Fragen und Hypothesen nachgegangen:

Unterscheiden sich die Einstellungen von Lehrkräften und Eltern? Vor dem Hintergrund der bisherigen Forschungsbefunde wird vermutet, dass Eltern positiver eingestellt sind als Lehrkräfte.

Welche Einflussfaktoren tragen zu einer höheren Bereitschaft zu Inklusion bei Lehrkräften und Eltern bei? Hier wird davon ausgegangen, dass Faktoren, die mit einer positiven Einstellung zu Inklusion einhergehen, wie z.B. Kontakt zu Menschen mit Behinderung, die berufsbezogene Kompetenzerwartung (Lehrer-Selbstwirksamkeit) sowie das Wissen über Inklusion und ihre Umsetzungsmöglichkeiten und eher niedrige Werte im individuellen Belastungsempfinden, auch die verhaltensnahe Bereitschaft zu Inklusion fördern. Zur Beschreibung des soziodemographischen Hintergrunds der Teilnehmenden wurden Alter, Geschlecht, Stadt (Mannheim vs. Heidelberg) und bei den Eltern der Bildungsstand erhoben. Als Kontrollvariable wurde die Tendenz zu sozial erwünschtem Antwortverhalten erfasst. Dabei handelt es sich um eine Kurzskaala auf Basis des BIDR (Paulhus, 1991).

Zur Erfassung der kognitiven Komponente von Einstellungen zu Inklusion wurde eine modifizierte Version der Skala Einstellungen zur Integration – EZI (Kunz et al., 2010) eingesetzt. Die EZI-Skala besteht aus elf Items mit einer 6-stufigen Likert-Skalierung. Neben dem Einstellungs-Gesamtwert erfasst die EZI-Skala mit vier der elf Items die Komponente soziale Inklusion und mit den verbleibenden sieben Items die Komponente schulische Förderung und Unterstützung.

Die verhaltensnahe Einstellungskomponente wurde als Bereitschaft zu Inklusion (Übernahme einer inklusiven Klasse bei den Lehrkräften bzw. Einschulung des eigenen Kindes in eine inklusive Klasse bei den Eltern) mit dichotomen Ein-Item-Messungen (ebenfalls 6-stufig likert-skaliert) erfasst.

Als Vorhersagevariablen zur Erklärung von Unterschieden in der Bereitschaft zu Inklusion wurden zum einen das Vorhandensein beruflicher und/oder privater Erfahrungen mit Menschen mit Behinderung, das Vorhandensein von Erfahrungen mit Inklusion (Tätigkeit bzw. eigenes Kind in einer inklusiven Klasse) sowie bei den Lehrkräften die Frage, ob Fortbildungen zum Thema Inklusion besucht worden waren, durch dichotome Ein-Item-Messungen erfasst. Zum anderen wurde die Lehrer-Selbstwirksamkeitserwartung mittels der Lehrer-Selbstwirksamkeits-Skala von Schwarzer und Schmitz (1999) erhoben (10 Items, 4-stufige Likert-Skalierung).

Insgesamt nahmen 1548 Personen, davon 652 Lehrkräfte und 896 Eltern, an 71 Schulen in den Stadtbezirken Mannheim und Heidelberg teil. Die Ergebnisse zeigen, dass die Einstellung und Bereitschaft zur Inklusion u.a. vom Vorhandensein beruflicher und/oder privater Erfahrungen mit Menschen mit Behinderungen, von der Teilnahme an themenspezifischen Fortbildungen sowie von der Lehrerselbstwirksamkeit abhängen. SonderpädagogInnen haben positivere Einstellungen als Lehrkräfte an anderen Schularten, während Eltern insgesamt eine positivere Einstellung und eine höhere Bereitschaft zu Inklusion zeigen als Lehrkräfte. Des Weiteren lässt sich feststellen, dass Lehrkräfte den Aspekt des sozialen Miteinanders von Inklusion signifikant positiver bewerten als Eltern, während diese die Auswirkungen auf die schulische Förderung durch Inklusion signifikant positiver einschätzen als die Lehrkräfte. Diese Ergebnisse zeigen wichtige Implikationen für eine Schulentwicklung in Richtung Inklusion. (Weiter-)Qualifizierungsangebote für Lehrkräfte sollten demnach beispielsweise Hospitationen in inklusiven Settings, Fortbildungen zu Differenzierungsmöglichkeiten im Unterricht und Maßnahmen zur Förderung der Selbstwirksamkeit beinhalten.

Bless, G. (2007): Zur Wirksamkeit der Integration. Forschungsüberblick, praktische Umsetzung einer integrativen Schulform, Untersuchungen zum Lernfortschritt (3. Aufl.). Bern: Haupt.

De Boer, A. A., Pijl, S. J. & Minnaert, A. (2011): Regular primary schoolteachers' attitudes towards inclusive education: a review of the literature. In: International Journal of Inclusive Education, 15 (3), 331-353.

Eberl, D. (2000): Gemeinsamer Unterricht von behinderten und nicht behinderten Schülern in der Beurteilung von Schulleitern und Lehrern. Eine Untersuchung an Grund- und Sonderschulen in Nordrhein-Westfalen. Psychologie und Medizin. Bd. 6. Witterschlick: Wehle..

Kunz, A., Luder, R. & Moretti, M. (2010): Die Messung von Einstellungen zur Integration (EZI). In: Empirische Sonderpädagogik, 2 (3), 83-94.

Paulhus, D.L. (1991): Measurement and control of response bias. In: Measures of personality and social psychological attitudes (Hrsgs: Robinson, J.P.; Shaver, P.R. und Wrightman, L.S.), 17–59. New York: Academic Press.

Schwarzer, R. & Schmitz, G. S. (1999): Kollektive Selbstwirksamkeitserwartung von Lehrern. Eine Längsschnittstudie in zehn Bundesländern. In: Zeitschrift für Sozialpsychologie, 30 (4), 262- 274.

Winkler, N., Kroh, M. & Spiess, M. (2006): Entwicklung einer deutschen Kurzskaala zur zweidimensionalen Messung von sozialer Erwünschtheit. Deutsches Institut für Wirtschaftsforschung: Berlin.

Christian Jäntsch (Universität Potsdam), Thorsten Luka (Universität Potsdam), Stefanie Bosse (Universität Potsdam), Jana Vogel (Universität Potsdam), Nadine Spörer (Universität Potsdam)

## Entwicklungen von Lehrer-Concerns im Kontext der inklusiven Grundschule

Inklusives Lernen ist derzeit eine der zentralen Herausforderungen für Schulentwicklung. Den Lehrkräften in den Schulen obliegt dabei die praktische Umsetzung des Inklusionsgedankens im Unterricht. Um die Akzeptanz von Innovationen durch Lehrkräfte zu erfassen, fokussieren sogenannte concerns-theoretische Ansätze auf die Beschreibung kognitiver und affektiver Reaktionen auf Implementationsvorgaben (Pant et al., 2008, S. 248). Der Concerns-Begriff umfasst dabei die Gesamtheit an Gefühlen, Vorannahmen, Gedanken und Überlegung gegenüber einzelnen Sachverhalten einer Innovation (Hall & Hord, 2011, S. 72). Aufbauend auf Fullers Arbeiten zu Concerns of Teachers (Fuller 1969) formulierten Hall und Hord (2011) das Concerns Based Adoption Model aus, das einen konzeptionellen und analytischen Rahmen zur Steuerung schulischer Implementationsvorhaben bietet. Mithilfe des darin enthaltenen diagnostischen Instruments der Stages of Concern (SoC) (Hall et al., 1977) kann die Perzeption von Innovationen durch Lehrkräfte abgebildet werden. Das SoC-Modell beinhaltet sieben distinkte Stufen: 0) Bewusstsein, 1) Informationsbeschaffung, 2) persönliche Betroffenheit, 3) Aufgabenmanagement, 4) Auswirkung auf Lernende, 5) Kooperationsmöglichkeiten sowie 6) Revision und Optimierung. Der idealtypische Entwicklungsverlauf sieht für einen gelingenden Implementationsprozess eine wellenförmige Verschiebung der kognitiven Beschäftigung von der ersten bis zur siebenten Stufe vor. Aufbauend auf dem SoC-Ansatz wird in diesem Beitrag der Frage nachgegangen, inwieweit Veränderungen aufseiten der Lehrerinnen und Lehrer in der Auseinandersetzung mit dem inklusiven Unterrichten im Laufe eines Jahres sichtbar werden. In diesem Zusammenhang sollen drei Fragestellungen beantwortet werden: 1) Wie zuverlässig und valide ist das SoC-Instrument im Kontext der inklusiven Grundschule? 2) Wie stabil erweist sich das Modell im Längsschnitt, d.h. nach einem Jahr? 3) Welche Lehrerprofile lassen sich mittels Clusteranalyse extrahieren und welche Entwicklung kann im Längsschnitt nachvollzogen werden? Grundlage dieser Studie bilden Befragungen von rund 900 Lehrerinnen und Lehrern im Rahmen des Pilotprojektes „Inklusive Grundschule“ des Landes Brandenburg, das seit dem Schuljahr 2012/13 von der Universität Potsdam wissenschaftlich begleitet wird. Auf Basis der Daten des ersten Messzeitpunktes erfolgte zunächst die Überprüfung der Güte des SoC-Fragebogens. Die Reliabilität der Subskalen erwies sich als zufriedenstellend bis gut ( $\alpha=.70$  bis  $\alpha=.85$ ). Übereinstimmend zu den Befunden von Cheung et al. (2001) zeigte sich, dass ein 5-Stufen-Modell ( $\chi^2=487$ ,  $df=196$ ,  $CFI=.910$ ,  $RSMEA=.064$ ,  $AIC=18407$ ) gegenüber der Annahme von sieben Stufen ( $\chi^2=1621$ ,  $df=539$ ,  $CFI=.822$ ,  $RSMEA=.070$ ,  $AIC=27905$ ) besser zu den vorliegenden Daten passt. Auf Grundlage des 5-Stufen-Modells konnten clusteranalytisch in einem nächsten Schritt drei Lehrerprofile extrahiert werden: 1) Passiv-rezeptive Lehrerinnen und Lehrer zeigen Bewusstsein für das inklusive Unterrichten. Sie wünschen sich in Hinblick auf ihre persönliche Betroffenheit in höherem Maße weitere Informationen dazu und machen sich Gedanken, wie sie inklusiven Unterricht umsetzen sollen. Nichtsdestotrotz sind sie nicht abgeneigt, inklusiven Unterricht in Kooperation mit dem Kollegium umzusetzen. 2) Innovatoren zeigen ein hohes Bewusstsein für das Thema Inklusion. Sie zeigen nur wenig Sorge, wie sie die Anforderungen erfüllen sollen und sind dementsprechend hoch motiviert, für die Umsetzung des inklusiven Unterrichts mit den Kollegen zu kooperieren. 3) Verunsicherte Skeptiker weisen ein unterdurchschnittliches Bewusstsein auf, sich mit dem inklusiven Unterrichten zu befassen. Hingegen haben sie große Sorge, wie sie inklusiven Unterricht realisieren können. Dementsprechend zeigen sie eine eher geringe Bereitschaft, in Kooperation mit den Kollegen inklusiv zu unterrichten. Die Daten des zweiten Messzeitpunktes, zu dem das SoC-Instrument eingesetzt wird, werden Anfang des Jahres 2014 vorliegen. Im Längsschnitt soll überprüft werden, wie die Entwicklung entlang der Stufen verläuft. Gemäß der Annahme eines wellenförmigen Entwicklungsverlaufs ist zu erwarten, dass die Anzahl der verunsicherten Skeptiker abnimmt und die Anzahl der Innovatoren zunimmt. Dabei ist von besonderem Interesse, in welchem Ausmaß Wanderungen zwischen den Lehrertypen zu finden sind.

Cheung, D., Hattie, J. & Ng, D. (2001). Reexamining the Stages of Concern Questionnaire: A Test of Alternative Models. *The Journal of Educational Research*, 94, 226-236.

Hall, G. E., George, A. & Rutherford, W. L. (1977). *Measuring Stages of Concern about the Innovation: A Manual for the Use of the SoC Questionnaire*. Austin: University of Texas.

Hall, G. E. & Hord, S. M. (2011). *Implementing Change. Patterns, Principles, Potholes*. Boston: Pearson.

Pant, H. A., Vock, M., Pohlmann, C. & Köller, O. (2008). Eine modellbasierte Erfassung der Auseinandersetzung von Lehrkräften mit den länderübergreifenden Bildungsstandards. In E.-M. Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 245-260). Münster: Waxmann.



Stefanie Czempiel (Friedrich-Schiller-Universität Jena), Sabine Sommer (Humboldt-Universität zu Berlin), Bärbel Kracke (Friedrich-Schiller-Universität Jena), Ada Sasse (Humboldt-Universität zu Berlin)

## ***Schulinterne, multiprofessionelle Kooperation als Voraussetzung für erfolgreichen Gemeinsamen Unterricht: Ergebnisse einer Untersuchung an allen Schulen in einer Kommune***

Bereits seit den 1970er Jahren, nachdem in der Bundesrepublik Deutschland die ersten Schulversuche mit Integrationsklassen starteten, arbeiten wissenschaftliche Begleitforschungen Gelingensbedingungen für Gemeinsamen Unterricht heraus (z.B. Wocken, Antor, & Hinz, 1988; Preuss-Lausitz, 2001; Prengel, 1990; Schöler, 2009; Heyer, 2009; Gold, 2011). Schulinterne Kooperation in multiprofessionellen Teams ist dabei eine zentrale und vielfach erforschte Voraussetzung für eine hohe Qualität des Gemeinsamen Unterrichts (Eberwein & Knauer, 2009). Insbesondere im Gemeinsamen Unterricht ist das Zusammenführen der Arbeitsfelder und Kompetenzen unterschiedlicher pädagogischer Professionen erforderlich. Intensive Kooperation zwischen Lehrkräften ist nach Gräsel et al. (2006) in Deutschland allerdings nur wenig verbreitet. Gräsel et al. (2006) systematisiert Ausprägungen der Kooperation nach ihrer Komplexität: (1) Austausch, (2) gemeinsame Arbeitsplanung, (3) ko-konstruktive Kooperation. Im Rahmen der Expertise „Gemeinsamer Unterricht im Kontext von Schul- und Unterrichtskultur in der Stadt Jena“ wurde die schulinterne Kooperation aus der Sicht der an schulischer Inklusion maßgeblich beteiligten Akteure, Schulleiter/innen, Lehrer/innen, Sonderpädagog/innen und Schulbegleiter/innen untersucht. Uns interessiert die Frage, wie die Kooperationen zwischen unterschiedlichen Professionen ausgeprägt sind. Welche Erfahrungen mit multiprofessioneller Kooperation berichten die einzelnen Akteure? Welche Bedeutung wird der Kooperation durch die einzelnen Akteure beigemessen? Welche Strukturen und Strategien zur gemeinsamen Arbeit werden genutzt?

Basierend auf dem ökosystemischen Ansatz wurden in der Jenaer Studie qualitative und quantitative Daten an 11 Grundschulen, 7 Gemeinschaftsschulen, 4 Gesamtschulen und 4 Gymnasien erhoben. In problemzentrierten Interviews wurden 26 Schulleiter/innen, 41 Lehrer/innen, 22 Sonderpädagog/innen und 18 Schulbegleiter/innen von 26 Schulen zu Kooperationsprozessen in der Unterrichtsvorbereitung, -ausgestaltung und -reflexion befragt. Hinzukommen quantitative Daten eines Lehrerfragebogens, der von 140 Lehrer/innen und 22 Sonderpädagog/innen bearbeitet wurde. Um das Bild zu den professionsübergreifenden Kooperationen zu vervollständigen, werden die Daten aus 280 Unterrichtsbeobachtungssequenzen á 20 Minuten an allen 26 Schulen hinzugezogen.

In der Untersuchung wird sichtbar, dass Kooperation, obwohl sie für den Gemeinsamen Unterricht eine wichtige Voraussetzung darstellt, noch sehr gering ausgeprägt ist und sich damit nach Gräsel auf der Ebene des Austausches befindet. Weiterhin kann beobachtet werden, dass Lehrer/innen von Schulen mit einer hohen Qualität des Gemeinsamen Unterrichts verstärkt gemeinsam planen oder bereits die Kooperationsform der Ko-Konstruktion umsetzen. Hier zeigt sich, dass Kooperation als ein wichtiges Element inklusiver Schulkultur im Schulkonzept verankert sein muss. Die Ergebnisse dieser Untersuchung geben Aufschluss über Bedürfnisse der Akteure bezüglich der Weiterentwicklung und Unterstützung von Schulen in der Umsetzung von Inklusion und zur Gestaltung von Aus- und Fortbildung.

Eberwein, H. & Knauer, S. (2009). Rückwirkungen integrativen Unterrichts auf Teamarbeit und Lehrerrolle. In H. Eberwein & Knauer, S. (Hrsg.). Handbuch Integrationspädagogik. Kinder mit und ohne Beeinträchtigung lernen gemeinsam. Weinheim: Beltz, 422-432.

Gold, A. (2011). Lernschwierigkeiten. Ursachen, Diagnostik, Intervention. Stuttgart: Kohlhammer.

Gräsel, C. & Fussangel, K. & Pröbstel, C. (2006). Lehrkräfte zur Kooperation anregen – eine Aufgabe für Sisyphos. Zeitschrift für Pädagogik 52/2 (2006). S. 205-219.

Heyer, P. (2009). Grundschule – Schule für alle Kinder. Grundsätze zur Entwicklung integrativer Arbeit. In Eberwein, H. & Knauer, S. (Hrsg.). Handbuch Integrationspädagogik. Kinder mit und ohne Beeinträchtigung lernen zusammen. Weinheim, Beltz, 191–200.

Prengel, A. (1990). Subjektive Erfahrungen mit Integration (Untersuchung mit qualitativen Interviews). In H. Deppe-Wolfinger, A. Prengel & H. Reiser. Integrative Pädagogik in der Grundschule. Weinheim und München: Juventa, 147-258.

Preuss-Lausitz, U. (2001). Qualitätsmerkmale, Leistungsmessung und Evaluation der pädagogischen Arbeit im Gemeinsamen Unterricht und in der Sonderschule. Zeitschrift für Heilpädagogik 2/2001, 46 -50.

Schöler, J. (2009). Alle sind verschieden. Auf dem Weg zur Inklusion in der Schule. Weinheim: Beltz.

Wocken, H., Antor, G. & Hinz, A. (1988). Integrationsklassen in Hamburger Grundschulen. Hamburg.

Stefan Voß (Universität Rostock), Yvonne Blumenthal (Universität Rostock)

## ***Das Rügener Inklusionsmodell (RIM) – Konzeption und Evaluation einer präventiven und inklusiven Grundschule nach dem Response to Intervention-Ansatz (RTI)***

Der in den USA in einer nunmehr 30 Jahre andauernden Forschungstradition gewachsene und nun gesetzlich verankerte Response to Intervention-Ansatz (RTI) zur präventiven und inklusiven Beschulung von Kindern mit Lern- und Entwicklungsproblematiken gewinnt gegenwärtig auch im deutschsprachigen Raum stetig an Bedeutung. Mit dem Rügener Inklusionsmodell (RIM) liegt ein erstes, umfassend und flächendeckend in die Praxis umgesetztes RTI-Konzept in Deutschland vor. Seit dem Schuljahr 2010/2011 werden auf der Insel Rügen Kinder mit und ohne sonderpädagogischem Förderbedarf gemeinsam in die erste Klasse der Grundschule eingeschult und kontinuierlich inklusiv entsprechend des Modellansatzes RIM unterrichtet.

Im Rahmen des Vortrags werden in einem ersten Teil die zentralen Kernelemente des RTI-Ansatzes sowie deren Umsetzung im Beschulungskonzeptes RIM im Überblick dargestellt. In einem zweiten Teil wird die Untersuchung der Wirksamkeit des RIM thematisiert. Forschungsleitend für die Studie ist die Frage, inwieweit der inklusive Unterricht nach dem RIM eine tragfähige Alternative zum herkömmlichen Unterrichts- und Fördersystem in Mecklenburg-Vorpommern darstellt. In einem Zweigruppenversuchsplan wurden dazu die Effekte von Unterricht und Förderung in der Schuleingangsphase in den Grundschulen auf Rügen (Experimentalgruppe) und in der Hansestadt Stralsund (Kontrollgruppe) vergleichend untersucht. Während die Rügener Grundschüler nach dem Rügener Inklusionsmodell (RIM) entsprechend dem Response to Intervention-Ansatz (RTI) unter Verzicht auf Diagnoseförder- und Sprachheilklassen sowie weitgehend ohne Zurückstellungen unterrichtet und gefördert wurden, fand dies in Stralsund entsprechend bisheriger in Mecklenburg-Vorpommern üblicher Arbeitsweisen statt. Ausgehend von den schulischen Lernvoraussetzungen zum Zeitpunkt der Einschulung wurden – getrennt nach Leistungsgruppen – verschiedene statistische Zwillingsgruppen gebildet, um präzise Vergleiche nach drei Schulbesuchsjahren ermöglichen zu können.

**GEBF** TAGUNG  
2014

G04 d

Symposiums-  
beitrag

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G191

Auf dem Weg zu einer inklusiven Schule: Aktuelle Perspektiven

2

## Diagnostische Fähigkeiten von Lehrkräften

### Chair(s): David Schmidt (Ruhr-Universität Bochum)

Ausgangspunkt dieser Dissertation ist zum einen die Bedeutsamkeit der diagnostischen Fähigkeiten von Lehrkräften für alle an der professionellen Lehrerbildung beteiligten Expertengruppen (Helmke & Schrader et al., 2008; Helmke, Hosenfeld & Schrader, 2004; Schrader, 2010; McElvany et al., 2009; Lorenz & Artelt, 2009; Horstkemper, 2006; Langenfeldt, 2006; Bos, Helmke, Leutner & Wirth, 2011), zum anderen die sehr breite Forschungslage zum Themenkomplex der diagnostischen Fähigkeiten von Lehrkräften, die bislang unsystematisch und weder einheitlich noch eindeutig ist (Brünken, 2009; Artelt, 2009; Jürgens, 2012; Schrader, 2009). Statt eines einheitlichen theoretischen Modells, scheinen vielmehr verschiedene subjektive Theorien eine Grundlage der unterschiedlichen Forschungs- und Definitionsansätze zu bilden (Groeben, Wahl, Schlee, & Scheele, 1988; Dann, 1983; Bergold & Breuer, 1987). Vor diesem Hintergrund war es das Ziel dieser Arbeit, die subjektiven Theorien der unterschiedlichen, an der professionellen Lehrerbildung beteiligten Expertengruppen empirisch zu untersuchen und gegenüberzustellen. Aufbauend auf einer Analyse der Unterschiede und Gemeinsamkeiten sollte zudem untersucht werden, ob ein gemeinsamer Definitionskern identifiziert werden kann.

Dazu wurde eine Delphi-Studie mit zwei Befragungsrounds durchgeführt. Die Stichproben wurden jeweils aus den unterschiedlichen, an der professionellen Lehrerbildung beteiligten Expertengruppen gezogen, d.h. aus der Gruppe der Dozenten, die an der universitären Lehrerbildung beteiligt sind, der Fachleiter, Referendare und Lehrkräfte.

Die erste Befragungsrunde diente mit einer Stichprobe von N=323 einer breit angelegten Exploration der subjektiven Theorien der unterschiedlichen Expertengruppen zum Thema diagnostische Fähigkeiten von Lehrkräften. Anhand von Mind Maps wurden hierbei die mentalen Repräsentationen zum Thema diagnostische Fähigkeiten von Lehrkräften erfasst und unter Verwendung der Visualisierungssoftware Akovia (Pirnay-Dummer & Ifenthaler, 2010) für jede Expertengruppe agglomeriert und visualisiert. Die so erhaltenen aggregierten Mind Maps ergaben wie erwartet unterschiedliche inhaltliche Schwerpunkte in den Expertengruppen, jedoch wiesen sie eine sehr begrenzte Struktur auf. Aus diesem Grund wurden nicht die aggregierten Mind Maps als Grundlage in die zweite Befragungsrunde gegeben, sondern im Rahmen von Strukturlegetechniken (Bonato, 1990; Scheele & Groeben, 1988) lediglich die jeweils zwanzig am häufigsten genannten Begriffe pro Expertengruppe.

In der darauf aufbauenden zweiten Runde des Delphis, wurden mit einer kleineren Stichprobe von insgesamt N=30 aus den unterschiedlichen Expertengruppen im Sinne der Strukturlegetechnik Concept Maps aus den jeweils häufigsten Begriffen der ersten Befragungsrunde generiert.

Als deskriptive Ergebnisse der ersten Studie, ließen sich unterschiedliche Inhaltsschwerpunkte in den am häufigsten genannten Begriffen beschreiben. Intergruppenspezifisch kam es dabei inhaltlich zu, zum Teil starken Differenzen, was die Subjektiven Theorien bezüglich des Konstruktes der Diagnostischen Fähigkeiten von Lehrkräften betraf. So ließen sich intergruppenspezifisch zum Beispiel inhaltliche Schwerpunktunterschiede bezüglich des testtheoretischen Hintergrundes dokumentieren. Unter den von Universitätsdozenten (6/20) genannten Begriffen lag der Fokus beispielsweise mehr auf testtheoretischen Hintergründen als bei Lehrkräften (3/20). Bei gleichem Inhaltsschwerpunkt lagen die Werte für Fachleiter (2/20) und Studenten (2/20) sowie bei Referendaren (1/20) noch darunter.

Aufgrund der, von der erstmaligen empirischen inhaltlichen Konstruktbeschreibung abgesehenen, verminderten Aussagekraft und der angesprochenen fehlenden Struktur, der auf den Häufigkeitslisten vorhandenen Begriffe, wurde in der zweiten Studie mittels Strukturlegetechniken eine Struktur über Concept Maps konstruiert.

Die gewonnenen Daten dieser zweiten Befragungsrunde, werden momentan ausgewertet und auf der Tagung präsentiert.

- Artelt, Cordula (2009): Diagnostische Kompetenz von Lehrkräften im Bereich der Lesekompetenz. 8. Göttinger Gespräche 2009. Göttingen, 17.03.2009. Bergold & Breuer, 1987
- Bonato, Marcellus (1990): Wissensstrukturierung mittels Struktur-lege-Techniken. Eine graphentheoretische Analyse von Wissensnetzen. Frankfurt am Main, New York: P. Lang (Europäische Hochschulschriften. Reihe VI, Psychologie Publications universitaires européennes. Série VI, Psychologie European university studies. Series VI, Psychology, 297). Bos, Helmke, Leutner & Wirth, 2011
- Brünken, Roland (2009): Diagnostik in der Lehrerbildung. Vortrag zum 43. BAK Seminartag. Saarbrücken, 02.10.2009.
- Dann, H.-D (1983): Subjektive Theorien: Irrweg oder Forschungsprogramm? Zwischenbilanz eines kognitiven Konstrukts. In: Leo Montada und Hans Aebli (Hg.): Kognition und Handeln. Hans Aebli zum 60. Geburtstag. Stuttgart: Klett-Cotta, S. 77–92.
- Groeben, Norbert; Wahl, Diethelm; Schlee, Jörg; Scheele, Brigitte (1988): Das Forschungsprogramm Subjektive Theorien. Eine Einführung in die Psychologie des reflexiven Subjekts. Tübingen: Francke.
- Schrader, Friedrich-Wilhelm; Helmke, Andreas; Groß Ophoff, Jana; Hosenfeld, Annette; Hosenfeld, Ingmar; Koch, Ursula (2008): Vergleichen und fördern. Von der zutreffenden Diagnose zu erreichbaren Zielen. In: Grundschulmagazin 76, 2008 (5), S. 8–12.
- Helmke, Andreas; Hosenfeld, Ingmar; Schrader, Friedrich-Wilhelm (2004): Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In: R. Arnold und C. Griese (Hg.): Schulmanagement und Schulentwicklung. Hohengehren: Schneider-Verlag, S. 115–135.
- Horstkemper, Marianne (2006): Fördern heißt Diagnostizieren. Pädagogische Diagnostik als wichtige Voraussetzung für individuellen Lernerfolg. In: Gerold Becker, Marianne Horstkemper, Erika Risse, Lutz Stäudel, Rolf Werning und Felix Winter (Hg.): Diagnostizieren und Fördern. Stärken entdecken - Können entwickeln. Friedrich Jahresheft 24, S. 4–7.
- Jürgens, Eiko (2012): Diagnosefunktion: Fehlanzeige. Ist eine Neuorientierung der Beurteilungspraxis nötig? In: Christian Fischer (Hg.): Diagnose und Förderung statt Notengebung? Problemfelder schulischer Leistungsbeurteilung. Münster [u.a.]: Waxmann, S. 23–44. Langenfeldt, 2006
- Lorenz, Christian; Artelt, Cordula (2009): Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. In: Zeitschrift für Pädagogische Psychologie 23 (3-4), S. 211–222.
- Pirnay-Dummer, Pablo; Ifenthaler, Dirk (2010): Automated Knowledge Visualization and Assessment. In: Dirk Ifenthaler, Pablo Pirnay-Dummer und Norbert M. Seel (Hg.): Computer-based diagnostics and systematic analysis of knowledge. 1. Aufl. New York: Springer, S. 77–115.
- Scheele, Brigitte; Groeben, Norbert (1988): Dialog-Konsens-Methoden zur Rekonstruktion subjektiver Theorien. Die Heidelberger Struktur-lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen. Tübingen: A. Francke. Schrader 2010; McElvany et al 2009;
- Schrader, Friedrich-Wilhelm (2009): Anmerkungen zum Themenschwerpunkt Diagnostische Kompetenz von Lehrkräften. In: Zeitschrift für Pädagogische Psychologie 23 (3-4), S. 237–245.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G161



## Zur prädiktiven Kraft diagnostischer Kompetenzen: Schrader vs. Karst – ein Vergleich

Diagnostische Kompetenzen (DK) von Lehrkräften sollten aus normativtheoretischer Perspektive positiv auf die Lernentwicklung ihrer Schülerinnen und Schüler (SuS) wirken. Denn, wenn Lehrkräften die Lernvoraussetzungen ihrer SuS bekannt sind, bietet dies die Möglichkeit zur adaptiven Gestaltung der Lernumgebung. Empirisch lässt sich diese These nur vage unterstützen, da die Befundlage uneinheitlich ist. Einerseits gibt es Studien, welche einen positiven Vorhersagewert für diagnostische Kompetenzen feststellen (Anders, Kunter, Brunner, Krauss & Baumert, 2010). Andererseits existieren Forschungsarbeiten, die keinen direkten Zusammenhang zwischen der Diagnosekompetenz von Lehrkräften und der Lernentwicklung belegen (Schrader & Helmke, 1987). Denjenigen Arbeiten, die sich mit der prädiktiven Kraft diagnostischer Kompetenzen beschäftigen, fehlt bislang der Einbezug einer umfassenden Systematisierung diagnostischer Kompetenzen wie sie von McElvany et al. (2009) oder Karst (2012) vorgeschlagen wird, bei der Diagnosekompetenzen in einem Gitter aus Urteilstebene/ unterrichtlicher Diagnosesituation und Urteilsbereich/ Wissensdomäne verortet werden. Meist basieren die Analysen auf ein bis zwei Komponenten, die sich aus der spezifischen Kombination eines Urteilsbereichs mit einer Urteilstebene ergeben (z.B. Anders et al., 2010), und resultieren erfahrungsgemäß aus einer für die zu berechnenden Modelle notwendigen Sparsamkeit sowie zum Zwecke einer inhaltlichen Fokussierung. Ein Vergleich situations-/ ebenenspezifischer diagnostischer Kompetenzen in ihrer prädiktiven Kraft wird schwieriger.

In diesem Beitrag werden insgesamt drei Ebenen diagnostischer Kompetenzen (klassenbezogene, schülerglobale und schülerspezifische DK) berücksichtigt (Karst, 2012). Dabei wird der Frage nachgegangen wie diese die Lernentwicklung von Grundschulkindern im Fach Mathematik beeinflussen. Zudem werden psychometrisch die Ansätze von Karst (2012) und Schrader (1989) verfolgt und in deren prädiktiver Wirkung verglichen.

Mit Schrader (1989) können je Ebene drei Komponenten diagnostischer Kompetenz (Niveau-, Differenzierungs- und Rangordnungskomponente) unterschieden werden. Karst (2012) differenziert je unterrichtlicher Diagnosesituation verschiedene Teilkompetenzen diagnostischer Kompetenz, die empirisch in einen situationsspezifischen Gesamtindex diagnostischer Kompetenz integriert werden können.

Die Stichprobe umfasst 680 SuS, die in der ersten (t1) und zweiten (t2) Klasse in ihren arithmetischen Kompetenzen getestet wurden. Zudem schätzten 32 Lehrkräfte die Testleistung ihrer Schüler ein (t1). Zur Erfassung der klassenbezogenen Diagnosekompetenz wurden die Lehrkräfte gefragt, wie viele SuS ihrer Klasse bestimmte Aufgaben lösen werden (Einschätzung der Klasse als Ganzes). Die schülerglobale Diagnosekompetenz wurde mit der Frage nach der Anzahl gelöster Aufgaben einzelner SuS erfasst (Einschätzung aller Schüler in Lernbereichen). Für 12 vom Forscherteam randomisiert ausgewählte SuS machten die Lehrkräfte schülerspezifische Aussagen („Wird ihr Schüler bei der jeweiligen Aufgabe zur richtigen oder falschen Lösung gelangen?“).

Pro Diagnosesituation werden drei Schrader-Kennwerte berechnet. Es sind dies der Niveau- und Differenzierungsfehler sowie die Rangordnungskomponente. Zudem wird mittels der Softwarepakete GENOVA 3.1 für balancierte Designs (Crick & Brennan, 2001) und urGENOVA 2.1 für unbalancierte Designs (Brennan, 2001) auf der Grundlage von Varianzkomponentenanalysen ein situationsspezifischer Basisindex ermittelt (vgl. Karst, 2012).

Mit Mehrebenenanalysen in MPlus 7 (Muthén & Muthén, 1998 – 2012) sowie Berechnung von Effektstärken  $\Delta$  nach Tymms (2004) wird die Fragestellung überprüft. Vergleicht man auf Grundlage der Effektstärken nach Tymms (2004) die Vorhersagekraft der Kennwerte nach Schrader mit den situationsspezifischen Basisindizes erscheinen die Unterschiede zunächst minimal. Jedoch zeigen sich Differenzen in der Signifikanz der Prädiktoren. Bei den Schraderkomponenten wird ausschließlich der Regressionskoeffizient für die schülerglobale Rangkomponente negativ signifikant. Für die situationsspezifischen Basisindizes kann festgehalten werden, dass die klassenbezogene DK die Lernentwicklung positiv vorhersagt, während die DK, bei der die Lehrkraft einzelne Schüler in Lernbereichen einschätzt (schülerglobale), die Lernentwicklung negativ prädiziert. Die schülerspezifische DK zeigt keine empirische Relevanz. Die Behauptung, diagnostische Kompetenzen hätten per se einen positiv-prädiktiven Wert, lässt sich demnach nicht untermauern und unterstützt den Bedarf diagnostische Kompetenzen nach unterrichtlichen Diagnosesituationen zu differenzieren. Gleichzeitig werfen die teilweise erwartungswidrigen Befunde die Frage auf, ob mit einer Zunahme diagnostischer Kompetenzen unbedingt eine Verbesserung der Unterrichtsqualität und damit der Lernentwicklung einhergehen muss.

Anders, Y., Kunter, M., Brunner, M., Krauss, S., & Baumert, J. (2010). Diagnostische Fähigkeiten von Mathematiklehrkräften und ihre Auswirkungen auf die Leistungen ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht*, 57, 175-193.

Brennan, R. L. (2001b). *urGENOVA – a generalized analysis of variance system 2.1*. Iowa City, IA: University of Iowa.

Crick, J.E., & Brennan, R.L. (2001). *GENOVA – a generalized analysis of variance system 3.1*. Iowa City, IA: University of Iowa.

Karst, K. (2012). *Kompetenzmodellierung des diagnostischen Urteils von Grundschullehrern*. Münster: Waxmann

McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., et al. (2009). Diagnostische Fähigkeiten von Lehrkräften: Bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit Instruktionellen Bildern. *Zeitschrift für Pädagogische Psychologie*, 23, 223-235.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Schrader, F.-W. (1989). Diagnostische Kompetenzen von Lehrern und ihre Bedeutung für die Gestaltung und Effektivität des Unterrichts. Frankfurt/ Main: Peter Lang.

Schrader, F.-W., & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. *Empirische Pädagogik*, 1, 27-52.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55-66). London: National Foundation for Educational Research.

## **Diagnostizieren findet nicht erst am Ende statt: Wie Lehrende und Lernende von formativen Diagnosen profitieren**

Theoretischer Hintergrund: Durch formatives Diagnostizieren können Lehrende im Verlauf des Lernens Informationen über den Lernenden sammeln. Die gesammelten Informationen können Lehrende nutzen, um ihre Instruktion an den Lernenden in lernförderlicher Weise anzupassen (Black & William, 1998). Sie sind aber auch für die summative Diagnose hilfreich, wenn es darum geht, nach der Instruktion zu beurteilen, welches Verständnis Lernende erworben haben (Birenbaum et al., 2006). Inwieweit formative Diagnosen Lernen fördern und summative Diagnosen verbessern, ist weitestgehend unklar (Kingston & Nash, 2011).

Im Tutoring bieten sich Tutoren besonders viele Gelegenheiten für die formative Diagnose (Snow & Swanson, 1992). Wir interessieren uns speziell dafür, wie gut Tutoren Wissensdefizite eines Tutees formativ diagnostizieren können. Die Diagnose von Wissensdefiziten ist besonders wichtig, weil Wissensdefizite das Lernen massiv behindern (Vosniadou, 1999). Wenn jedoch ein Tutee Wissensdefizite äußert, kann der Tutor ein besseres Verständnis dafür erwerben, was der Tutee nicht weiß, und seine Instruktion entsprechend anpassen.

In der Tutoringforschung wird diskutiert, welche Rolle die Lehrerfahrung für die Genauigkeit von Diagnosen und die Effektivität der Instruktion spielt. Studien deuten an, dass lehrerfahrene Tutoren sowohl in der formativen Diagnose als auch in der summativen Diagnose besser als lehrunerfahrene Tutoren sind (Cromley & Azevedo, 2005; Herppich, Wittwer, Nückles & Renkl, 2013). Systematische Vergleiche von Tutoren mit unterschiedlicher Lehrerfahrung fehlen allerdings bislang.

Fragestellung: Wir untersuchten, ob formatives Diagnostizieren von Wissensdefiziten eines Tutees den Tutee beim Lernen unterstützt und die Genauigkeit der summativen Diagnose des Tutors nach dem Tutoring erhöht. Weiterhin prüften wir, inwieweit Lehrerfahrung das formative Diagnostizieren beeinflusst. Konkret testeten wir folgende Hypothesen:

1. Häufigeres formatives Diagnostizieren nach Wissensdefizitäußerungen fördert (1) das Lernen des Tutees und (2) die Genauigkeit der summativen Diagnose des Tutors nach dem Tutoring.

2. Lehrerfahrene Tutoren (1) fördern das Lernen des Tutees stärker und (2) diagnostizieren genauer summativ nach dem Tutoring als lehrunerfahrene Tutoren, weil lehrerfahrene Tutoren häufiger nach Wissensdefizitäußerungen diagnostizieren als lehrunerfahrene Tutoren.

Methode: Insgesamt  $N = 46$  zufällig zusammengesetzte Tutor-Tutee Dyaden diskutierten jeweils einen Text zum menschlichen Herz-Kreislauf-System. Lehrerfahrene Tutoren waren 21 Biologielehrkräfte. Lehrunerfahrene Tutoren waren 25 Biologiestudierende. Tutees waren Siebtklässlerinnen und Siebtklässler. Die abhängigen Variablen waren (1) die relative Häufigkeit formativer Diagnosen des Tutors nach Wissensdefizitäußerungen des Tutees während der Tutoringsitzung, (2) der Lerngewinn des Tutees, gemessen als Differenz zwischen Prä- und Posttest zu Konzepten des Herz-Kreislauf-Systems und (3) die summative Diagnosegenauigkeit des Tutors nach dem Tutoring, gemessen als Anzahl der Übereinstimmungen zwischen Tutee-Antworten im Konzeptposttest und Tutor-Einschätzungen für den Konzeptposttest.

Ergebnisse: Zwei Regressionsanalysen zeigten, dass häufigere formative Diagnosen nach Wissensdefizitäußerungen das Lernen der Tutees unterstützten,  $R$ -Quadrat = .11,  $F(1, 44) = 5.07$ ,  $\beta = .33$ ,  $p < .05$ ,  $f = .35$  (mittlerer Effekt) und die Genauigkeit der summativen Diagnose des Tutors nach dem Tutoring erhöhten,  $R$ -Quadrat = .16,  $F(1, 44) = 6.17$ ,  $\beta = .41$ ,  $p < .05$ ,  $f = .43$  (großer Effekt).

Weiterhin zeigte das 95% Bootstrappingkonfidenzintervall mit 10000 Stichprobenziehungen im Rahmen einer regressionsbasierten Mediationsanalyse (Hayes, 2009), dass erfahrene Tutoren das Lernen ihres Tutees stärker förderten als unerfahrene Tutoren, weil sie häufiger nach Wissensdefizitäußerungen formativ diagnostizierten (unstandardisierter indirekter Effekt = 0.025,  $\kappa$ -Quadrat = .09 (mittlerer Effekt; Preacher & Kelley, 2011), KI [0.004, 0.063]). Eine analoge Analyse ergab außerdem, dass erfahrene Tutoren genauer summativ diagnostizierten als unerfahrene Tutoren, weil sie häufiger nach Wissensdefizitäußerungen formativ diagnostizierten (unstandardisierter indirekter Effekt = 0.415,  $\kappa$ -Quadrat = .15 (mittlerer Effekt), KI [0.155, 0.853]).

Unsere Ergebnisse zeigen, dass das Lernpotenzial geäußerter Wissensdefizite davon abhängt, ob sie vom Tutor durch formatives Diagnostizieren berücksichtigt werden. Außerdem entwickeln Tutoren durch formatives Diagnostizieren eine genauere Vorstellung davon, was der Tutee nach der Tutoringsitzung gelernt hat. Schließlich fördert Lehrerfahrung Instruktionseffektivität und summative Diagnosegenauigkeit dadurch, dass Tutoren mit Lehrerfahrung besonders häufig formativ diagnostizieren.

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J., ... Nickmans, G. (2006). A learning integrated assessment system. *Educational Research Review*, 1, 61-67. doi: 10.1016/j.edurev.2006.01.001

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-68. doi: 10.1080/0969595980050102

Cromley, J. G., & Azevedo, R. (2005). What do reading tutors do? A naturalistic study of more and less experienced tutors in reading. *Discourse Processes*, 40, 83-113. doi: 10.1207/s15326950dp4002\_1

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76, 408-420. doi: 10.1080/03637750903310360

Herppich, S., Wittwer, J., Nückles, M., & Renkl, A. (2013). Does it make a difference? Investigating the assessment accuracy of teacher tutors and student tutors. *The Journal of Experimental Education*, 81, 242-260. doi: 10.1080/00220973.2012.699900

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30, 28-37.

doi: 10.1111/j.1745-3992.2011.00220.x

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16, 93-115. doi: 10.1037/a0022658

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology*, 43, 583-626. doi: 10.1146/annurev.ps.43.020192.003055

Vosniadou, S. (1999). Conceptual change research: State of the art and future directions. In W. Schnotz, S. Vosniadou, & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 3-13). Amsterdam, The Netherlands: Pergamon.





## **Förderung von Diagnosekompetenz in verschiedenen Domänen: Differenzierte Effekte von Selbsterklärungsprompts**

Diagnosekompetenz gilt in unterschiedlichen Domänen als wichtige Kompetenz, bei der die Fähigkeit zur Integration von wissenschaftlichem Wissen und individueller Erfahrung entscheidend ist. Abhängig von der Domäne unterscheiden sich die Ziele einer Diagnose. In der Pflege hat eine Diagnose zum Ziel, die Auswirkungen einer gesundheitlichen Einschränkung zu erkennen (Evers, 1997). Für eine schulische Lehrkraft ist es ein wichtiges Ziel ein pädagogisches Konzept im Klassenzimmer zu implementieren und die Umsetzung zu diagnostizieren (Vogt & Rogalla, 2009).

Um Lehrkräfte bzw. Pflegekräfte beim Aufbau von Diagnosekompetenz im Klassenzimmer bzw. bei der Patientenversorgung zu unterstützen liegt ein fallbasierter Ansatz nahe. Eine erfolgreich zur Diagnosekompetenzförderung eingesetzte Methode sind Lösungsbeispiele mit integrierten Fehlern (Stark et al., 2011). Eine Möglichkeit um Lerner zu aktiver Verarbeitung der Fehler anzuregen sind Selbsterklärungsprompts (Schworm & Renkl, 2007). Jedoch gibt es Hinweise, dass unterschiedliche Prompts differenzierte Effekte hervorrufen können (Berthold et al., 2009). Um aus den integrierten Fehlern zu lernen ist es notwendig, dass Lerner die Fehler erkennen und verstehen (Schank, 1999). Dazu kann Feedback notwendig sein. Eine Möglichkeit Feedback an die Bedürfnisse des Lerners anzupassen, ist die Nutzung von adaptierbarem Feedback (Leutner, 2002). Im vorliegenden Projekt wurde untersucht ob Lerner mit Selbsterklärungsprompts und adaptierbares Feedback mehr Diagnosekompetenz erwarben. Außerdem wurden die Effekte von verschiedenen Prompts auf den Diagnosekompetenzerwerb analysiert. Inwiefern sich Unterschiede zwischen den Domänen auf die Wirksamkeit der instruktionalen Maßnahmen auswirken war ein weiterer Untersuchungsgegenstand.

Dazu wurden zwei experimentelle Studien mit einem 2x2-faktoriellen Design durchgeführt. In beiden Studien bearbeiteten Lerner fehlerhafte Lösungsbeispiele in einer Online-Lernumgebung. Die Teilnehmer lernten entweder mit oder ohne Selbsterklärungsprompts und mit oder ohne adaptierbarem Feedback. Die Lerner wurden durch drei verschiedene Prompts zur Selbsterklärung aufgefordert. Der erste Prompt zielte auf das Erkennen des Fehlers ab, der zweite auf Problemlösestrategien und Heuristiken und der dritte auf die Verbindung des Falles zum theoretischen Hintergrund. Diagnosekompetenz wurde mittels unterschiedlicher Tests (MC Test, Key Feature- und Problemlöseaufgaben) erfasst. An Studie 1 nahmen N = 108 Studierende des Lehramts und der Pädagogik teil. In Studie 2 setzten sich die Teilnehmer aus N = 152 Krankenpflegeschülern zusammen.

Die Ergebnisse zeigten, dass die Gruppen die mit Selbsterklärungsprompts lernte, schlechter in den Tests abschnitten die auf Handlungswissen abzielten. Dieser Effekt zeigte sich im Lehramt ( $F(1,103)=12.60$ ,  $p<.01$ , partial  $\eta^2=0.11$ ) und in der Pflege ( $F(1,147)=20.92$ ,  $p<.001$ , partial  $\eta^2=0.13$ ). Adaptierbares Feedback hatte keinen signifikanten Einfluss in Studie 1 und 2. Eine Korrelation der Lernzeit der Prompts und Diagnosekompetenz zeigte unterschiedliche Auswirkungen der drei Prompts (Lehramt: Prompt 1  $r=-0.07$ ,  $p=0.56$ ; Prompt 2  $r=0.21$ ,  $p=0.13$ ; Prompt 3  $r=0.32$ ,  $p<0.05$ ; Pflege: Prompt 1  $r=-0.01$ ,  $p=0.92$ ; Prompt 2  $r=-0.06$ ,  $p=0.56$ , Prompt 3  $r=-0.29$ ;  $p<.01$ ). In Studie 1 korrelierte nur der Prompt der die Verbindung des Falls mit dem theoretischen Hintergrund erforderte, positiv und signifikant mit dem Diagnosekompetenzerwerb. In Studie 2 zeigte sich bei diesem Prompt eine negative signifikante Korrelation.

Während Prompts oft generell mit positiven Wirkungen assoziiert werden, fehlen bislang systematische Untersuchungen der differenzierten Effekte insbesondere im Kontext des Lerners mit Fehlern. In unseren Studien hatten Prompts insgesamt einen unerwarteten negativen Einfluss auf den Diagnosekompetenzerwerb. Ähnlich wie in einer weiteren Studie (Berthold et al., 2009) gibt diese Studie Hinweise darauf, dass auch im Kontext des Fehlerlerner die Art des Prompts entscheidend für dessen Wirksamkeit ist. Das Erkennen von Fehlern und von Problemlösestrategien könnten möglicherweise Prozesse sein, die durch die Integration von Fehlern bereits angeregt werden. Prompts würden in diesen Fall keine wirkungsvolle Unterstützung darstellen und gegebenenfalls den Lerner behindern. Die Verbindung zum theoretischen Hintergrund hingegen könnte ein Prozess sein, der nicht durch das bloße Vorhandensein der Fehler ausgelöst wird. Die beiden berichteten Studien liefern zudem Hinweise darauf, dass die Lernwirksamkeit von Prompts abhängig von der Domäne ist.

Berthold, K., Eysink, T. H. S., & Renkl, A. (2009). Assisting Self-Explanation Prompts Are More Effective than Open Prompts when Learning with Multiple Representations. *Instructional Science: An International Journal of the Learning Sciences*, 37(4), 345–363.

Evers, G. C. M. (1997). *Theorien und Prinzipien der Pflegekunde*. Berlin, Wiesbaden: Ullstein Medical.

Leutner, D. (2002). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In *Information und Lernen mit Multimedia*. Weinheim: PVU.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science: A Multidisciplinary Journal*, 21(1), 1–29.

Schank, R. C. (1999). *Dynamic Memory Revisited* (2nd ed.). Cambridge University Press.

Schworm, S., & Renkl, A. (2007). Learning Argumentation Skills through the Use of Prompts for Self-Explaining Examples. *Journal of Educational Psychology*, 99(2), 285–296.

Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, 21(1), 22–33. doi:10.1016/j.learninstruc.2009.10.001

Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency through coaching. *Teaching and Teacher Education*, 25(8), 1051–1060. doi:10.1016/j.tate.2009.04.002

## **Beobachtung und Bewertung von Schülerinteraktionen in kooperativen Lernsettings: Wie kann diese Lehrkompetenz gefördert werden?**

Kooperatives Lernen ist eine gut untersuchte Unterrichtsform (klassische Übersichtsartikel: Lou et al., 1996; Slavin, 1980), deren Wirksamkeit wesentlich von der Qualität der Interaktion der Lernenden abhängt (Dillenbourg, Baker, Blaye, & O'Malley, 1996). Eine hohe Qualität der Interaktion zeichnet sich durch lernförderliche Aktivitäten wie zum Beispiel gegenseitiges Nachfragen und Erklären (Webb, 1989) während der Gruppenarbeit aus. Nur wenn eine hohe Qualität der Interaktion gegeben ist, wird das Ziel dieser Unterrichtsform erreicht, nämlich eine gemeinsame Wissenskonstruktion und daraus folgend der individuelle Lernerfolg jedes Gruppenmitglieds.

Lehrkräfte, die kooperative Lernformen effektiv in ihren Unterricht integrieren wollen, müssen die Qualität der Schülerinteraktionen angemessen beurteilen können. Dazu ist es notwendig, lernförderliche Verhaltensweisen der Schülerinnen und Schüler beim kooperativen Lernen im Unterricht zu erkennen. Bislang wurde die professionelle Wahrnehmung von Unterricht (vgl. Sherin, 2009) angehender Lehrkräfte vor allem bezüglich Klassenführung untersucht (bspw., Seidel, Blomberg, & Stürmer, 2010). In Bezug auf kooperatives Lernen verstehen wir unter einer professionellen Wahrnehmung das Erkennen lernförderlicher Verhaltensweisen während der Gruppenarbeit. Zur Förderung dieser Beobachtungskompetenz bieten sich Videos als Trainingsinstrument an (Janík & Seidel, 2009).

In der hier vorgestellten Studie (durchgeführt im Juli 2013) wurde untersucht, ob die Beobachtung und Bewertung von Schülerinteraktionen durch ein eintägiges, videobasiertes Training bei Lehramtsstudierenden gefördert werden kann.

Zur situativen Erfassung der Beobachtungs- und Bewertungskompetenz wurde ein computerbasiertes Assessment Tool entwickelt. Drei Videos von Schülerinteraktionen sollten auf zwei Ebenen beurteilt werden: ganzheitliche Urteile zur Bewertung der Qualität der gezeigten Schülerinteraktion (Die Qualität der Schülerinteraktion ist insgesamt: sehr niedrig; niedrig; mittel; hoch; sehr hoch) sowie spezifischere Urteile bezüglich Indikatoren lernförderlicher Verhaltensweisen, wie Die Gruppenmitglieder begründen ihre Aussagen. (Skala trifft zu versus trifft nicht zu). Eine korrekte Beurteilung der Interaktion wurde als absolute Übereinstimmung mit einem Expertenurteil operationalisiert. Vor dem Training erhielten die Lehramtsstudierenden eine kurze, computerbasierte Einführung in die Beobachtung und Bewertung von Schülerinteraktionen. Es wurde ein Video einer beispielhaften Schülerinteraktion gezeigt und daran demonstriert, wie die Interaktion bewertet werden kann. Anschließend bearbeiteten die Lehramtsstudierenden das Assessment Tool (Prämesszeitpunkt). Nun begann das eigentliche Training, das nur die Bewertung von Schülerinteraktionen auf der Ebene der Verhaltensindikatoren fokussierte, nicht die Bewertung auf der Ebene der ganzheitlichen Urteile. Dazu wurden die Verhaltensindikatoren in Form einer Checkliste ausgeteilt. Anhand weiterer Videos von Schülerinteraktionen wurde das Beobachten und Bewerten der Verhaltensindikatoren in Kleingruppen geübt und anschließend im Plenum ausführlich besprochen. Nach dem Training bearbeiteten die Lehramtsstudierenden das Assessment Tool erneut (Postmesszeitpunkt).

Es zeigte sich ein Lernzuwachs vom Prä- zum Postmesszeitpunkt für die Beobachtung und Bewertung von Schülerinteraktionen auf spezifischer Ebene mit Verhaltensindikatoren. Des Weiteren konnte gezeigt werden, dass die untersuchte Beobachtungs- und Bewertungskompetenz auf spezifischer Ebene bei Lehramtsstudierenden anfänglich gering ausgeprägt war und erst nach einem Training reliabel erfassbar war. Ein Transfereffekt der trainierten spezifischen Beobachtung und Bewertung der Verhaltensindikatoren auf die Ebene der ganzheitlichen Urteile konnte nicht nachgewiesen werden.

Im Herbst 2013 finden weitere Erhebungen statt, um die Trainingseffekte auf mögliche Moderatoren hin genauer zu analysieren. Neben Lehramtsstudierenden wird das Training auch bei Lehreranwärtern/innen eingesetzt werden. Da Lehreranwärter/innen bereits über praktisches Erfahrungswissen verfügen, soll untersucht werden, ob es Unterschiede zwischen Lehramtsstudierenden und Lehreranwärtern/innen hinsichtlich der Ausprägung und des Effekts des Trainings auf die Beobachtungs- und Bewertungskompetenz gibt.

Nach dem bisherigen Stand fördert das Training die professionelle Wahrnehmung von Schülerinteraktionen. Diese Beobachtungs- und Bewertungskompetenz ist eine wichtige Voraussetzung für eine angemessene adaptive Unterstützung der Schülerinnen und Schüler beim kooperativen Lernen. Die Checkliste mit konkreten Verhaltensindikatoren hilft angehenden Lehrkräften, einzuschätzen, ob die Gruppenarbeit gerade gelingt oder nicht. Sie dient der selektiven Aufmerksamkeit und gewährleistet somit, einen Überblick im komplexen Unterrichtsgeschehen zu bewahren.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines. Towards an interdisciplinary learning science* (1st ed., pp. 189–211). Oxford, New York: Pergamon.

Janík, T., & Seidel, T. (Eds.). (2009). *The power of video studies in investigating teaching and learning in classroom*. Münster: Waxmann.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4), 423-458. doi:10.3102/00346543066004423

Seidel, T., Blomberg, G., & Stürmer, K. (2010). „Observer“ – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Zeitschrift für Pädagogik*, 56. Beiheft, 296-306.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education* 60(1), 20-37.

Slavin, R. E. (1980). Cooperative Learning. *Review of Educational Research*, 50(2), 315–342. doi:10.3102/00346543050002315

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1), 21-39. doi:10.1016/0883-0355(89)90014-1



Chair(s): Ulrich Trautwein (Universität Tübingen)

Diskutant/in: Julia Dietrich (Universität Tübingen)

Seit der Veröffentlichung der Ergebnisse der Studie PISA-2000 hat die Empirische Bildungsforschung einen starken Aufschwung erlebt und gilt heute als unverzichtbarer Grundpfeiler einer evidenzorientierten Bildungspolitik und als Bezugsdisziplin für Fachwissenschaften, Fachdidaktik, Schulpädagogik und Pädagogische Psychologie. Gleichzeitig sind Möglichkeit und Notwendigkeit Empirischer Bildungsforschung in vielen wissenschaftlichen Disziplinen noch immer umstritten. Es wird befürchtet, dass die Empirische Bildungsforschung zu Fehlpriorisierungen der (relativ einfach zu messenden) „Grundkompetenzen“ in den Bereichen Lesen, Mathematik, Fremdsprache und Naturwissenschaften im schulischen Fächerkanon führe, dass sie mit der Gefahr eines reduktionistischen Fehlschlusses vom Messbaren auf das Wirksame und Erstrebenswerte einhergehe und dass die Reflexivität und Vielschichtigkeit menschlichen Denkens und Erlebens, das die Breite der „Bildung“ auszeichne, in der Empirischen Bildungsforschung (bislang) keinen Platz finde. In der Tat sind ganze Kompetenzbereiche, denen von der Gesellschaft und den jeweiligen Fachvertretern und Fachvertreterinnen eine wichtige Rolle zugesprochen wird, in den Arbeiten der Empirischen Bildungsforschung bislang wenig oder gar nicht berücksichtigt.

Dieses Symposium zeigt anhand von zwei „schwer messbaren“ Kompetenzbereichen, nämlich literarästhetischen und historischen Kompetenzen, die Möglichkeiten (und Grenzen) der empirischen „Vermessung“ von Bildung auf. Die ersten beiden Beiträge fokussieren die literarästhetische Urteils- bzw. Verstehenskompetenz (LUK). Im ersten Beitrag stellen Volker Frederking und Kollegen theoretische Grundlagen und empirische Befunde eines um literarisches Fachwissen, ästhetische Aufmerksamkeit und textseitig intendierte Emotionen erweiterten Modells literarästhetischer Verstehenskompetenz vor und diskutieren die damit verbundenen Chancen für die Entwicklung eines evidenzbasierten Förderkonzepts zum Aufbau bzw. zur Vertiefung literarischer Textverstehenskompetenz. Im zweiten Beitrag zeigen Jörn Brüggemann und Kollegen, dass und wie sich literarisch kodierte Emotionen, denen traditionell in der Literaturdidaktik eine exponierte Bedeutung für literarisches Textverstehen zugesprochen wird, kompetenztheoretisch modellieren und empirisch erheben lassen und welche Bedeutung dies für eine evidenzorientierte Literaturdidaktik besitzt. Waltraud Schreiber und Kollegen zeigen auf, dass eine Messung historischer Kompetenzen mithilfe von Multiple-Choice-Items sowohl unterschiedliche Kompetenzbereiche abzudecken vermag als auch zu einem intern reliablen Gesamttest führen kann. Im abschließenden Beitrag stellen Johannes Meyer-Hamme und Andreas Körber ein Konzept zur Erfassung einer Kernkompetenz der im FUER-Modell definierten "Historischen Fragekompetenz" vor und diskutieren den gegenwärtigen Stand und die sich daraus ergebenden Herausforderungen für die weitere Entwicklung von Messinstrumenten.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G165

Volker Frederking (Universität Erlangen- Nürnberg), Jörn Brüggemann (Universität Erlangen-Nürnberg), Thorsten Roick (Humboldt-Universität zu Berlin), Christel Meier (Universität Erlangen-Nürnberg)

## **Ein erweitertes Modell literarästhetischer Verstehenskompetenz – theoretische Grundlagen und empirische Befunde**

Theoretischer Hintergrund: Das Verstehen literarischer Texte steht traditionell im Zentrum literaturwissenschaftlicher und literaturdidaktischer Forschung. Seit PISA 2000 und dem dadurch initiierten Übergang von der Input- zur Outputorientierung hat sich der Fokus auf damit korrespondierende Kompetenzen erweitert (z. B. Zabka, 2012). Auch in den Bildungsstandards und neueren schulischen Curricula für das Fach Deutsch (z. B. KMK, 2003, 2012) wird dem kompetenten Umgang mit literarischen Texten eine exponierte Bedeutung zugesprochen. Dass damit Kompetenzanforderungen verbunden sind, die sich von denen im Umgang mit faktualen Texten unterscheiden, ist eine Annahme, die in Literaturdidaktik und Literaturwissenschaft weithin geteilt wird (z. B. Kammler, 2006) – allerdings lange Zeit ohne empirische Absicherung. In dem interdisziplinär angelegten Forschungsprojekt ‚Literarästhetische Urteilskompetenz‘ (LUK), das Teil des DFG-Schwerpunktprogramms 1293 ‚Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen‘ war, ist der Versuch unternommen worden, diese Forschungslücke zu schließen (Frederking, Roick & Steinhauer, 2010; Frederking et al., 2012; Roick et al., 2010; Roick, Frederking, Henschel, & Meier, 2013; Henschel, Roick, Martin, & Stanat, 2013). Auf Basis der ästhetischen Semiotik Umberto Ecos (1972, 1990) ist ein Modell literarästhetischer Textverstehenskompetenz entwickelt und in mehreren Querschnittstudien empirisch überprüft worden, in dem zwischen zwei Teildimensionen unterschieden wird, einer semantischen und einer idiolektalen. Erstere bezeichnet die Fähigkeit zum Erschließen zentraler Inhalte, Sinnstrukturen und Deutungsspielräume eines literarischen Textes, letztere die Fähigkeit zum Erfassen seiner formalen Spezifika und deren ästhetischer Funktion.

Fragestellung: Im Rahmen des Beitrags soll der Frage nachgegangen werden, ob sich das Modell literarästhetischer Textverstehenskompetenz um vier Subfacetten erweitern lässt, denen traditionell im Literaturunterricht eine exponierte Bedeutung für literarisches Verstehen zugeschrieben wird: literarisches Fachwissen (Meier, Henschel, Roick, & Frederking, 2012), ästhetische Aufmerksamkeit und literarisch thematisierte bzw. textseitig intendierte Emotionen (Frederking & Brüggemann, 2012).

Methode: 872 Jugendliche (50% Mädchen; Durchschnittsalter 16,72) aus 42 Klassen der Jahrgangsstufe 10 an bayerischen Realschulen und Gymnasien haben Testaufgaben zum literarischen Verstehen bearbeitet. Diese bestehen aus Items zu semantischen und idiolektalen Verstehenskomponenten, zum Fachwissen, zur ästhetischen Aufmerksamkeit und zu literarisch thematisierten und intendierten Emotionen.

Ergebnisse: Die Strukturanalysen weisen darauf hin, dass literarisches Fachwissen, ästhetische Aufmerksamkeit und das Erkennen intendierter Emotionen abgrenzbare kognitive Anforderungen im literarischen Verstehensprozess darstellen, die das bisherige Modell erweitern. Literarische Textverstehenskompetenz lässt sich mithin als ein Konstrukt beschreiben, das neben semantischen und idiolektalen Teildimensionen mit Fachwissen, ästhetischer Aufmerksamkeit und intendierten Emotionen drei weitere Subfacetten aufweist.

Diskussion: Für eine evidenzorientierte Literaturdidaktik liefern die Ergebnisse erste Hinweise darauf, wie im Literaturunterricht literarische Verstehenskompetenz systematisch entwickelt bzw. verbessert werden kann. Über spezifische Problemstellungen im semantischen und idiolektalen Bereich hinaus (Brüggemann, 2013; Frederking, 2013) haben sich mit Fachwissen, ästhetischer Aufmerksamkeit und intendierten Emotionen weitere konkrete Ansatzpunkte ermitteln lassen, die zur Entwicklung eines empirisch abgesicherten Förderkonzepts zum Aufbau bzw. zur Vertiefung literarischer Verstehenskompetenz im Literaturunterricht genutzt werden können.

Brüggemann, J. (2013). Literarische Verstehenskompetenz und ihre Förderung in der Sekundarstufe II: Konstruktionsprinzipien, normative Implikationen und intendierte Wirkungen von Lernaufgaben. In S. Gailberger & F. Wietzke (Hrsg.): Handbuch Kompetenzorientierter Deutschunterricht: Diagnostizieren – Binnendifferenzieren – Fördern (S. 145-171). Weinheim, Basel: Beltz.

Eco, U. (1972). Einführung in die Semiotik. München: Fink.

Eco, U. (1990). Die Grenzen der Interpretation. München: dtv.

Frederking, V., Roick, T. & Steinhauer, L. (2010). Literarästhetische Urteilskompetenz – Forschungsansatz und Zwischenergebnisse. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, M. Rothgangel, L. Schön, H. J. Vollmer & G. Weigand (Hrsg.), Empirische Fundierung in den Fachdidaktiken. Fachdidaktische Forschung (Band 1) (S. 75-94). Münster: Waxmann.

Frederking, V., Henschel, S., Meier, Ch., Roick, T., Stanat, P., & Dickhäuser, O. (2012). Beyond Functional Aspects of Reading Literacy: Theoretical Structure and Empirical Validity of Literary Literacy. Special issue guest edited by I. Pieper & T. Janssen. L1- Educational Studies in Language and Literature, 12, 35-56.

Frederking, V. & Brüggemann, J. (2012). Literarisch kodierte, intendierte bzw. evozierte Emotionen und literarästhetische Verstehenskompetenz. Theoretische Grundlagen einer empirischen Erforschung. In D. A. Frickel, C. Kammler, & G. Rupp (Hrsg.): Literaturdidaktik im Zeichen von Kompetenzorientierung und Empirie (S. 15-40). Freiburg/Br.: Fillibach.

Frederking, V. (2013). Literarisches Verstehen. In S. Gailberger & F. Wietzke (Hrsg.): Handbuch Kompetenzorientierter Deutschunterricht (S. 117-144). Weinheim, Basel: Beltz Juventa Verlag.

Henschel, S., Roick, T., Martin, B. & Stanat, P. (2013). Leseselbstkonzept und Textart: Lassen sich literarisches und faktuales Leseselbstkonzept trennen? Zeitschrift für Pädagogische Psychologie, 27, 181-191.

Kammler, Clemens (Hrsg.). (2006). Literarische Kompetenzen – Standards im Literaturunterricht. Modelle für die Primar- und Sekundarstufe. Seelze: Kallmeyer / Klett.

KMK (2003). Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. München: Luchterhand.

KMK (2012). Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife. Retrieved from <http://www.kmk.org/fileadmin/veroeffentlichungenbeschluesse/2012/20121018-Bildungsstandards-Deutsch-Abi.pdf>.

Meier, Ch., Henschel, S., Roick, Th. & Frederking, V. (2012). Literarästhetische Textverstehenskompetenz und fachliches Wissen: Möglichkeiten und Probleme domänenspezifischer Kompetenzforschung. In I. Pieper & D. Wieser (Hrsg.): Fachliches Wissen und literarisches Verstehen (S. 237-258). Frankfurt: M. Lang.

Roick, T., Stanat, P., Dickhäuser, O., Frederking, V., Meier, C. & Steinhauer, L. (2010). Strukturelle und kriteriale Validität der literarästhetischen Urteilskompetenz. Projekt literarästhetische Urteilskompetenz. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. 56. Beiheft der Zeitschrift für Pädagogik (S. 165-174). Weinheim: Beltz.

Roick, T., Frederking, V., Henschel, S. & Meier, C. (2013). Literarische Textverstehenskompetenz bei Schülerinnen und Schülern unterschiedlicher Schulformen. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), Literalität erfassen: bildungspolitisch, kulturell, individuell (S. 69-84). Weinheim: Juventa.

Zabka, T. (2012). Hinweise zum Aufbau literarischer Kompetenz in der Sekundarstufe II. *ide*, 36 (1), 108-118.

GEBF TAGUNG  
2014

G06 a

Symposiums  
beitrag

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G165



Literarästhetische und Historische Kompetenzen



Jörn Brüggemann (Friedrich-Alexander-Universität Erlangen- Nürnberg), Volker Frederking (Friedrich-Alexander-Universität Erlangen-Nürnberg), Sofie Henschel (IQB, Humboldt-Universität zu Berlin), Christel Meier (Friedrich-Alexander-Universität Erlangen-Nürnberg), Thorsten Roick (Humboldt-Universität zu Berlin)

## **Literarisch kodierte Emotionen im Fokus der literarischen Textverstehenskompetenzforschung**

Theoretischer Hintergrund: In der Literaturdidaktik ist die Überzeugung vorherrschend, dass literarisches Textverstehen auf Imagination und Emotionen basiert (Spinner, 2006; Kämper-van den Boogaart, & Pieper, 2008; Jesch, 2012). Die Bedeutung von Emotionen als Komponenten literarischer Textverstehenskompetenz ist bislang jedoch kaum systematisch untersucht worden, da die meisten empirisch validierten Kompetenzmodelle kognitive Aspekte akzentuieren oder die kognitive Verarbeitung emotionaler Aspekte ignorieren (Kneepkens & Zwaan, 1994). Nachdem in den ersten beiden Laufzeiten des Projekts „Literarästhetische Urteilskompetenz“ (LUK) der Fokus auf der Erforschung der kognitiven Verstehensdimensionen literarästhetischer Urteilskompetenz gelegen hat, wurde in der dritten Projektlaufzeit die Frage untersucht, inwieweit sich auch emotionale Aspekte literarischen Verstehens kompetenztheoretisch modellieren und empirisch erfassen lassen. Diese Frage hat fachspezifisch eine besondere Bedeutung, weil innerhalb der Literaturdidaktik die Erweiterung kognitiv ausgerichteter Kompetenzmodelle um emotionale Aspekte als Gradmesser für die Legitimität bislang hoch umstrittener quantitativer empirischer Forschungen betrachtet wird (Steinbrenner, 2009). Im Rahmen des Projekts wurden auf der Basis theoretischer Vorüberlegungen (Frederking & Brüggemann, 2012) drei Typen literarisch kodierter Emotionen theoretisch hergeleitet und empirisch überprüft: literarisch präsentierte, intendierte und evozierte Emotionen. Während präsentierte Emotionen im Text in Form von emotionsbezeichnenden und -beschreibenden lexikalischen Ausdrucksvarianten auf der Wortebene sowie mittels Propositionen auf der Satzebene repräsentiert sind, beziehen sich intendierte Emotionen auf das Spektrum der Emotionen, die durch den Text beim Ideal-Leser (im Sinne der ästhetischen Semiotik) ausgelöst werden sollen. Evozierte Emotionen sind die tatsächlich im Leser ausgelösten Emotionen, die einer kompetenztheoretischen Modellierung im Rahmen des LUK-Konstrukts nicht zugänglich sind.

Fragestellung: Untersucht werden soll, ob die Fähigkeit zum Erfassen und Verarbeiten literarisch kodierter präsentierte und intendierter Emotionen eine unterscheidbare Subfacette bzw. Teildimension literarischer Textverstehenskompetenz darstellt und sich von semantischen und idiolektalen Verstehensanforderungen abgrenzen lässt. Auf der Basis der Ergebnisse soll diskutiert werden, in welchen Bereichen Ansatzpunkte für weitere evidenzbasierte Untersuchungen zur Entwicklung von literarischer Textverstehenskompetenz bestehen.

Methode: Im Zuge der Erhebungen haben 872 Zehntklässler aus Gymnasien und Realschulen jeweils drei von acht Testlets zum literarischen Textverstehen mit insgesamt ca. 30 Items bearbeitet. Die acht Testlets wurden in einem Multi-Matrix-Design eingesetzt, sodass jeder Proband nur eine Teilmenge der erhobenen Testaufgaben zu bearbeiten hatte. Die präsentierten Emo-Items, die in Analogie zu den semantischen und idiolektalen LUK-Items konstruiert wurden, sowie die intendierten und evozierten Emo-Items wurden für jedes Testlet separat erfasst, wobei die evozierte Emotionalität direkt nach der Lektüre erfragt wurde und die Items zur intendierten Emotionalität relativ weit am Ende der Unit bearbeitet wurden. Um intendierte und evozierte Emotionalität zu vergleichen, wurden beide Itemtypen zweistufig kodiert (kann zutreffen / trifft nicht zu); anschließend wurden die evozierten Items im Sinne der intendierten als richtig/falsch bewertet. Auf dieser Basis wurde für jeden Proband einen „Fähigkeitsparameter“ für beide Typen bestimmt. Zusätzlich wurden bei den Schülerinnen und Schülern mittels Fragebogen für das literarische Textverstehen bedeutsame Bedingungsfaktoren –z.B. Geschlecht und Empathie – erhoben.

Ergebnisse: Die empirischen Ergebnisse zeigen, dass die unterschiedlichen Typen literarisch kodierter Emotionalität (präsentiert, intendiert) voneinander abgegrenzt werden können. Während die Fähigkeit zum Erfassen und Verarbeiten von präsentierten Emotionen einen hohen Zusammenhang zu den semantischen und idiolektalen Verstehensanforderungen aufweist, stützen die Ergebnisse zum Erfassen textseitig intendierter Emotionen hingegen die Annahme, dass diese Fähigkeitskomponente als unterscheidbare Teildimension literarischer Textverstehenskompetenz zu betrachten ist.

Diskussion: Die Befunde liefern erste Hinweise darauf, in welchen Bereichen Ansatzpunkte für eine Überschreitung der rein kognitiv orientierten evidenzbasierten Kompetenzforschung zugunsten emotionaler Komponenten des literarischen Textverstehens bestehen. Hierzu zählen auch die Untersuchung des Verhältnisses von textseitig intendierten und real evozierten Emotionen und die Bedeutung von Empathie (Henschel & Roick, 2013) für die Ausprägung von Teildimensionen literarischer Textverstehenskompetenz. Vor diesem Hintergrund soll skizziert werden, mit welchen Testverfahren emotionale Aspekte literarischer Lektüren erhoben werden könnten, um die Anwendbarkeit des LUK-Modells für die empirische Unterrichtsforschung zu optimieren.

Frederking, V. & Brüggemann, J. (2012). Literarisch kodierte, intendierte bzw. evozierte Emotionen und literarästhetische Verstehenskompetenz. Theoretische Grundlagen einer empirischen Erforschung. In D. Frickel, C. Kammler, & G. Rupp (Hrsg.), Literaturdidaktik im Zeichen von Kompetenzorientierung und Empirie: Perspektiven und Probleme (S. 15-41). Freiburg: Fillibach.

Henschel, S. & Roick, T. (2013). Zusammenhang zwischen Empathie und dem Verstehen literarischer Texte. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 45, 103-113.

Jesch, T. (2012). Aktuelles Fachwissen zur narrativen Perspektive und sein Ertrag für den Literaturunterricht. In I. Pieper & D. Wieser (Hrsg.), Fachliches Wissen und literarisches Verstehen: Studien zu einer brisanten Relation (S. 91-111). Frankfurt a.M., Berlin, Bern u.a.: Peter Lang.

Kämper-van den Boogaart, M. & Pieper, I. (2008). Literarisches Lesen. Didaktik Deutsch. Sonderheft, 46-65.

Kneepkens, E. W. E. M., & Zwaan, R. A. (1994). Emotions and literary text comprehension. Poetics, 23, 125-138.

Roick, T., Stanat, P., Dickhäuser, O., Frederking, V., Meier, C. & Steinhilber, L. (2010). Strukturelle und kriteriale Validität der literarästhetischen Urteilskompetenz. In E. Klieme, D. Leutner, M. Kenk (Hrsg.), Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes (56. Beiheft der Zeitschrift für Pädagogik) (S. 165-174). Weinheim, Basel: Beltz.

Spinner, K. H. (2006). Literarisches Lernen. Praxis Deutsch, 200, 6-16.

Steinbrenner, M. (2009). "Ausgehen von der Individualität" (Schleiermacher). Zu Karlheinz Fingerhuts Beitrag in Didaktik Deutsch 24 und Iris Winklers Erwiderung in Didaktik Deutsch 25. Didaktik Deutsch, 26, 5-11.





## **Historische Kompetenzen: Zur Entwicklung eines standardisierten Testinstruments für Large-Scale-Studien**

Theoretischer Hintergrund: Bislang wurden in deutschen und internationalen Large-Scale-Untersuchungen historische Kompetenzen nicht erfasst. Dies liegt unter anderem an einer lange vorherrschenden Tradition von Leistungsmessung, welche sich auf konkrete, im Unterricht vermittelte historische Narrative und Deutungen bezog, nicht aber auf transferable Kompetenzen historischen Denkens (z. B. VanSledright, 2014, aktuell für die USA). Frühere (z. B. v.Borries, 1974) und aktuelle (z. B. VanSledright, 2014) Ansätze einer Testung von historischen Denken in einem solchen thematischen Rahmen haben sich nicht durchgesetzt, bzw. erfüllen nicht die Anforderung an historische Kompetenzen im Sinne der Weinert'schen Definition als fachspezifische Problemlösefähigkeiten. Andere, diesbezüglich einschlägigere, Konzepte betreffen zumeist lediglich Teilfähigkeiten (z. B. Hartmann 2008, 2009; Hartmann, Sauer, & Hasselhorn, 2009; Bertram, Wagner, & Trautwein, 2013) bzw. die Verfügung über einschlägige fachliche Konzepte, vor allem die wichtigen „second order concepts“ (Seixas, 2008). Ein zeitökonomisch einsetzbares Messinstrument, welches eine reliable, zugleich aber auch die Einzeldimensionen übersteigende Erfassung historischer Kompetenzen ermöglicht, fehlt hingegen.

In dem Verbundprojekt „Historical Thinking – Competencies in History“ (HiTCH), das vom BMBF gefördert wird, geht es um die Entwicklung eines solchen Kompetenztests, der auch für den Einsatz in Large-Scale-Assessments gegen Ende der Sekundarstufe I geeignet ist. Mit dem Test soll gemessen werden, inwiefern Schüler über die Kompetenz „how to think history“ (also nicht bzw. nicht nur über „Wissen“) verfügen. Als heuristischer Rahmen bei der Itemgenerierung und Validierung des Tests wird das in der Fachdidaktik Geschichte prominente Kompetenz-Strukturmodell historischen Denkens (Körber, Schreiber, & Schöner, 2007) der sogenannten FUER-Gruppe zugrunde gelegt, da es (1) mit seinem Kompetenzbegriff direkt an Schulleistungsstudien angeschlossen ist (Kompetenzdefinition von Weinert, 2001), (2) relativ umfassend ist und damit auch Kompetenzdefinitionen anderer Modelle sowie Messkonzepte zu einzelnen Fähigkeiten, etwa aus der internationalen Forschung, integriert werden können (Second Order Concepts aus dem anglo-amerikanischen Raum, u.a. Seixas, 2008) (3) eine Differenzierung der Kompetenzbereiche vornimmt, die als Ausgangspunkt für die Entwicklung von Items dienen können, und (4) auf der Basis einer klaren Logik eine Niveauunterscheidung (Graduierung) historischer Kompetenzen erlaubt. Nach dem FUER-Modell konstituieren deutlich überlappende Kompetenzen in vier Bereichen (historische Frage-, Methoden-, Orientierungs- und Sachkompetenzen) die historische Kompetenz (Schreiber et al., 2007).

Fragestellung: Untersucht werden soll, ob sich mit Multiple-Choice-Fragen die historische Kompetenz reliabel erfassen lässt. Zudem soll geprüft werden, ob sich Hinweise darauf finden lassen, dass die im FUER-Modell postulierten Domänen modellierbare Subfacetten darstellen.

Methode: In dem Vortrag werden die Befunde der ersten empirischen Teilstudien vorgestellt werden, an denen mehr als 1.500 Schüler/innen der 9. Jahrgangsstufe aller Schulformen in Deutschland teilnahmen.

Ergebnisse: Es zeigt sich, dass standardisierte Tests mit Multiple-Choice-Format auf große Akzeptanz der Schüler/innen stoßen. Es fanden sich zudem in umfassenden konfirmatorischen Faktorenanalysen bereits deutliche Belege dafür, dass sich mit einer überschaubaren Anzahl von Aufgaben ein reliabler Test konzipieren lässt. Neben einem Generalfaktor, der als „globale historische Kompetenz“ bezeichnet werden kann und aus Items aus allen vier Kompetenzbereichen des FUER-Modells gespeist wird, fanden sich auch erste Hinweise auf eine Mehrdimensionalität. In den Vortrag sollen auch die Befunde einer zweiten Datenerhebung mit 1.700 Schüler/innen der Klassenstufen 9 bis 12 in Belgien einfließen, die im Juni 2013 stattfand.

Diskussion: Die Befunde werden im Zusammenhang mit der Debatte um „schwer messbare“ (Frederking, 2008) Kompetenzen kritisch diskutiert.

Borries, B. v.(1974). Lernziele und Testaufgaben für den Geschichtsunterricht, dargestellt an der Behandlung der Römischen Republik in der 7. Klasse (2. Aufl.). Stuttgart: Klett.

Borries, B. v. (1976). "Wie man Geschichtstests nicht machen darf". Geschichtsdidaktik, 1 (2), 22–37.

Bertram, C., Wagner, W. & Trautwein, U. (2013). Chancen und Risiken von Zeitzugbefragungen – Entwicklung eines Messinstruments für eine Interventionsstudie. In J. Hodel & B. Ziegler (Hrsg.), Forschungswerkstatt Geschichtsdidaktik 12: Beiträge zur Tagung "geschichtsdidaktik empirisch 12" (S. 108-119). Bern: Hep-Verlag.

Frederking, V. (Hrsg). (2008). Schwer messbare Kompetenzen: Herausforderungen für die empirische Fachdidaktik. Baltmannsweiler: Schneider Hohengehren.

Hartmann, U. (2008). Perspektivenübernahme als eine Kompetenz historischen Verstehens (Dissertation). Retrieved from <http://ediss.uni-goettingen.de/bitstream/handle/11858/00-1735-0000-0006-AD13-2/hartmann.pdf?sequence=1>.

Hartmann, U. (2009). "Kompetenzprofile historischer Perspektivenübernahme in Klasse 7 und 10." In J. Hodel, B. Ziegler (Hrsg.), Forschungswerkstatt Geschichtsdidaktik 07 (S. 79-89). hep: Bern.

Hartmann, U., Sauer, M., Hasselhorn, M. (2009). "Perspektivenübernahme als Kompetenz für den Geschichtsunterricht". Zeitschrift für Erziehungswissenschaft 12 (2), 321–342.

Körber, A., Schreiber, W. & Schöner, A. (Hrsg.). (2007). Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Neuried: ars una.

Schreiber, W., Körber, A., Borries, B. v., Krammer, R., Leutner-Ramme, S., Mebus, S.,... Ziegler, B. (2007). Historisches Denken. Ein Kompetenz-Strukturmodell (Basisbeitrag). In A. Körber, W. Schreiber & A. Schöner (Hrsg.),

Kompetenzen historischen Denkens: Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik (S. 17-53). Neuried: ars una.

Seixas, P. (2008). "Scaling up" the benchmarks of Historical Thinking: A Report on the Vancouver Meetings, February 14-15, 2008. Retrieved from <http://www.historybenchmarks.ca/sites/default/files/Scaling%20Up%20Meeting%20Report.pdf>.

VanSledright, B. A. (2014). Assessing Historical Thinking and Understanding: Innovative Designs for New Standards. New York, Oxon: Taylor & Francis.

Weinert, Franz E. (2001). „Schulleistungen – Leistungen der Schule oder der Schüler?“ In F. E. Weinert (Hrsg.), Leistungsmessungen in Schule (S.72-101). Weinheim,Basel.



## **Erfassung Historischer Fragekompetenz als Teildimension historischen Denkens**

Theoretischer Hintergrund: Historische Kompetenz besteht dem FUER-Modell (Körber, Schreiber, & Schöner, 2007) zufolge aus einer Kombination dreier prozeduraler Teilkompetenzen, die entlang einem regelkreisförmigen Modell zeitlicher Orientierung angeordnet sind, wie auch einer zu allen beitragenden nicht-prozeduralen Teilkompetenz. In diesem Modell nimmt die Fähigkeit, aktuelle eigene (individuelle wie kollektive) Orientierungsbedürfnisse in bearbeitbare historische Fragen umzusetzen sowie historische Fragen anderer hinsichtlich der ihnen zugrunde liegenden Orientierungsbedürfnisse zu analysieren (und auf eigene zu beziehen) eine besondere Rolle ein. Dazu gehört insbesondere die Fähigkeit, den spezifisch historisch-narrativen Charakter von Fragestellungen zu erkennen. Dem Modell zufolge bedeutet dies, einerseits die (nicht notwendig eindeutige) Zuordnung vorliegender Fragen zu den „Fokussierungen“ (Vergangenheit; Geschichte; Gegenwart/Zukunft) des Modells vorzunehmen und andererseits die Abgrenzung zu nicht-historischen Fragen zu erfassen.

Fragestellung: Untersucht werden soll, inwiefern die so verstandene Fragekompetenz mit Hilfe des genannten Messkonzepts reliabel gemessen werden kann. Angesichts des Anspruchs von Kompetenzmessung, thematisch transferable Fähigkeiten zu messen, stellt die thematische „Einkleidung“ des Instruments und ihr Einfluss auf die Reliabilität eine besondere Herausforderung dar. Vorgestellt und diskutiert werden mehrere verschiedene Operationalisierungen, die sich auf unterschiedliche historische Themen beziehen.

Methode: In dem Vortrag werden entsprechende Befunde aus der ersten empirischen Teilstudie des Projekts HiTCH vorgestellt, an denen mehr als 1500 Schülerinnen und Schüler der 9. Jahrgangsstufe unterschiedlicher Schulformen in mehreren Bundesländern teilnahmen, vorgestellt.

Ergebnisse: Es zeigt sich, dass mehrere Fassungen des Instruments reliable Skalen mit relevanten Interkorrelationen zu anderen Teilkonstrukten und Beiträgen zu einem Gesamtfaktor historischer Kompetenz beitragen ergeben, die allerdings teilweise deutlich unterschiedliche sachliche Struktur aufweisen.

Diskussion: Diskutiert werden diese Operationalisierungen und Ergebnisse im Zusammenhang mit der Debatte um „schwer messbare Kompetenzen“ mit Blick auf die Möglichkeit, den Einfluss der thematischen Einkleidung und somit der konkreten Orientierungen zugunsten einer Messung thematisch transferabler Fähigkeiten zu reduzieren.

Hauschke, L. M. (2012). Empirische Untersuchungen zur Relevanz historischer Fragen im Geschichtsunterricht aus Lehrerperspektive (Masterarbeit). Münster(Westf.): Lit.  
Körber, A., Borries, B. v., Pflüger, C., Schreiber, W. & Ziegler, B. (2008) "Sind Kompetenzen historischen Denkens messbar?" In: V. Frederking (Hrsg.), Schwer messbare Kompetenzen: Herausforderungen für die empirische Fachdidaktik (S. 65-84). Baltmannsweiler: Schneider Hohengehren.  
Körber, A., Schreiber, W., Schöner, A. (Hrsg.).(2007). Kompetenzen historischen Denkens: Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik: Neuried: ars una Verlags-Gesellschaft.

Chair(s): Steffi Pohl (Freie Universität Berlin), Claus Carstensen (Otto-Friedrich-Universität Bamberg)

Diskutant/in: Frank Goldhammer (Deutsche Institut für Internationale Pädagogische Forschung)

Das Nationale Bildungspanel (NEPS, Blossfeld, Roßbach, & von Maurice, 2011) ist eine großangelegte längsschnittliche Studie, die zum Ziel hat Kompetenzen, Bildungsverläufe und –entscheidungen sowie deren Einflussfaktoren über die gesamte Lebensspanne zu erfassen. Kompetenzen in beispielsweise Lesen, Mathematik und Naturwissenschaften werden kohärent über die gesamte Lebensspanne erfasst (Weinert et al., 2011). Als eine spezielle Zielgruppe nimmt das NEPS Schüler mit sonderpädagogischem Förderbedarf in die Erhebungen auf (Heydrich et al., 2013). Die Daten werden in Form von Scientific Use Files der Öffentlichkeit zur Verfügung gestellt (Barkow et al., 2011).

Das NEPS bietet durch das längsschnittliche Design, die breite Altersspanne, die Vielfalt an erhobenen Kontext-, Entscheidungs- und Bildungsrenditevariablen aus unterschiedlichsten Disziplinen ein großes Potential zur Untersuchung verschiedenster inhaltlicher Fragestellungen. Dazu müssen die Daten jedoch bestimmten Anforderungen genügen. Zum Beispiel müssen Kompetenzen reliabel und valide erfasst sowie Messfehler modelliert und berücksichtigt werden. Weiterhin erfordern Analysen zur Kompetenzentwicklung, dass die Kompetenztests kohärent über die Lebensspanne dasselbe Konstrukt messen. In diesem Symposium wird ein Einblick in die methodischen Herausforderungen, die bei der Kompetenzmessung und -skalierung in (längsschnittlichen) Large-Scale-Assessments wie dem NEPS auftreten, gegeben, die Verknüpfung zu inhaltlichen Fragestellungen aufgezeigt und Lösungsansätze präsentiert. Dazu gehen wir von der Erhebung der Kompetenzen über die Bereitstellung von Daten für den Scientific Use File bis zur Analyse von Fragestellungen auf verschiedenen Bereiche der Kompetenzmessung ein.

Im Symposium wird konkret vorgestellt, wie Förderschüler in die Erhebung einbezogen werden können und für welche Förderschüler es möglich ist, reliable und valide Kompetenzmessungen zu erhalten. In einem weiteren Beitrag wird die Testbearbeitung anhand von Informationen zu fehlenden Werten genauer untersucht und es wird der Frage nachgegangen, ob das Überspringen von Aufgaben eine stabile Personeneigenschaft ist und wie sich verschiedene Altersgruppen in ihrer Testbearbeitung unterscheiden. Nachdem die ersten beiden Vorträge die reliable und valide Erfassung von Kompetenzen betrachten, befassen sich die letzten beiden Vorträge mit der Nutzung der Daten zur Untersuchung von inhaltlichen Fragestellungen. Im dritten Vortrag wird der Frage nachgegangen, mit welchen Modellen man dem Nutzer von Scientific Use Files plausible values der Kompetenzen zur Verfügung stellen kann, die eine Modellierung latenter Zusammenhänge mit Kontextvariablen und persönlichkeitsbezogenen Konstrukten erlauben. Im vierten Beitrag wird die längsschnittliche Perspektive des NEPS betrachtet und es wird untersucht, inwieweit Tests für unterschiedliche Altersstufen tatsächlich dasselbe Konstrukt messen und somit die Untersuchung von Kompetenzentwicklung ermöglichen. Die Beiträge des Symposiums zeigen die Vielfalt an methodischen Herausforderungen bei der Kompetenzmessung in längsschnittlichen Large-Scale-Assessments sowie die enge Verknüpfung der methodischen Themen mit den verschiedensten inhaltlichen Fragestellungen.

Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (eds.) (2011). Education as a lifelong process – the German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft: Special Issue 14.

Barkow, I., Leopold, T., Raab, M., Schiller, D., Wenzig, K., Blossfeld, H.P. & Rittberger, M. (2011) RemoteNEPS: data dissemination in a collaborative workspace. Zeitschrift für Erziehungswissenschaft, 14, 315-325.

Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., & Carstensen, C. H. (2011). Development of competencies across the life span. Zeitschrift für Erziehungswissenschaft, 14, 67-86.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G131



Steffi Pohl (Freie Universität Berlin), Katinka Hardt (University of Luxembourg), Anna Südkamp (Technische Universität Dortmund), Lena Nusser (Otto-Friedrich-Universität Bamberg), Claus Carstensen (Otto-Friedrich-Universität Bamberg), Sabine Weinert (Otto-Friedrich-Universität Bamberg)

## **Identifikation von testbaren Förderschülern**

Large-scale assessments wie das Nationale Bildungspanel (NEPS, Blossfeld, Roßbach, & von Maurice, 2011) versuchen Förderschüler in die Erhebung und damit die Kompetenzmessung zu integrieren. Die Herausforderung dabei ist Kompetenztests so zu entwickeln, dass diese auch bei Förderschülern zu reliablen und zu Regelschülern vergleichbaren Messungen führen. Bisherige Ergebnisse aus Large-Scale-Assessments (z.B. Abedi, Leon, & Kao, 2008; Cormier, Altmann, Shjyyan, & Thurlow, 2010; Pitoniak, & Royer, 2001) sowie Machbarkeitsstudien im NEPS (Südkamp, Pohl, & Weinert, submitted), die neben Standardtests für Regelschüler auch in Schwierigkeit und Länge angepasste Tests implementiert haben, zeigen, dass dies kein leichtes Unterfangen ist. Häufig treten Probleme mit der Passung des Messmodells und der Vergleichbarkeit der Aufgaben zu den Regelschülern auf.

In diesem Vortrag setzen wir einen anderen Fokus auf die Frage nach der Testbarkeit von Förderschülern. Wir widmen wir uns nicht der Frage danach welche Aufgaben bei Förderschülern zu reliablen und vergleichbaren Messungen führen, sondern bei welchen Förderschülern wir mit den in Large-Scale-Assessments einsetzbaren Tests belastbare Kompetenzmessungen erhalten können. Dazu wollen wir 1) identifizieren für welche Gruppen von Förderschülern man mit Tests für Large-Scale-Assessments adäquate Kompetenzmessungen erhält sowie 2) untersuchen anhand welcher Testbearbeitungsmerkmale zwischen testbaren und nicht testbaren Förderschülern differenziert werden kann.

Dazu wurden Daten aus Machbarkeitsstudien des Nationalen Bildungspanels (Heydrich, Weinert, Nusser, Artelt, & Carstensen, 2013) von N=403 Schülern mit sonderpädagogischem Förderbedarf im Schwerpunkt Lernen der Klassenstufe 9 zur Messung von Lesekompetenz analysiert. Zunächst wurde ein Rasch-Modell auf die Daten angewendet und Itempassung sowie Messinvarianz zur Regelschulstichprobe (N=13933) evaluiert. In einem restringierten Mixed-Rasch-Modell wurde untersucht für wie viele Förderschüler Messinvarianz zur Regelschule vorliegt. Danach wurden weitere Mixed-Modelle spezifiziert, die basierend auf unterschiedlichen Bearbeitungsstrategien Subgruppen identifizieren sollten. Als Bearbeitungsstrategien untersuchen wir unsystematisches Ankreuzen sowie Auslassen und Nichterreichen von Aufgaben. In einem ersten Modell (Rate-Modell) werden zwei Klassen spezifiziert, wobei in einer Klasse das Rasch-Modell gilt, während in der anderen Klasse unsystematisches Ankreuzen modelliert wird. In einem zweiten Modell (Missing-Modell) werden die Informationen zu den fehlenden Werten auf unterschiedlichen Antwortformaten aufgrund von Auslassen sowie aufgrund des Nichterreichens von Aufgaben verwendet und unterschiedliche Patterns von fehlenden Werten über die Klassen zugelassen. Dann wurde evaluiert, ob sich mit Hilfe der Mixed-Modelle Subgruppen identifizieren lassen konnten, für die Reliabilität der Items (Itempassung) und Vergleichbarkeit (Messinvarianz zur Regelschulstichprobe) gilt. Die Subgruppen wurden bezüglich Ihrer Größe und beim Missing-Modell bezüglich der Auslassungsstrategien beurteilt.

Die Analyseergebnisse zeigen, dass nur eine sehr kleine Gruppe von Förderschülern reliabel und zu Regelschülern vergleichbar testbar ist. Es konnte eine kleine Gruppe von Förderschülern identifiziert werden, die unsystematisch ankreuzen. Ein Ausschluss dieser Schüler führt zu einer teilweisen Verbesserung der Itemkennwerte. Es zeigt sich weiterhin, dass unterschiedliche Bearbeitungsstrategien bezüglich ausgelassener Antworten vorliegen. Während eine Subgruppe mit tendenziell besseren Itemkennwerten wenig fehlende Werte aufweist, die konsistent über verschiedene Aufgabenformate auftreten, gibt es eine andere Gruppe, welche viele fehlende Werte aufweist, die Antwortformat-spezifisch sind. Die Ergebnisse legen nahe, dass nicht reliabel testbare Förderschüler zum Teil unsystematisch Antworten ankreuzen oder Aufgaben nicht bearbeiten. Ursachen dafür können mangelnde Motivation, kognitive Überforderung, sowie Probleme mit komplexeren Antwortformaten sein. Konsequenzen der Ergebnisse für die Einbeziehung von Förderschülern in Large-Scale-Assessments werden diskutiert.

- Abedi, J., Leon, S., & Kao, J. (2008). Examining differential item functioning in reading assessments for students with disabilities (CRESST Report 744). Los Angeles, CA: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.
- Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (eds.) (2011). Education as a lifelong process – the German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft: Special Issue 14.
- Cormier, D. C., Altman, J., Shyyan, V., & Thurlow, M. L. (2010). A summary of the research on the effects of test accommodations: 2007-2008 (Technical Report 56). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Heydrich, J., Weinert, S., Nusser, L., Artelt, C. & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: Challenges and approaches within the German National Educational Panel Study (NEPS). Journal for Educational Research Online (JERO), 5, 217-240.
- Pitoniak, M. J., & Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. Review of Educational Research, 71, 53-104.
- Südkamp, A., Pohl, S., Weinert, S. (submitted). Assessing reading comprehension of students with special educational needs - Identification of appropriate testing accommodations.

## **Untersuchung der Tendenz zum Überspringen von Aufgaben in Kompetenztests**

Bei der Bearbeitung von Kompetenztests groß angelegter Bildungsstudien überspringen Teilnehmer häufig einzelne Aufgaben. Diese müssen bei der Skalierung angemessen berücksichtigt werden, wobei Forscher bis dato keinen Konsensus bezüglich einer idealen Handhabung fehlender Werte gefunden haben. In diesem Zusammenhang erscheint es wichtig, die Ursachen für das Überspringen genauer zu betrachten. Ein eingehendes Verständnis darüber könnte einerseits Hinweise über den Umgang mit fehlenden Werten geben und andererseits Testentwicklern Anhaltspunkte darüber liefern, wie eine hohe Antwortrate Seitens der Testteilnehmer zu erreichen ist. Studien haben gezeigt, dass übersprungene Aufgaben personen-, situations- und aufgabenbedingt verursacht sind (Jakwerth, Stancavage, & Reed, 1999; Koretz et al., 1993). Untersuchungen diesbezüglich fanden jedoch überwiegend in der Testkohorte der Schüler und einer begrenzten Anzahl an Kompetenzbereichen statt. Die vorliegende Arbeit widmet sich der Fragestellung, wie sich die Tendenz zum Überspringen einer Aufgabe über verschiedenen Kompetenzdomänen hinweg verhält und inwieweit sie mit der Fähigkeit und anderen Persönlichkeitseigenschaften zusammenhängt. Unterschiede zwischen Altersstufen werden hierbei ermittelt. Grundlage der Analysen bilden Daten des Nationalen Bildungspanels (NEPS; Blossfeld et al., 2011) in der Kohorte der Neuntklässler (N = 15 239), der Studierenden (N = 5 951) sowie der Erwachsenen (N = 7 256), wobei jede Kohorte in mehreren Kompetenzbereichen getestet wurde. Die Tendenz einer Person, ein Item zu überspringen, wurde mithilfe von Modellen der Item Response Theorie pro Kompetenzdomäne erfasst. Latente Korrelationen zwischen diesen Tendenzen in verschiedenen Domänen geben Aufschluss darüber, ob das Auslassverhalten einer Person stabil oder eher domänenspezifisch auftritt. Korrelationen zwischen der Tendenz und der Fähigkeit sowie anderen Persönlichkeitsvariablen informieren über personenbedingte Einflüsse und ob diese sich in den verschiedenen Domänen unterschiedlich stark auf das Auslassverhalten auswirken. Die Analysen finden in allen Alterskohorten statt und ermöglichen somit Aussagen über Unterschiede zwischen verschiedenen Altersstufen. Vorläufige Ergebnisse lassen darauf schließen, dass die Tendenzen zum Überspringen über verschiedene Testdomänen hinweg hoch korrelieren. Die Fähigkeit einer Person beeinflusst vor allem in Mathematik und in höheren Altersstufen ihr Auslassverhalten. Das Wissen über die Stabilität sowie personenbedingte Einwirkungen liefert Erklärungsansätze für das Auftreten von fehlenden Werten in verschiedenen Domänen und Alterskohorten, welche in Bezug auf praktische Implikationen bei der Entwicklung und Skalierung von Kompetenztests diskutiert werden.

Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (eds.) (2011). Education as a lifelong process – the German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft: Special Issue 14.

Koretz, D., Lewis, E., Skewes-Cox, T. & Burstein, L. (1993). Omitted and not-reached items in mathematics in the 1990 National Assessment of Educational Progress (CSE Tech. Rep. No. 357). Los Angeles: Center for Research on Evaluation, Standards and Student Testing, University of California, Los Angeles.

Jakwerth, P. M., Stancavage, F. B. & Reed, E. D. (1999). An investigation of why students do not respond to questions, NAEP Validity Studies, Working Paper Series, American Institutes for Research, Palo Alto, CA.



Maike Krannich (Otto-Friedrich-Universität Bamberg), Claus Carstensen (Otto-Friedrich-Universität Bamberg),  
Steffi Pohl (Freie Universität Berlin)

## **Wie robust sind latente Zusammenhangsschätzungen in der längsschnittlich angelegten large-scale assessment Studie NEPS gegen unterspezifizierte Hintergrundmodelle?**

Das nationale Bildungspanel (NEPS) ist eine großangelegte Längsschnittstudie mit einem Multikohorten-Sequenz-Design (Blossfeld, von Maurice, & Schneider, 2011). Ein wesentliches Ziel des NEPS besteht darin, zusammenhangsanalytische Untersuchungen zwischen Kompetenzen und erklärenden Variablen zu ermöglichen. Auf Individualebene werden hierfür Punktschätzer der individuellen Fähigkeit zur Verfügung gestellt (Pohl & Carstensen, 2012). Bei der Untersuchung von Zusammenhängen auf der Populationsebene bieten jedoch sogenannte „plausible values“ (Mislevy, 1984; Mislevy, 1991) zahlreiche methodologische Vorteile (Beaton, 1987) – ganz zentral die Analyse auf latenter Ebene. Das übliche Vorgehen bei einer Schätzung dieser plausible values besteht darin, möglichst alle erhobenen Kontextvariablen mit in das Modell aufzunehmen. Aufgrund der komplexen Datenlage im NEPS, ist dies jedoch kaum möglich. Die Bereitstellung von plausible values stellt daher für das nationale Bildungspanel eine besondere Herausforderung dar.

Im Vortrag wird der Frage nachgegangen, inwieweit reduzierte (d.h. unterspezifizierte) Hintergrundmodelle bei der Schätzung der plausible values ausreichen, um die untersuchten latenten Zusammenhänge annähernd adäquat abzubilden.

Hierfür wird exemplarisch untersucht, welcher Zusammenhang zwischen den im NEPS erhobenen schülerbezogenen motivationalen und persönlichkeitsbezogenen Konstrukten (hierzu zählen beispielsweise Leistungsmotivation, persönliche Ziele, Interessensorientierungen sowie allgemeines und domänenspezifisches Selbstkonzept) und der Lese- und Mathematikfähigkeit besteht. Um plausible values zu schätzen, werden zwei verschiedene Hintergrundmodelle entwickelt:

Modell 1 stellt ein vollständiges Modell dar, in dem alle erhobenen relevanten schülerbezogenen Variablen mit Hilfe latenter Regressionen in das Hintergrundmodell aufgenommen werden; Modell 2 hingegen stellt ein sparsameres Modell dar, welches mit ausgewählten Kontextvariablen auskommt. In beiden Modellen dient ein Item-Response-Modell dazu, die latente Fähigkeit der Schüler im Bereich Mathematik und Lesen zu schätzen.

In einem zweiten Schritt werden die aus der marginalen posterioren Verteilung geschätzten plausible values als Populationsschätzer (Beaton, 1987) der Mathematik- bzw. Lesekompetenz herangezogen, um mit diesen den latenten Zusammenhang zwischen Kompetenzen und motivationalen und persönlichkeitsbezogenen Variablen zu untersuchen. Im Vortrag wird dargestellt, in welchem Ausmaß die Verwendung des sparsameren Modells 2 zu verzerrten Populationsschätzern und damit auch zu verzerrten Zusammenhangsschätzungen der Kompetenzen mit den erklärenden Variablen führen kann.

Beide Modelle werden mit Daten aus der Startkohorte 3 des NEPS (Fünftklässler an Regel- und Förderschulen sowie Eltern, Lehrer/innen und Schulleiter/innen als Kontextpersonen) (Pohl, Haberkorn, Hardt & Wiegand, 2012) und anhand einer Simulationsstudie getestet. Im Rahmen des Vortrags werden erste Ergebnisse vorgestellt und diskutiert.

Allen, N. L., Donoghue, J. R. & Schoeps, T. L. (2001). The NAEP 1998 Technical Report (NCES 2001–509). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Beaton, A.E. (1987). The NAEP 1983-1984 Technical Report. Princeton: Educational Testing Service.

Blossfeld, H.-P., von Maurice, J. & Schneider, T. (2011). Grundidee, Konzeption und Design des Nationalen Bildungspanels für Deutschland (NEPS Working Paper No. 1). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.

Mislevy, R. J. (1984). Estimating latent distributions. Psychometrika, 49 (3), 359-381.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. Psychometrika, 56(2), 177–196.

Pohl, S. & Carstensen, C. H. (2012). NEPS Technical Report – Scaling the Data of the Competence Tests (NEPS Working Paper No. 14). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.

Pohl, S., Haberkorn, K., Hardt, K., & Wiegand, E. (2012). NEPS Technical Report for Reading – Scaling Results of Starting Cohort 3 in Fifth Grade (NEPS Working Paper No. 15). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.

## **Vertikales Linking von Kompetenzdaten über die Lebensspanne**

Bildungsverläufe und Kompetenzentwicklungen spielen durch die wachsende Bedeutung von Bildung und Ausbildung in allen Phasen des Lebenslaufs eine zunehmend größere Rolle in modernen Industrieländern (Blossfeld, von Maurice, & Schneider, 2011). In diesem Zusammenhang ist es Ziel des Nationalen Bildungspanels (NEPS), bildungsrelevante Kompetenzen über die Lebensspanne hinweg zu erfassen und wichtige Voraussetzungen und Bedingungen für deren Erwerb zu analysieren. Eine besondere Herausforderung besteht dabei darin, domänenspezifische Kompetenzen über die Lebensspanne kohärent zu messen. Denn wie von Davier, Carstensen und von Davier (2008) beschreiben, ist eine wichtige Voraussetzung für den Vergleich verschiedener Altersstufen, dass die untersuchte Kompetenz über verschiedene Messzeitpunkte hinweg vergleichbar ist. Mit Blick darauf, Kompetenzentwicklungen über die Zeit verfolgen zu können, wurde im Rahmen der NEPS-Studie deshalb der zentralen Frage nachgegangen, inwiefern Kompetenzmessungen einer Domäne, welche in verschiedenen Altersstufen stattfinden, vergleichbare Konstrukte abbilden.

Für die Beantwortung der vorliegenden Frage wurden Daten der basalen Kompetenzbereiche Lesen und Mathe herangezogen. Da zur Erfassung von Kompetenzen für eine Domäne im NEPS in verschiedenen Altersstufen unterschiedliche Tests administriert werden, bearbeitete in sogenannten Link-Studien eine zusätzliche Stichprobe jeweils die Tests zweier unterschiedlicher Altersstufen. Die Messinvarianz der Tests wurde zwischen Schulkohorten und zwischen Schul- und Erwachsenenstichprobe geprüft. Unter Verwendung verschiedener Item-Response-Modelle wurden differentielle Veränderungen auf Itemebene und Zusammenhänge der Messinstrumente, die für unterschiedliche Altersstufen konstruiert wurden, untersucht.

Erste Ergebnisse weisen darauf hin, dass für die beiden Domänen Lesen und Mathe zwischen der Schulstichprobe und der Erwachsenenstichprobe teilweise starke Verletzungen der Messinvarianz vorliegen. Für das Schulalter scheint die Messinvarianz zwischen Altersstufen weniger gefährdet zu sein. Tiefergehende Analysen legen nahe, dass verschiedene Altersstufen unterschiedliche Strategien bei der Testbearbeitung benutzen könnten.

Die Befunde verdeutlichen die Schwierigkeiten bei der Verfolgung der Kompetenzentwicklung im Lebenslauf und regen zur Diskussion an, wie eine vergleichbare Messung einer Kompetenz über die Lebensspanne gelingen kann und wie Veränderungen der Kompetenz adäquat modelliert werden können.

Blossfeld, H.-P., von Maurice, J. & Schneider, T. (2011). Grundidee, Konzeption und Design des Nationalen Bildungspanels für Deutschland (NEPS Working Paper No. 1). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.

Von Davier, A. A., Carstensen, C. H. & Von Davier, M. (2008). Linking Competencies in Horizontal, Vertical and Longitudinal Settings and Measuring Growth. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of Competencies in Educational Contexts* (S. 121–150). New York: Hogrefe & Huber.

Chair(s): Linda Wirthwein (TU Dortmund), Ricarda Steinmayr (TU Dortmund)

Diskutant/in: Jörn Sparfeldt (Universität des Saarlandes)

Zielorientierungen gelten als zentrale motivationale Determinanten des Lern- und Leistungsverhaltens. Dabei werden mittlerweile sowohl Antezedenzen als auch Auswirkungen von Zielorientierungen untersucht. Ziel des Symposiums ist es, aktuelle Befunde der Zielorientierungsforschung an unterschiedlichen Personengruppen (Schülerinnen und Schüler, Lehrkräfte) zu präsentieren sowie den Einfluss unterschiedlicher Kontextvariablen (z.B. Familie, Migrationshintergrund) hinsichtlich der Entwicklung von Zielorientierungen bzw. hinsichtlich der Korrelate von Zielorientierungen zu thematisieren.

Die Beiträge berücksichtigen unterschiedliche Kontexte bzw. theoretische Perspektiven (Soziologie, Psychologie, Erziehungswissenschaft) und greifen sowohl auf quer- als auch auf längsschnittliche Forschungsdesigns zurück. Während sich die ersten beiden Beiträge des Symposiums mit etwaigen Antezedenzen von Zielorientierungen befassen, thematisieren die letzten beiden Beiträge die (reziproken) Zusammenhänge zwischen Zielorientierungen und Schulleistung sowie selbstbezogenen Kognitionen und Interesse. Ludwig und Watermann befassen sich in ihrer Studie mit dem Einfluss familialer sozialer und kultureller Praxen und elterlichen schulbezogenen Involvements auf die Veränderungen von Zielorientierungen bei Schülerinnen und Schülern. Dabei greifen die Autoren neben soziologischen Theorien u.a. auf die Selbstbestimmungstheorie von Deci und Ryan (1985) zurück, die auch als theoretische Grundlage für die Untersuchung von Janke, Nitsche und Dickhäuser herangezogen wird. Diese Studie befasst sich mit der Vorhersage des berufsbezogenen Erlebens und Verhaltens von Lehrkräften und verweist auf ein integratives Modell zu Antezedenzen von Zielorientierungen im Lehrerberuf. Wirthwein und Steinmayr befassen sich mit den reziproken Zusammenhängen zwischen Zielorientierungen und Schulnoten bei Schülerinnen und Schülern und untersuchen damit ein im Bereich der Pädagogischen Psychologie zwar bereits häufig angesprochenes, jedoch bislang kaum an Längsschnittdaten untersuchtes Thema. Die Bedeutung der Kontextvariablen „Migrationshintergrund“ wird in dem Beitrag von Gebauer, McElvany, Ferdinand, Bos, Köller und Schöber analysiert. Genauer befassen sich die Autoren damit, welche Bedeutung Lernziele für Selbstwirksamkeitsüberzeugungen sowie Interesse haben und inwiefern sich differentielle Zusammenhangsmuster in Abhängigkeit vom Migrationshintergrund ergeben.

Jörn Sparfeldt stellt im Anschluss an die Vorträge eine Gesamtdiskussion der Beiträge vor.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G135

## ***Konsequenzen elterlichen schulbezogenen Involvements für die Entwicklung von Zielorientierungen im frühen Jugendalter***

Als wichtige Determinanten akademischer Motivation im frühen Jugendalter gelten die schulische Lernumwelt sowie die Rolle des Elternhauses. Bei der Analyse dieser Determinanten wurde die Rolle des Elternhauses - anders als die Rolle der schulischen Lernumwelt - jedoch häufig vernachlässigt. Nur wenige Studien untersuchen etwa den Zusammenhang von Merkmalen der Familie und Zielorientierungen und die bisher bestehenden Untersuchungen verfolgen mehrheitlich ein querschnittliches Design. Die bestehenden Studien zeigen jedoch, dass Merkmale des Elternhauses für die Analyse von Zielorientierungen eine wichtige Rolle spielen (Duchesne & Ratelle, 2010; Régner et al., 2009; Gonzalez et al., 2002; Gurland & Grolnick, 2005; Marjoribanks & Mboya, 2001). Auf Basis eines längsschnittlichen Designs soll in der vorliegenden Studie unter Rückgriff auf die Selbstbestimmungstheorie von Deci und Ryan (1985) der Einfluss familialer sozialer und kultureller Praxen sowie elterlichen schulbezogenen Involvements auf die Veränderung von Zielorientierungen geprüft werden. Es wird angenommen, dass schulbezogenes Involvement der Eltern die Befriedigung psychologischer Grundbedürfnisse unterstützen und damit motivationale Ressourcen der Schüler begünstigen kann (Deci & Ryan, 1985). Weiterhin wird in Anlehnung an die theoretischen Arbeiten Bourdieus (1977) und Colemans (1988) erwartet, dass die familialen sozialen und kulturellen Praxen über das elterliche schulbezogene Involvement vermittelt sind. Im Rahmen der TIMSS-Übergangsstudie wurden N=1783 Viertklässler am Ende der Grundschule hinsichtlich der in der Familie vorherrschenden sozialen und kulturellen Aktivitäten sowie hinsichtlich verschiedener schulbezogener Verhaltensmuster ihrer Eltern befragt. Die familialen sozialen Praxen wurden als gemeinsame häusliche Aktivitäten der Eltern mit ihren Kindern erfasst, die familialen kulturellen Praxen als gemeinsame Aktivitäten der Eltern mit ihren Kindern im kulturellen Bereich. Elterliches schulbezogenes Involvement ließ sich den Dimensionen Autonomieunterstützung und Kontrolle zuordnen. Autonomieunterstützung beschreibt ein elterliches Engagement zur Förderung selbstinitiiert lernbezogener Aktivitäten, wohingegen mit Kontrolle ein elterliches Verhalten beschrieben wird, das darauf abzielt, das Leistungsverhalten und -ergebnis des Kindes zu kontrollieren, zu belohnen oder zu sanktionieren. Die beiden Dimensionen gehen auf eine Konzeptualisierung elterlichen Involvements von Lorenz und Wild (2007) zurück. Die Zielorientierungen der Schüler wurden schließlich als Lern-, Annäherungsleistungs- und Vermeidungsleistungsziele (Ames & Archner, 1988; Elliot & Church, 1997; Elliot & Dweck, 2005) sowohl am Ende der 4. Klasse (T1), als auch zu Beginn der 7. Klasse (T2) erfasst. Vor dem Hintergrund der theoretischen Annahmen sowie der bestehenden querschnittlichen Befunde zum Einfluss elterlichen schulbezogenen Involvements auf Zielorientierungen wird erwartet, dass sich elterliche Autonomieunterstützung am Ende der 4. Klasse positiv auf die Veränderung der Lernzielorientierung auswirkt, wohingegen sich elterliche Kontrolle zu T1 positiv auf die Veränderung der Leistungszielorientierung auswirken sollte. Multiple Regressionsanalysen, in denen für die Zielorientierung sowie die schulische Leistung zu T1 und für Merkmale der sozialen Herkunft (sozioökonomischer Status) kontrolliert wurde, zeigten einen längsschnittlichen positiven Effekt von Autonomieunterstützung auf die Veränderung der Lernzielorientierung sowie einen positiven Effekt von Kontrolle auf die Veränderung der Vermeidungsleistungszielorientierung. Anders als theoretisch angenommen zeigten sich keine über das elterliche schulbezogene Involvement vermittelten Effekte der sozialen und kulturellen Praxen. Die bestehenden querschnittlichen Befunde konnten weitgehend repliziert werden und deuten auf unterschiedliche, von den familialen sozialen und kulturellen Praxen unabhängige Wirkungsweisen elterlichen schulbezogenen Involvements hin. Weiterhin zeigte sich, dass Autonomie Lernziele und Kontrolle Leistungsziele fördert

- Ames, C., & Archner, J. (1988). Achievement in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80, 260-267.
- Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: University Press.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95-120.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Duchesne, S., & Ratelle, C. (2010). Parental behaviours and adolescents achievement goals at the beginning of middle school: Emotional problems as potential mediators. *Journal of Educational Psychology*, 2, 497-507.
- Gonzalez, A., Doan Holbein, M. F., & Quilter, S. (2002). High school students' goal orientations and their relationship to perceived parenting styles. *Contemporary Educational Psychology*, 27, 450-470.
- Gurland, S. T., & Grolnick, W. S. (2005). Perceived threat, controlling parenting, and children's achievement orientations. *Motivation and Emotion*, 29, 103-121.
- Lorenz, F. & Wild, E. (2007). Parental involvement in schooling - results concerning its structure and impact on students' motivation. In M. Prenzel & L. Allolio-Näcke (Eds.), *Studies on the educational quality of schools. The final report on the DFG Priority Program (299-316)*. Münster: Waxmann.
- Régner, I., Loose, F., & Dumas, F. (2009). Students' perceptions of parental and teacher academic involvement: Consequences on achievement goals. *European Journal of Psychology of Education*, 24 (2), 263-277.
- Elliot, A. J., & Church, M. A. (1997). A Hierarchical model of approach and achievement motivation. *Journal of Personality and Social Psychology*, 72, 218-232.
- Elliot, A. J., & Dweck, C. S. (2005). *Handbook of competence and motivation*. New York, NY, Guilford Press.
- Marjoribanks, K., & Mboya, M. (2001). Familial capital, goal orientations and south african adolescents' self-concept: A moderation-mediation model. *Educational Psychology*, 21, 3, 333-350.



## **Zur Bedeutung selbstbestimmter Arbeit für Lehrerziele - Ein integratives Modell zu Antezedenzen von Zielorientierungen im Lehrerberuf**

Das Konzept der beruflichen Zielorientierung geht davon aus, dass Lehrkräfte das eigene berufliche Handeln mehr oder weniger lern- und leistungsthematisch wahrnehmen und im Beruf unterschiedlich stark das Ziel verfolgen, im Beruf hinzuzulernen (Lernzielorientierung), sich als gute Lehrkraft zu erweisen (Annäherungsleistungszielorientierung) und sich nicht als schlechte Lehrkraft zu erweisen (Vermeidungsleistungszielorientierung). Berufliche Zielorientierungen von Lehrkräften haben sich in der Vergangenheit als fruchtbares Konzept zur Vorhersage des berufsbezogenen Erlebens und Verhaltens von Lehrkräften erwiesen. Der Fokus vergangener Forschungsarbeiten lag dabei vor allem auf den Konsequenzen von Zielorientierungen, beispielsweise für selbstreguliertes Lernen im Lehrerberuf (z.B. Nutzung von Fortbildungsgelegenheiten) oder das berufliche Belastungserleben. Auf Grund der nachgewiesenen Bedeutsamkeit beruflicher Zielorientierungen von Lehrkräften für diese Variablen ist es wichtig, auch mögliche Antezedenzen von Zielorientierungen genauer in den Blick zu nehmen. Dabei sind besonders solche Antezedenzen interessant, welche im Arbeitsumfeld der Lehrkräfte angelegt sind und sich motivationsförderlich beeinflussen lassen. Um solche Faktoren aufzudecken, integrieren wir Erkenntnisse aus der Selbstbestimmungstheorie und dem Zielorientierungsansatz. Unter Rückbezug auf das TARGET-Modell (Ames, 1992) und Arbeiten zu intrinsischen Zielen (Vansteenkiste, Lens, & Deci, 2006) gehen wir davon aus, dass die Befriedigung der psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit im Arbeitsalltag von Lehrkräften mit einer erhöhten beruflichen Lernzielorientierung assoziiert ist. Sowohl das Ausmaß der beruflichen Lernzielorientierung, als auch die Befriedigung sozialer Grundbedürfnisse am Arbeitsplatz sollten entsprechend bisheriger Literatur positiv mit dem Ausmaß der intrinsischen Berufsmotivation korrespondieren. In Bezug auf Leistungsziele lässt sich annehmen, dass diese vor allem in kompetitiven Kollegien mit geringem kollegialem Zusammenhalt verfolgt werden. Entsprechend sollte das Ausmaß an erlebter sozialer Eingebundenheit die beruflichen Leistungszielorientierungen negativ vorhersagen. Leistungszielorientierungen sollten aber auch mit dem beruflichen Kompetenzerleben in Zusammenhang stehen. Dabei ist zu beachten, dass Leistungszielorientierungen entsprechend dem trichotomen Modell der Zielorientierungen (Elliot & Harackiewicz, 1996) in eine Annäherungs- und Vermeidungskomponente unterteilt werden. Nach Elliot und Church (1997) ist ein direkter Zusammenhang zwischen der eigenen Kompetenzeinschätzung und der Einnahme eines Annäherungs- respektive Vermeidungsfokus zu erwarten: Menschen nehmen vor allem dann einen Annäherungsfokus ein, wenn Sie glauben gute Leistung demonstrieren zu können. Umgekehrt wird ein Vermeidungsfokus wahrscheinlicher, wenn Menschen der Auffassung sind, dass Sie unter Umständen nicht fähig sind, den Anforderungen der Leistungssituation zu entsprechen. Wir gehen entsprechend dieser Überlegung davon aus, dass die Erfüllung des Bedürfnisses nach Kompetenzerleben im Arbeitsalltag von Lehrkräften differenziell prädiktiv dafür ist, ob Lehrkräfte eine Annäherungs- oder Vermeidungsleistungszielorientierung einnehmen. Ergänzend legt das hierarchische Modell der Zielorientierungen (Elliot & Church, 1997) nahe, dass das Ausmaß beruflicher Vermeidungsleistungszielorientierung negativ mit intrinsischer Berufsmotivation assoziiert ist. Um das aus diesen Theoriesträngen resultierende integrative Modell zu testen, wurde eine Stichprobe von 334 Lehrkräften erhoben. Es handelte sich um eine eher heterogene Stichprobe von Lehrkräften aus verschiedenen Schultypen und mit unterschiedlichen Fachkombinationen. Die Erhebung erfolgte in Form einer Onlinebefragung, in deren Verlauf die Lehrkräfte zu den modellrelevanten Variablen mit Hilfe etablierter Instrumente befragt wurden. Das postulierte Modell wurde auf Basis linearer Strukturgleichungsmodelle getestet. Der Modellfit wies auf eine gute Passung an die erhobenen Daten hin. Alle postulierten Pfade waren signifikant und wiesen in hypothesenkonformer Richtung; lediglich der angenommene Zusammenhang zwischen Vermeidungsleistungszielorientierung und intrinsischer Motivation ließ sich in unserer Stichprobe nicht nachweisen. Obwohl die beschriebenen Analysen lediglich auf Querschnittsdaten basiert, lassen sich aus dem integrativen Modell erste vorläufige Implikationen zur Gestaltung des Arbeitsumfeldes von Lehrkräften ableiten. So scheint es erstrebenswert, die Arbeitsbedingungen von Lehrkräften an deren psychologische Grundbedürfnisse anzupassen (z.B. durch Schaffen von Freiräumen bei der Unterrichtsgestaltung, Feedbackroutinen oder Teamentwicklungsmaßnahmen). Weitere Implikationen, sowie offene theoretische Fragen werden diskutiert.



## ***Henne oder Ei? - Zum reziproken Zusammenhang von Zielorientierungen und Schulnoten***

Zielorientierungen werden häufig als zentrale motivationale Determinanten schulischen Lern- und Leistungsverhaltens untersucht (vgl. Wolters, 2004). Meta-Analysen (z.B. Wirthwein, Sparfeldt, Piquart, Wegerer, & Steinmayr, 2013; Hulleman, Schrager, Bodmann, & Harackiewicz, 2010) konnten insgesamt geringe positive Zusammenhänge zwischen unterschiedlichen Leistungsmaßen und Lernzielen (Ziel, eigene Fähigkeiten zu erweitern) sowie Annäherungs-Leistungszielen (Ziel, anderen gegenüber hohe Kompetenzen zu zeigen) aufzeigen; geringe negative Zusammenhänge ergeben sich für Vermeidungs-Leistungsziele (Ziel, Schwächen vor anderen zu verbergen) sowie Arbeitsvermeidung (Ziel, möglichst wenig Arbeit zu investieren). Frühe theoretische Überlegungen von Dweck (1986) oder Dweck und Leggett (1988) verweisen auf Zielorientierungen als Determinanten von Leistung. Vor allem das Verfolgen einer Lernzielorientierung wurde als förderlich für günstiges Lern- und Leistungsverhalten erachtet. Analog zur Entwicklung akademischer Fähigkeitsselbstkonzepte (vgl. Marsh & Craven, 2006) ist es jedoch auch denkbar, dass die jeweilige Leistungsausprägung Einfluss auf die Entwicklung von Zielorientierungen hat bzw. dass sich reziproke Beziehungen zwischen Zielorientierungen und Leistung ergeben. In Abhängigkeit von der jeweiligen Zielorientierung könnten sich in diesem Zusammenhang differentielle Befunde ergeben: Während Lernziele vermutlich vor allem Auswirkungen auf nachfolgendes Lern- und Leistungsverhalten haben, könnte die Tendenz, eigene Fähigkeiten demonstrieren zu wollen (Annäherungs-Leistungsziele) vielmehr von der vorherigen Leistungsausprägung abhängig sein. Senko, Hulleman und Harackiewicz (2011) gehen darüber hinaus davon aus, dass eher normativ-orientierte Annäherungs-Leistungsziele (besser sein wollen als andere) aufgrund ihres herausfordernden Charakters im Vergleich zu kompetenz-orientierten Annäherungs-Leistungszielen (Kompetenz demonstrieren) positive Auswirkungen auf die zukünftige Leistung haben. Es ist jedoch auch durchaus denkbar, dass erst gute Leistungen zum Ziel, die eigenen Kompetenzen zeigen zu können, führen. Bisher liegen jedoch keine längsschnittlichen Untersuchungen zum reziproken Zusammenhang von Zielorientierungen und Schulnoten vor, bei denen sowohl die Veränderung in den Schulnoten als auch die Veränderung in den Zielorientierungen untersucht wurden. Ziel der vorliegenden Studie ist die Untersuchung der Wirkrichtung des Zusammenhangs von Zielorientierungen und Schulnoten.

Untersucht wurden N = 343 Schülerinnen und Schüler (n = 206 weiblich) dreier Gymnasien, die an zwei Messzeitpunkten untersucht wurden (erster Messzeitpunkt: Zweites Halbjahr 11. Klasse; zweiter Messzeitpunkt: Zweites Halbjahr 12. Klasse). Die Zielorientierungen wurden sowohl schulfachübergreifend als auch spezifisch für die Schulfächer Mathematik und Deutsch erfasst; des Weiteren wurden Schulnoten erhoben. Mittels SEM cross lagged-Modellen wurden für jede fachspezifische Zielorientierung getrennt die reziproken Beziehungen zu Schulnoten analysiert. Entsprechende Modelle bieten sich an, um die Veränderung einer Variablen über die Zeit auf eine andere Variable zurückzuführen. Der Modellfit für die jeweiligen Modelle war durchweg zufriedenstellend. Es ergaben sich Hinweise für reziproke Beziehungen zwischen Lernzielen und Schulleistung. Für normativ-orientierte Annäherungs-Leistungsziele sagte jedoch lediglich die vorherige Schulleistung die Veränderung dieser Zielorientierung vorher. Für die übrigen Zielorientierungen waren keine statistisch signifikanten Pfade zu verzeichnen. Basierend auf diesen Analysen lassen sich relevante Implikationen für die Gestaltung von Maßnahmen zur Förderung der Motivation in der Schule ableiten. Vor allem durch die differenzierte Betrachtung von Annäherungs-Leistungszielen können differentielle Empfehlungen gegeben werden.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, 136, 422-449.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. *Perspectives on Psychological Science*, 1, 133-163.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads. Old controversies, current challenges, and new directions. *Educational Psychologist*, 46, 26-47.

Wirthwein, L., Sparfeldt, J.R., Piquart, M., Wegerer, J., & Steinmayr, R. (2013). Achievement goals and academic achievement: A closer look at moderating factors. *Educational Research Review*, 10, 66-89.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict student's motivation, cognition, and achievement. *Journal of Educational Psychology*, 96, 236-250.

Miriam M. Gebauer (TU Dortmund), Nele McElvany (TU Dortmund), Hanna D. Ferdinand (TU Dortmund), Wilfried Bos (TU Dortmund), Olaf Köller (Christian-Albrechts-Universität zu Kiel), Christian Schöber (Christian-Albrechts-Universität zu Kiel)

## ***Bedeutung von Zielorientierung für fachbezogene Selbstwirksamkeit und Interesse unter Berücksichtigung des Migrationshintergrundes von Schülerinnen und Schülern***

Theoretische Motivationsmodelle postulieren einen Zusammenhang zwischen Zielorientierung, selbstbezogenen Kognitionen und dem Interesse von Schülerinnen und Schülern, welches wiederum von Bedeutung für die schulische Performanz ist (z.B. Wigfield & Eccles, 2000). Empirische Evidenzen, welche die Bedeutung von selbstbezogenen Kognitionen wie Selbstwirksamkeitsüberzeugungen - das Vertrauen in die eigene Handlungskompetenz, welche die Planung, Organisation und Handlungsdurchführung beeinflusst (Bandura, 1997) - oder Interesse - dem subjektiven Bezug zwischen Person und Gegenstand (Schiefele, 1996) - für Lernen und Leistungen von Schülerinnen und Schülern belegen, liegen umfangreich vor (z.B. Pajares, 1996; Schraw & Lehman, 2001). Forschungen, die die Selbstwirksamkeitsüberzeugung und das Interesse gleichzeitig betrachten, liegen hingegen nur in geringer Zahl vor (z.B. Niemivirta & Tapola, 2007) und untersuchen in erster Linie die Bedeutung dieser Konstrukte für Leistungen der Schülerinnen und Schüler. Studien, welche die Bedeutung der Zielorientierung für selbstbezogene Kognitionen untersuchen, nutzen eher das Fähigkeits-selbstkonzept als die Selbstwirksamkeitsüberzeugung (z.B. Spinath & Steinmayr, 2012). Weitgehend unbeachtet blieb in den bislang vorliegenden Studien der Migrationshintergrund der Schülerinnen und Schüler, wobei zu vermuten ist, dass dieser von Bedeutung für die motivationale Dynamik ist (Klassen, 2004; Stanat & Christensen, 2006). Vor diesem Hintergrund beschäftigt sich der Beitrag mit der Bedeutung der Lernzielorientierung für die fachbezogenen Selbstwirksamkeitsüberzeugungen und die fachspezifischen Interessen bei Schülerinnen und Schülern der siebten Jahrgangsstufe unter Berücksichtigung des Migrationshintergrunds. Die Fragestellungen dieses Forschungsvorhabens lauten wie folgt:

1. Ist die Lernzielorientierung von Bedeutung für die fachspezifische Selbstwirksamkeitsüberzeugung und diese wiederum für das fachspezifische Interesse?
2. Gibt es indirekte Zusammenhänge von der Lernzielorientierung über die fachbezogene Selbstwirksamkeitsüberzeugung für das fachspezifische Interesse?
3. Können Unterschiede bei den Zusammenhangsmustern zwischen Kindern mit und ohne Migrationshintergrund festgestellt werden?

Die Fragestellungen wurden mit Daten des ersten Messzeitpunkts der BMBF-geförderten Studie „Se-Mig“ analysiert. Mit Daten von N = 1591 Schülerinnen und Schülern in der gesamten Stichprobe (davon n = 599 mit Migrationshintergrund und n = 992 ohne Migrationshintergrund) wurden anhand des Programms Mplus (Muthén & Muthén, 1998-2012) Strukturgleichungsmodelle mit latenten Variablen und unter Berücksichtigung indirekter Effekte gerechnet. Zur differenziellen Betrachtung der Gruppen mit und ohne Migrationshintergrund wurden Mehrgruppenmodelle gerechnet.

Die Ergebnisse der Analysen weisen darauf hin, dass die Lernzielorientierung die fachspezifische Selbstwirksamkeitsüberzeugung und die fachspezifische Selbstwirksamkeitsüberzeugung das fachbezogene Interesse vorhersagt. Bei der differenziellen Betrachtung der Gruppen mit und ohne Migrationshintergrunds zeigen sich nur kleine Unterschiede: Bei der Gruppe der Jugendlichen ohne Migrationshintergrund ist der Zusammenhang zwischen der Lernzielorientierung und der mathebezogenen Selbstwirksamkeitsüberzeugung im Vergleich zu den Zusammenhängen der Lernzielorientierung mit der lesebezogenen Selbstwirksamkeitsüberzeugung oder der bildungsbezogenen Selbstwirksamkeitsüberzeugung am geringsten. Bei der Gruppe der Jugendlichen mit Migrationshintergrund sind bei diesen Zusammenhängen zwischen den Konstrukten keine bedeutsamen Unterschiede festzustellen. Bei der Gruppe der Schülerinnen und Schüler ohne Migrationshintergrund hat der Zusammenhang zwischen der mathebezogenen Selbstwirksamkeitsüberzeugung und dem Mathematikinteresse eine sehr hohe Ausprägung im Vergleich zu den Zusammenhängen zwischen der lesebezogenen Selbstwirksamkeitsüberzeugung und dem Leseinteresse sowie der bildungsbezogenen Selbstwirksamkeitsüberzeugung und dem bildungsbezogenen Interesse. Dies kann bei der Gruppe der Jugendlichen mit Migrationshintergrund ebenfalls beobachtet werden.

Die Ergebnisse unterstützen die Annahme, dass die Lernzielorientierung von entscheidender Bedeutung für selbstbezogene Kognitionen und Interesse ist. Darüber hinaus wurden differentielle Zusammenhangsmuster in Abhängigkeit vom Migrationshintergrund festgestellt, deren Implikationen für die weitere Forschung und Praxis diskutiert werden.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Klassen, R. M. (2004). Optimism and realism: A review of self-efficacy from a cross-cultural perspective. *International Journal of Psychology*, 39, 205-230.

Muthén, L.K. & Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA

Niemivirta, M. & Tapola, A. (2007). Self-efficacy, interest, and task performance: Within-task changes, mutual relationships, and predictive effects. *Zeitschrift für Pädagogische Psychologie*, 21, 241-250.

Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*, 66, 543-578.

Spinath, B. & Steinmayr, R. (2012). The roles of competence beliefs and goal orientations for change in intrinsic motivation. *Journal of Educational Psychology*, 104, 1135-1148.

Stanat, P. & Christensen, G. S. (2006). Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003. Paris: OECD.

Schiefele, U. (1996). Motivation und Lernen mit Texten. Göttingen: Hogrefe.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13, 23-52.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81.

## **Kompetenzorientierter Mathematikunterricht am Gymnasium und sein Zusammenhang mit der Leistungsentwicklung in der Sekundarstufe I**

Chair(s): Holger Gärtner (ISQ, Freie Universität Berlin)

Seit der Einführung von Bildungsstandards durch die KMK für den Hauptschulabschluss und den Mittleren Schulabschluss (KMK 2004; 2005) wird die Gestaltung und Implementation kompetenzorientierten Unterrichts in Deutschland intensiv diskutiert (Köller, 2008).

Kompetenzorientierung wird hierbei zum einen als Perspektivwechsel bei der Unterrichtsplanung und -Gestaltung beschrieben: „Die wichtigste Frage ist nicht „Was haben wir durchgenommen?“, sondern „Welche Vorstellungen, Fähigkeiten und Einstellungen sind entwickelt worden?“ (Blum, Drüke-Noe, Hartung & Köller, 2006, S.15). Zum anderen wird versucht, Kompetenzorientierung als Muster bestimmter Aspekte von Unterrichtsqualität zu beschreiben (Kunter & Voss, 2011). Empirische Untersuchungen zum Zusammenhang zwischen Ausgestaltung kompetenzorientierten Unterrichts und dem Erreichen der Bildungsstandards liegen bislang jedoch kaum vor.

In der vorliegenden Studie wird daher folgenden Forschungsfragen nachgegangen: 1) In welchem Ausmaß werden Aspekte kompetenzorientierten Mathematikunterrichts in der zehnten Klasse an Gymnasien wahrgenommen? 2) Passen Lehrkräfte ihren Unterricht hinsichtlich von Aspekten kompetenzorientierten Unterrichts an die Eingangsvoraussetzungen der Schülerinnen und Schüler an? 3) Lässt sich ein Zusammenhang zwischen der Ausgestaltung kompetenzorientierten Unterrichts und der Leistungsentwicklung der Schülerinnen und Schüler nachweisen?

Es handelt es sich um eine Längsschnittstudie mit insgesamt vier Messzeitpunkten. Von der achten (VERA8) bis zur zehnten Klasse (MSA) wurden je Messzeitpunkt ca. 1.200 Schülerinnen und Schüler aus 45 Klassen Berliner Gymnasien zur Entwicklung ihrer mathematischen Kompetenz untersucht. Eingesetzt wurden hierfür jeweils VERA-Mathematiktesthefte bzw. Zusammenstellungen mehrerer Testhefte. In der zehnten Klasse wurden zudem die Mathematikergebnisse im Mittleren Schulabschluss (MSA) sowie die Wahrnehmung des Mathematikunterrichts aus Schülersicht erhoben. Die Operationalisierung kompetenzorientierten Unterrichts wurde hierbei direkt aus den Bildungsstandards für die allgemeinen mathematischen Kompetenzen (u.a. Mathematisch Kommunizieren, Problemlösen etc.) abgeleitet (Beispielitem Mathematisches Kommunizieren: Im Unterricht sollen wir unsere Lösungswege anderen Schülerinnen und Schülern verständlich erklären). Anhand von Mehrebenenmodellen wurde zum einen der Zusammenhang zwischen den Eingangsvoraussetzungen der Klasse und dem Ausmaß kompetenzorientierten Unterrichts untersucht. Zum anderen wurde in einem Mediationsmodell die Leistung im MSA sowohl durch die Vorjahresleistung als auch über die Ausgestaltung der Kompetenzorientierung vorhergesagt.

Deskriptive Ergebnisse verdeutlichen eine hohe Variabilität der sechs Aspekte kompetenzorientierten Mathematikunterrichts aus Sicht der Schülerinnen und Schüler ( $.16 < ICC1 < .28$ ). Trotz erheblicher Unterschiede in den Vorjahresleistungen ( $ICC1 = .31$ ) passen Lehrkräfte diese Aspekte ihres Unterrichts nicht an die Eingangsvoraussetzungen ihrer Schüler an. Da die Leistungsentwicklung der Klassen nahezu parallel verläuft, können nur geringe Einflüsse der wahrgenommenen Kompetenzorientierung auf die MSA-Ergebnisse nachgewiesen werden. Gründe für die nahezu parallele Entwicklung der untersuchten Klassen trotz unterschiedlicher Realisierung von Lerngelegenheiten werden diskutiert.

Blum, W., Drüke-Noe, K., Hartung, R., & Köller, O. (2006). Bildungsstandards Mathematik konkret: Sekundarstufe I: Aufgabenbeispiele, Unterrichtsideen und Fortbildungsmöglichkeiten. Berlin: Cornelsen.

Köller, O. (2008). Bildungsstandards in Deutschland: Implikationen für die Qualitätssicherung und Unterrichtsqualität. Zeitschrift für Erziehungswissenschaft, 10(9), 47-59.

Kunter, M., & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften (S. 85-115). Münster: Waxmann.

## Argumentieren Lernen in der Mathematik: Die Bedeutung der Transaktivität im Kollaborativen Lernprozess

Theoretischer Hintergrund: Obwohl Argumentationsfertigkeiten beim mathematischen Beweisen unverzichtbar sind, haben Studierende häufig Probleme, sich an einem argumentativen Diskurs im mathematischen Kontext zu beteiligen. Für die Förderung von Argumentationsfertigkeiten haben sich Argumentationsskripts als hilfreich erwiesen, die Lernende bei der Strukturierung argumentativer Dialoge unterstützen (Noroozi, Weinberger, Biemans, Mulder & Chizari, 2013). Zur Förderung mathematischen Beweisens werden auch heuristische Lösungsbeispiele eingesetzt, die inhaltliche Unterstützung zu mathematischen Beweisprozesse anbieten (Reiss & Renkl, 2002). Allerdings ist bislang unklar, welche kollaborativen Lernprozesse durch diese Unterstützungsmaßnahmen stimuliert werden und in welcher Beziehung diese mit einer Verbesserung der Argumentationsfertigkeiten stehen. In dieser Studie werden sogenannte transaktive Aktivitäten (Teasley, 1997) als vermittelnde Prozesse in den Blick genommen. Transaktive Aktivitäten zeichnen sich dadurch aus, dass im Gegensatz zu konstruktiven Aktivitäten die Lernenden sich auf Beiträge der Lernpartner beziehen, diese also etwa widerlegen oder ergänzen (Chi, 2009). Es ist jedoch nicht geklärt, inwiefern sich transaktive Aktivitäten mit argumentativem Charakter (transaktive Argumentation, z.B. Kritisieren und Integrieren von Argumenten) und transaktive Aktivitäten, die die Beiträge der Lernpartner weiter auf- und ausbauen (transaktive Ko-Konstruktion) in ihrer Effektivität für das Lernen unterscheiden. Auch ist unklar, ob es eine Rolle spielt, ob transaktive Aktivitäten des Lernenden selbst oder auch die von den jeweiligen Lernpartnern generierten transaktiven Aktivitäten bedeutsam für den individuellen Fertigkeitserwerb sind, da die Quantität und Qualität von Beiträgen zwischen den Lernpartnern einer Lerngruppe großen Schwankungen unterliegen kann (Jiménez-Aleixandre, Bugallo Rodríguez & Duschl, 2000) und Lernende nicht zwingend die transaktiven Beiträge ihrer Lernpartner verarbeiten müssen (z.B. Asterhan & Schwarz, 2009).

Fragestellung: Inwiefern werden Effekte von Argumentationsskripts und heuristischen Lösungsbeispielen auf den Erwerb von Argumentationsfertigkeiten über bestimmte kollaborative Lernprozesse (konstruktive Aktivitäten, transaktive Ko-Konstruktion und transaktive Argumentation) vermittelt?

Inwiefern gibt es einen Unterschied zwischen den Effekten selbst-generierter und vom Lernpartner generierten kollaborativen Lernprozessen auf den Erwerb von Argumentationsfertigkeiten?

Methode: Die Studie wurde im Rahmen eines einwöchigen Vorbereitungskurses für angehende Lehramtsstudierende (Fach Mathematik) in einem 2x2-Design mit den Faktoren Argumentationsskript und heuristische Lösungsbeispiele durchgeführt. Die N=101 Teilnehmer (57 weiblich, Durchschnittsalter: 20 Jahre) wurden randomisiert einer der vier Bedingungen zugewiesen. In drei Treatment-Sitzungen arbeiteten die Teilnehmer in Dyaden an einer mathematischen Beweisaufgabe (z.B. „Summiere ungerade Anzahlen von aufeinander folgenden Zahlen und versuche, Regelmäßigkeiten zu finden. Formuliere eine Vermutung und beweise diese.“). Je nach zugewiesener Bedingung wurden das Argumentationsskript, die heuristischen Lösungsbeispiele, eine Kombination beider Unterstützungsmaßnahmen oder keine Maßnahmen auf zwei verbundenen Laptops angeboten. In den Bedingungen mit Argumentationsskript wurde die Diskussion der Lernpartner in die Phasen Argument, Gegenargument und Synthese strukturiert (Leitão, 2000). In den Bedingungen mit heuristischem Lösungsbeispiel wurde den Lernenden schrittweise ein möglicher Lösungsweg für die mathematische Beweisaufgabe angeboten (Boero, 1999). Zur Messung der Argumentationsfertigkeiten wurde ein Wissenstest eingesetzt, der anschließend von zwei trainierten Beurteilern bewertet wurde (Cohens  $\kappa=.82$ ). Für die Messung der kollaborativen Lernprozesse, wurden die Beiträge der Lernenden während der Treatment-Sitzungen von zwei trainierten Kodierern jeweils einer der drei Aktivitäten (konstruktive Aktivitäten, transaktive Ko-Konstruktion und transaktive Argumentation) oder einer Restkategorie zugewiesen (ICC=.84).

Ergebnisse: Sowohl das Argumentationsskript ( $F(1,97)=14.98, p<.01, \eta^2=.08$ ) als auch die heuristischen Lösungsbeispiele ( $F(1,97)=7.19, p=.04, \eta^2=.04$ ) hatten einen signifikanten positiven Effekt auf den Erwerb von Argumentationsfertigkeiten. Mediatoranalysen zeigen, dass lediglich die von den Lernenden selbst generierte transaktive Argumentation den positiven Effekt des Argumentationsskripts (Sobel  $Z=1.89, p=.03$ ) sowie den positiven Effekt der heuristischen Lösungsbeispiele (Sobel  $Z=2.08, p=.02$ ) erklären. Die vom Lernenden selbst generierten transaktiven Ko-Konstruktionen und konstruktiven Aktivitäten, sowie die von den Lernpartnern generierten Aktivitäten, erklären die Effekte der beiden Unterstützungsmaßnahmen nicht. Zur Unterstützung des Erwerbs von Argumentationsfertigkeiten ist also vor allem die Stimulation transaktiver Argumentation der Lernenden selbst notwendig (Chi, 2009). Da Beiträge der Lernpartner für den kollaborativen Lernprozess unabdingbar sind, müssten Lernende unterstützt werden, von der transaktiven Argumentation ihrer Lernpartner besser zu profitieren.

Asterhan, C. S. C. & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science*, 33(3), 374–400. doi:10.1111/j.1551-6709.2009.01017.x

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. Preuve: International Newsletter on the Teaching and Learning of Mathematical Proof, (July/August 1999). Retrieved from <http://www.lettredelapreuve.it/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html>

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. doi:10.1111/j.1756-8765.2008.01005.x

Jiménez-Aleixandre, M. P., Bugallo Rodríguez A., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84(6), 757–792. doi:10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

Leitão, S. (2000). The Potential of Argument in Knowledge Building. *Human Development*, 43(6), 332–360. doi:10.1159/000022695

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M. & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education*, 61, 59–76. doi:10.1007/s11412-012-9162-z

Reiss, K. & Renkl, A. (2002). Learning to prove: The idea of heuristic examples. *ZDM The International Journal on Mathematics Education*, 34(1), 29–35. doi:10.1007/BF02655690

Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaborations? In C. O'Malley (Ed.), *Discourse, Tools, and Reasoning: Situated Cognition and Technologically Supported Environments* (pp. 361–384). Berlin: Springer.



## **Lehr-Lern-Unterstützung beim Problemlösen am Beispiel mathematischer Textaufgaben - Eine videobasierte Analyse**

Theoretischer Hintergrund: Konstruktivistisch orientierte Vorstellungen zum Wissenserwerb gehen davon aus, dass in der sozialen Interaktion mittels Sprache eine adaptive Lernsituation gestaltet werden kann, welche den aktuellen Lernstand der Schülerin oder des Schülers als auch kognitive Herausforderungen einer Lernaufgabe gleichzeitig berücksichtigt. In einem ko-konstruktiven Lerngespräch, etwa im Rahmen eines tutoriellen Lerndialogs, können individueller Lernstand und Aufgabenanforderungen so koordiniert werden, dass für die Schülerin oder den Schüler ein Lernen in der Zone ihrer nächsten Entwicklung möglich wird (Vygotskij, 1978).

Neben einer differenzierten didaktischen Aufbereitung einer problemorientierten Lernaufgabe ist für die Lernenden die Form der Begleitung durch die Lehrperson entscheidend für ein erfolgreiches Problemlösen (Reusser, 2005). Die Lernwirksamkeit eines tutoriellen Lerndialogs wird jedoch neben pädagogischen Fähigkeiten der Lehrperson auch durch die Beiträge der Lernenden sowie die Qualität der sozialen Interaktion selbst beeinflusst. Auf sprachlicher Ebene zeigt sich, dass offene, kognitiv-aktivierende Hinweise der Lehrperson ebenso wirksam sind wie Erklärungen und Rückmeldungen (Hogan, Nastasi & Pressley, 2000; Chi, Siler, Jeong, Yamauchi & Hausmann, 2001).

Fragestellungen: In der vorliegenden Arbeit wurde untersucht, in welcher Form und mit welchem Anteil Lehrperson wie auch Schülerin oder Schüler im Kontext tutorieller Lerndialoge aufgabenbezogen am Problemlösungsprozess sprachlich beteiligt sind. Des Weiteren wurde der Frage nachgegangen, welcher Form der sozialen Interaktion die tutoriellen Lerndialoge entsprechen. Im Speziellen geht es hierbei um die Frage, ob Gesprächsmuster im Sinne von Scaffolding (Wood, Bruner & Ross, 1976) vorliegen, d.h. Gespräche, in denen es der Lehrperson gelingt, dass sich die Lernenden mit substantiellen Gesprächsbeiträgen am Problemlösungsprozess beteiligen.

Methode: Aus dem Datenmaterial einer schweizerisch-deutschen Videostudie zum Mathematikunterricht (Klieme, Pauli & Reusser, 2009) wurden 12 auf Video aufgezeichnete, tutorielle Lerndialoge (6 aus der Schweiz und 6 aus Deutschland), in welchen je dieselbe mathematische Textaufgabe (8./9. Schuljahr) gelöst wurde, zufällig ausgewählt und mittels zweier Codiersysteme hinsichtlich der sprachlichen Interaktion inhaltlich wie auch strukturell analysiert.

Das Analyseinstrument beruht auf einem Instrument für naturwissenschaftliche Gruppengespräche (Hogan et al., 2000). Es wurde für die vorliegende Analyse mathematischer Problemlösungsprozesse weiterentwickelt. Der erste Codierdurchgang erfasst die Gesprächsbeiträge von Lehrperson und Schülerin bzw. Schüler (z.B. substantielle vs. nicht-substantielle Beiträge, metakognitive Beiträge), um festzustellen, wer mit welchen Gesprächsbeiträgen an der Entwicklung der Problemlösung beteiligt ist. Auf dieser Basis werden im zweiten Codierdurchgang drei Gesprächsmustertypen identifiziert, die sich qualitativ in der Gesprächsbeteiligung unterscheiden: konsensuelles Gesprächsmuster (kaum substantielle Schülerinnen- und Schülerbeiträge), konstruktives Gesprächsmuster (substantielle Beiträge von Lehrperson und Schülerin oder Schüler ohne gegenseitige Elaboration), transaktives Gesprächsmuster (gegenseitige Elaboration). Die zweite Codierung gibt daher Auskunft über die Partizipation von Lehrperson und Schülerin oder Schüler an der inhaltlichen Entwicklung der Problemlösung.

Ergebnisse und Diskussion: Die tutoriellen Lerndialoge dauern je nach gewähltem Lösungsweg zwischen 6.5 bis 17 Minuten. Unabhängig von der sehr unterschiedlichen Gesprächsdauer zeigt sich in allen 12 Dialogen eine ähnliche Verteilung der drei Gesprächsmustertypen. Am häufigsten kommt das konstruktive Muster vor, einen etwas kleineren Anteil umfasst das konsensuelle Muster, und eher selten tritt das transaktive Muster auf.

Die Dominanz des konstruktiven Gesprächsmusters macht sichtbar, dass es den Lehrpersonen gelingt, ihre Gesprächsbeiträge adaptiv auszurichten. Die Lehrpersonen orientieren sich mit ihren Gesprächsbeiträgen am aktuellen Lernstand der Schülerin bzw. des Schülers und formulieren Fragen mehr oder weniger offen, geben Erklärungen oder Hilfestellungen so, dass sich diese im Problemlösungsprozess als verbale Unterstützung im Sinne von Scaffolding (Wood et al., 1976) zeigen. Trotz der hohen Orientierung der Lehrperson am aktuellen Lernstand der Schülerin bzw. des Schülers bleiben die Dialoge dennoch deutlich von der Lehrperson gesteuert, was durch das Ausmaß ihrer Gesprächsbeiträge zum Ausdruck kommt. Eine Übertragung der Verantwortung für die Problemlösung an die Lernenden, im Sinne von „transfer of responsibility“ (Wood et al., 1976), zeigt sich nur selten.

Chi, M. T. H., Siler, St. A., Jeong, H., Yamauchi, T. & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471-533.

Hogan, K., Nastasi, B. K. & Pressley, M. (2000). Discourse Patterns and Collaborative Scientific Reasoning in Peer and Teacher-Guided Discussions. *Cognition and Instruction*, 17 (4), 379-432.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom* (pp. 137-160). Münster: Waxmann.

Reusser, K. (2005). Problemorientiertes Lernen – Tiefenstruktur, Gestaltungsformen, Wirkung. *Beiträge zur Lehrerbildung* 23 (2), 159-182.

Vygotskij, L. S. (1978). *Mind and Society*. Cambridge, MA: Harvard University Press.

Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.





## **Computergestützte Erfassung mathematikdidaktischer Kompetenzen und Überzeugungen von ErzieherInnen und GrundschullehrerInnen**

Aktuelle Kompetenzmodelle für LehrerInnen (Baumert & Kunter, 2006; Blömeke, Kaiser & Lehmann, 2010) ordnen fachdidaktisches Wissen und professionsbezogene Überzeugungen stets als zentrale Komponenten der professionellen Kompetenz ein, neben anderen Wissenskomponenten (Fachwissen, pädagogisches Wissen, Organisationswissen und Beratungswissen) und weiteren Komponenten (wie Motivation und Selbstregulation).

Für LehrerInnen im Anfangsunterricht Mathematik sowie für ErzieherInnen, die im Kindergarten mathematische Lernprozesse fördern sollen, bedeutet dies: Ihre Kompetenz hängt unter anderem ab (1) vom vorhandenen mathematikdidaktischen Wissen (etwa über die Entwicklung des Zählens und Rechnens einschließlich typischer Zähl- und Rechenfehler sowie ihrer Ursachen oder über die Anregung und Begleitung entsprechender Lernprozesse einschließlich der Auswahl sinnvoller Materialien) sowie (2) von den auf das Mathematiklernen bezogenen Überzeugungen (insbesondere solchen, die handlungsrelevant sind, sich also speziell auf die Gestaltung mathematischer Lernsituationen beziehen).

Im Sinne eines anschlussfähigen Mathematiklernens für alle Kinder müssen beim Übergang vom Kindergarten in die Grundschule auftretende Diskontinuitäten entwicklungsförderlich sein und deshalb gezielt gestaltet werden (Roßbach, 2006); sie dürfen nicht das Ergebnis von Zufälligkeiten aufgrund eines nicht abgesprochenen Handelns beider Institutionen sein. Im Hinblick auf die Kooperation von Kindergarten und Grundschule gelten deshalb neben den Kompetenzen insbesondere auch die Überzeugungen der beteiligten Fachkräfte als ein wichtiger Einflussfaktor für deren Gelingen (von Bülow, 2011; Faust, Wehner & Kratzmann, 2011). Aufgrund der unterschiedlichen Ausbildungen und institutionellen Traditionen besteht jedoch die Hypothese, dass in Bezug auf das mathematikdidaktische Wissen und die Überzeugungen zum Mathematiklernen erhebliche Unterschiede (1) zwischen ErzieherInnen und GrundschullehrerInnen und (2) innerhalb der Gruppe der GrundschullehrerInnen zwischen jenen mit Mathematikstudium und fachfremd Unterrichtenden bestehen.

In einer Teilstudie des Projekts Anschlussfähigkeit mathematikdidaktischer Überzeugungen und Praktiken von ErzieherInnen und GrundschullehrerInnen als Voraussetzung für die Kooperation von Elementar- und Primarbereich (kurz: AnschlussM; [www.anschluss-m.de](http://www.anschluss-m.de)), einem vom BMBF geförderten Verbundprojekt der Universität Bremen und der Pädagogischen Hochschule Freiburg, wurde diese Hypothese untersucht. Als Mittel der Wahl für eine handlungsnahere Erfassung fachdidaktischer Kompetenzen und Überzeugungen gelten computerbasierte Settings mit Bild- und Videovignetten (Lindmeier, 2011). Im Sommer 2013 wurden so N = 161 ErzieherInnen und GrundschullehrerInnen in Baden-Württemberg und Bremen befragt. Zu sechs Bild- und vier Videovignetten sollten die TeilnehmerInnen jeweils Fragen in Bezug auf (1) das mathematische Potenzial der Situation, (2) die Förderung allgemeiner mathematischer Kompetenzen und (3) mögliche Reaktionen in einer konflikthaltigen Lernsituation in Form von Freitext direkt am Computer beantworten.

Die Antworten wurden in einem konsensualen Verfahren zweifach unabhängig voneinander kodiert (u.a. im Hinblick auf aufgetretene fachliche Fehler, angesprochene Leitideen, erkennbare Überzeugungen) bzw. geratet (bezüglich der mathematikdidaktischen Qualität der Antwort). Die Auswertung der so gewonnen hochinferenten Daten erfolgte entsprechend der Hypothesen varianzanalytisch.

Es ergeben sich signifikante Unterschiede zwischen den jeweiligen Gruppen. Auch wenn die mathematikdidaktische Qualität der Antworten insgesamt niedrig ist und unter den Erwartungen liegt, zeigt sich, dass GrundschullehrerInnen mit Mathematikstudium im Unterschied zu Fachfremden und ErzieherInnen das Potenzial einer Situation signifikant treffender einschätzen können und in der Folge auch ihre Handlungsvorschläge signifikant besser sind. Auch in Bezug auf die hinter den Handlungsbeschreibungen zutage tretenden Überzeugungen treten Differenzen auf. Insgesamt unterstreicht die Studie damit, dass eine aus heutiger mathematikdidaktischer Sicht adäquate Lernbegleitung nur gelingen kann, wenn die pädagogischen Fachkräfte auch eine entsprechende Ausbildung erfahren haben. Angesichts der Tatsache, dass in Baden-Württemberg beispielsweise über 40% der GrundschullehrerInnen Mathematikfachfremd unterrichten (Richter et al., 2011), ist diese eine alarmierende Konsequenz.

Baumert, J. & Kunter, M. (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft, 9, H. 4, 469–520.

Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.) (2010): TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich. Münster: Waxmann.

Bülow, K. von (2011): Anschlussfähigkeit von Kindergarten und Grundschule. Rekonstruktion von subjektiven Bildungstheorien von Erzieherinnen und Lehrerinnen. Bad Heilbrunn: Klinkhardt.

Faust, G., Wehner, F. & Kratzmann, J. (2011). Zum Stand der Kooperation von Kindergarten und Grundschule. Maßnahmen und Einstellungen der Beteiligten. In: Journal für Bildungsforschung Online, 3, H. 2, 38–61.

Lindmeier, A. (2011). Modeling and Measuring Knowledge and Competencies of Teachers: A Threefold Domain-Specific Structure Model for Mathematics. Münster: Waxmann.

Richter, D. et. al. (2011): Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe. In: Stanat, P. et al. (Hrsg.): Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Länder-ver-gleichs 2011. Münster: Waxmann, 237–250.

Roßbach, H.-G. (2006): Institutionelle Übergänge in der Frühpädagogik. In: Fried, L. & Roux, S. (Hrsg.): Pädagogik der frühen Kindheit. Weinheim: Beltz, 280–292.

## **Effekte eines lehrervermittelten Trainings zur Förderung von Transformationsprozessen beim Lernen mit multiplen Repräsentationen**

Theoretischer Hintergrund: Der verständige Umgang mit multiplen Repräsentationen für den mathematischen Wissenserwerb ist von grundlegender Bedeutung. Dies wird auf curriculärer Ebene (KMK, 2012) sowie auf Ebene bildungswissenschaftlicher und mathematikdidaktischer Forschung immer wieder hervorgehoben (Vogel, 2006). Bei der Verarbeitung multipler Repräsentationen werden an die Lernenden verschiedene Anforderungen gestellt: So müssen sie Informationen aus den Repräsentationen entnehmen (Rezeption), verschiedene Formen von Repräsentationen kombinieren (Integration; Seufert, 2003), neue Repräsentationen erstellen (Produktion; Monk, 1989) oder die Information einer Repräsentation in eine andere Repräsentation überführen (Übersetzung; vgl. Schnotz, Baadte, Müller & Rasch, 2010). Empirische Studien zeigen, dass Lernende über verschiedene Altersstufen hinweg Schwierigkeiten im Umgang mit multiplen Repräsentationen haben (Ainsworth, 1999).

In der vorliegenden Studie wurden die Effekte eines lehrervermittelten Trainings zur Förderung von Transformationsprozessen beim Lernen mit multiplen Repräsentationen in der Mathematik analysiert. Es wurde erwartet, dass sich die Prozesse der Rezeption, Produktion, Integration und Übersetzung durch das Training verbessern und diese Verbesserungen langfristig sind.

Methode: Es wurde ein quasiexperimentelles Wartekontrollgruppendesign in den 8. und 9. Klassen eines Gymnasiums durchgeführt, um mögliche Verbesserungen der Lernenden hinsichtlich Rezeption, Produktion, Integration und Übersetzung zu untersuchen. Die Stichprobe setzte sich aus einer Experimental- (EG), einer Kontroll- (KG1) und einer Wartekontrollgruppe (KG2) zusammen. In der EG waren  $n = 92$  Schüler und in der KG  $n = 130$  Schüler. Für die Analyse des Langzeiteffekts wurden in der EG  $n = 92$  und in der KG  $n = 44$  Probanden einbezogen. Die Lehrkräfte der EG und KG2 nahmen an einem Workshop zum Projekt „Fit in der Mathematik“ teil, welcher sich in einen kognitionspsychologischen und einen didaktischen Teil gliederte. Ziel war es den Lehrkräften die Transformationsprozesse beim Lernen mit multiplen Repräsentationen und deren Förderung zu vermitteln. Anschließend wurden die Schülerinnen und Schüler der EG über zwei Monate im Unterricht trainiert. Um die Effekte des Trainings zu erfassen, wurde als abhängige Variable die Differenz der errechneten Lösungswahrscheinlichkeit entsprechender Aufgaben zur Erfassung der Rezeption, Produktion, Integration und Übersetzung bestimmt. Dabei umfasste die Studie drei Messzeitpunkte: vor dem Training ( $t_0$ ), nach dem Training in der EG ( $t_1$ ) und nach zwei Monaten nach dem Training ( $t_2$ ). Als Kontrollvariablen wurden verbale und logische Fähigkeiten, Alter und Geschlecht verwendet.

Ergebnisse: Zur Analyse der Forschungsfrage wurden einfaktorielle Kovarianzanalysen für die jeweiligen abhängigen Variablen durchgeführt. Die abhängige Variable stellte die Differenz der errechneten Lösungswahrscheinlichkeit zu  $t_1$  und  $t_0$  dar. Es zeigten sich positive Trainingseffekte bei der Rezeption ( $F(1, 219) = 65.00, p < .01; \eta^2 = .23$ ), Produktion ( $F(1, 218) = 4.36, p < .05; \eta^2 = .02$ ) und Integration ( $F(1, 219) = 8.56, p < .01; \eta^2 = .04$ ), so dass sich diese Fähigkeiten bei den trainierten Schülern im Vergleich zu den untrainierten Schülern verbesserten. Dagegen wurde kein signifikanter Trainingseffekt für die Übersetzung gefunden ( $F(1, 231) < 1, n.s.$ ).

Um die langfristigen Effekte zu analysieren, wurden die Differenzen der jeweiligen Lösungswahrscheinlichkeiten zu  $t_2$  und  $t_0$  bestimmt. Es zeigte sich ein langfristiger Trainingseffekt für alle Transformationsprozesse (Rezeption:  $F(1, 133) = 4.29, p < .05; \eta^2 = .03$ ; Produktion:  $F(1, 133) = 3.37, p < .05; \eta^2 = .03$ ; Integration:  $F(1, 133) = 2.91, p < .05; \eta^2 = .02$ ; Übersetzung:  $F(1, 133) = 3.25, p < .05; \eta^2 = .02$ ).

Diskussion: Die Fähigkeiten von Lernenden zur Rezeption, Produktion und Integration multipler Repräsentationen in der Mathematik konnten anhand eines Trainings gefördert werden. Für den Übersetzungsprozess zeigten sich keine Verbesserung im Vergleich zu einer untrainierten Schülergruppe, wobei Lehrkräfte seltener die Strategien zur Förderung der Übersetzungsprozesse im Unterricht einsetzten als die anderen Strategien. Darüber hinaus konnte eine langfristige Wirkung des Trainings für alle untersuchten Transformationsprozesse bestätigt werden. Obwohl sich hinsichtlich der Übersetzung kein Effekt über den Trainingszeitraum hinweg zeigte, fand sich ein langfristiger Effekt. Der Prozess der Übersetzung scheint insbesondere intensives Üben zu benötigen. Um nachhaltige Veränderungen zu bewirken, sollten die Prozesse über einen längeren Zeitraum häufig trainiert werden.

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, (33), 131–152.

Kultusministerkonferenz (KMK) (2012). Bildungsstandards im Fach Mathematik für die allgemeine Hochschulreife. [verfügbar unter [www.kmk.org](http://www.kmk.org)]. [Zugriff am 28.10.2013].

Monk, G. S. (1989). Students' Understanding of Functions in Calculus Courses. *Humanistic Mathematics Network Journal*, (9), 21–27.

Schnotz, W., Baadte, C., Müller, A., & Rasch, R. (2010). Creative thinking and problem solving with depictive and descriptive representations. In L. Verschaffel, E. Corte, J. Elen, & T. Jong de (Eds.), *Use of External Representations in Reasoning and Problem Solving*. (pp. 11–35). Amsterdam: Elsevier.

Seufert, T. (2003). Wissenserwerb mit multiplen Repräsentationen: Wirksamkeit von Kohärenzbildungshilfen (1st ed.). Berlin: Logos Verlag Berlin.

Vogel, M. (2006). Mathematisieren funktionaler Zusammenhänge mit multimedialer Supplimentation. Hildesheim: Franzbecker.

## Kompetenzunterschiede von Quereinsteigern und Lehramtsabsolventen im Fach Physik

### Chair(s): Lars Oettinghaus (Goethe-Universität Frankfurt)

Theoretischer Hintergrund und Ausgangslage: Im den Jahren zwischen 2004 und 2008 lag die gymnasiale Quereinsteigerquote im Fach Physik zwischen 40 und 64 Prozent. Folglich durchläuft nur noch ein Teil der künftigen Physiklehrkräfte die universitäre Phase der Lehrerbildung (Lamprecht, 2011). Welche Folgen diese Programme der Kultusministerien auf die „professionelle Kompetenz“ (Baumert & Kunter, 2006) der künftigen Physiklehrkräfte besitzt, ist empirisch noch nicht ausreichend belegt. Für die Mathematik konnten im Rahmen von COACTIV-R erste Erkenntnisse bezüglich des Professionswissens gewonnen werden (Kleickmann & Anders, 2011). In der Physik werden Lehramtsabsolventen und Quereinsteiger seit 2008 durch das Projekt proΦ untersucht.

Professionelle Überzeugungen werden in proΦ im Sinne der „Cognitiv Guided Instruktion“ von Fennema, Carpenter und Loef (1990) aufgefasst, die von Staub und Stern (2002) als „Cognitiv Constructivist Orientation“ bezeichnet werden. Das etablierte Messinstrument, das diese Überzeugung abbildet, wird oft zweigeteilt in einer rezeptartigen/transmissiven und einer selbständigen/konstruktivistischen Subkategorie verwendet, die sich auf andere Instrumente übertragen lassen (Oettinghaus, Lamprecht, & Korneck, 2012). Um dem ursprünglichen Ansatz zu entsprechen, richtet proΦ beide Subskalen positiv aus und bezieht den konstruktivistischen Begriff auf das gesamte Messinstrument.

Das Fachwissen und fachdidaktisches Wissen wird ineinander übergehend betrachtet (Ball, Thames, & Phelps, 2008), wobei proΦ den verwendete Test zum fachdidaktischen Wissen von Riese (2009), in Anlehnung an Ball und Kollegen (2008) und von Riese abweichend, als eine Form des fachlichen Unterrichtswissens, das neben dem Fachwissen das fachdidaktische Wissen enthält, interpretiert.

Forschungsfrage: Im Projekt proΦ wird untersucht, ob der Umfang der Lerngelegenheiten von Referendarsgruppen mit unterschiedlichen Studienabschlüssen in den professionellen Überzeugungen, dem Fachwissen und dem fachdidaktischen Wissen abgebildet werden.

Studiendesign: Von Dezember 2008 bis Mai 2011 wurden in drei Erhebungen insgesamt 368 Physikreferendare mit verschiedenen universitären Ausbildungswegen erfasst. Die erste Erhebung erfasst 90% aller zum ersten Befragungszeitpunkt beginnenden Referendare im gymnasialen Bereich in den Bundesländern Baden-Württemberg, Hamburg, Hessen und Niedersachsen.

Für diese Studie wurden Items und Aufgaben aus etablierten Instrumenten verwendet, um an die aktuelle, empirisch fundierte Diskussion anzuschließen. Das Testheft enthielt Instrumente zu den Persönlichkeitsmerkmalen, zu den selbstregulativen Fähigkeiten (Schaarschmidt & Fischer, 2008), zur Motivation (Mayr, 1994) und zu den Überzeugungen (Fennema, Carpenter, & Loef, 1990; Staub & Stern, 2002; Seidel, Prenzel, & Kobarg, 2005). Die zweite Erhebung im Haupt-, Real- und Gesamtschulbereich (HRGe) erfasste Referendare in allen ausbildenden Studienseminaren in Bremen, Hessen und Niedersachsen, wobei das Testheft um Instrumente zum physikalischen Wissen und physikdidaktischen Wissen (Riese, 2009), ergänzt wurde. Um auch für den gymnasialen Bereich das Professionswissen erfassen zu können, wurden in einer dritten Erhebung erneut Studienseminare in Baden-Württemberg und Bremen im gymnasialen Bereich befragt.

Die Recherche der universitären Lerngelegenheiten der Probanden zeigt, dass die vier größten Probandengruppen folgende Zusammensetzungen der Lerngelegenheiten repräsentieren:

- 148 Referendare mit einem gymnasialen Lehramtsstudium (Physik 70 SWS, Physikdidaktik 6 SWS, Bildungswissenschaften 11 SWS),
- 63 Referendare mit einem Haupt-, Real- und Gesamtschullehramtsstudium (Physik 29 SWS, Physikdidaktik 15 SWS, Bildungswissenschaften 30 SWS),
- 66 Quereinsteiger im Referendariat mit einem Physikstudium (Physik 112 SWS, Physikdidaktik 0 SWS, Bildungswissenschaften 0 SWS),
- 60 Quereinsteiger im Referendariat mit einem Chemie- oder Ingenieurstudium (Physik 33 SWS, Physikdidaktik 0 SWS, Bildungswissenschaften 0 SWS).

Methode und Ergebnisse: Für das Messmodell wurde jedes Merkmal durch drei Dimensionen rasch-skaliert, wobei die drei WLE-Personenschätzer eine intervallskalierte Datenbasis für ein gemeinsames Strukturmodell darstellen.

Die Modelle sind in der Lage sowohl direkt gemessene Merkmale der Testinstrumente als auch indirekt erfasste Merkmale (durch Strukturannahmen) auszuwerten. Die Unterschiede der Referendargruppen werden aus der Perspektive mehrerer Modelle diskutiert. Für die Quereinsteiger-Situation kann so ein differenziertes Bild gezeichnet werden (Oettinghaus, Korneck, & Lamprecht, eingereicht).

Ball, D., Thames, M., & Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59(5), S. 389-407.

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Handlungskompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft* Heft 4, S. 469-520.

Fennema, E., Carpenter, T., & Loef, M. (1990). Teacher belief scale: Cognitively guided instruction project. 1, S. 195-221.

Kleickmann, T., & Anders, Y. (2011). Lernen an der Universität. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften - Ergebnisse des Forschungsprogramms COAKTIV* (S. 305-315). Münster. New York. MÜNchen. Berlin: Waxmann Verlag GmbH.

Lamprecht, J. (2011). *Ausbildungswege und Komponenten professioneller Handlungskompetenz*. Berlin: Logos verlag Berlin GmbH.

Mayr, J. (Hrsg.). (1994). *Lehrer/in werden*. Innsbruck: Österreichischer Studienverlag .

Oettinghaus, L., Korneck, F., & Lamprecht, J. (eingereicht). Analyse der professionellen Kompetenz von Referendaren. *phydid b*.

Oettinghaus, L., Lamprecht, J., & Korneck, F. (2012). Überzeugungen zum Unterrichtsfach und zur Wissenschaft Physik - Ein Skalenvergleich. (V. Nordmeier, & H. Grötzebauch, Hrsg.) *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung in Mainz*.

Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Berlin: Logos Verlag Berlin GmbH.

Schaarschmidt, U., & Fischer, A. W. (2008). *AVEM - Arbeitsbezogenes Verhaltens- und Erlebensmuster*. London: Pearson.

Seidel, T., Prenzel, M., & Kobarg, M. (2005). *How to run a video study: Technical report of the IPN Video Study*. Münster: Waxmann.

Staub, F. C., & Stern, E. (2002). The nature of Teachers' Pedagogical Content Beliefs Matters for Students' Achievement Gains: Quasi-Experimental Evidence From Elementary Mathematics. *Journal of Educational Psychology*, 94(2), S. 344-355.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G150

## Analyse von Unterrichtshandeln in komplexitätsreduzierten Physikunterrichtssequenzen

Ausgangspunkt und theoretischer Hintergrund der Videostudie  $\Phi$ actio:  $\Phi$ actio hat das Ziel, den Zusammenhang zwischen den fachbezogenen Kompetenzkomponenten „Überzeugungen“ und „Professionswissen“ von Physiklehrkräften auf das Unterrichtshandeln und die Unterrichtsqualität zu analysieren. Insbesondere bezüglich des Zusammenhangs zwischen Überzeugungen und Unterrichtshandeln existieren bisher heterogene Befunde: Die COACTIV-Studie (Kunter et al., 2011; Dubberke et al., 2008) erhob die Überzeugungen von Mathematiklehrkräften sowie die Qualitätsmerkmale ihres Unterrichts mit dem Ergebnis, dass Lehrkräfte mit transmissiven Lernüberzeugungen den Unterricht weniger herausfordernd und aktivierend sowie eher Fehler vermeidend gestalten. Die geringere kognitive Aktivierung wirkte sich in dieser Studie auf den Lernerfolg der Schülerinnen und Schüler aus. Eine aktuelle Veröffentlichung (Kunter et al., 2013) belegt bei univariater Analyse einen Zusammenhang zwischen den konstruktivistischen Überzeugungen der Lehrkräfte und den Mathematikleistungen der Schülerinnen und Schüler. Bei multivariaten Modellen wurde der Effekt nicht mehr nachgewiesen, allerdings ein negativer Zusammenhang mit der Qualitätsdimension „Klassenführung“. Die Qualitätsmerkmale des Unterrichts wurden in der COACTIV-Studie durch Schülerfragebögen und Aufgabenanalysen erhoben.

Die IPN-Videostudie (Seidel et al., 2008) analysierte Videos aus dem Bereich des Physikunterrichts. Sie fand keine systematischen Zusammenhänge zwischen den Überzeugungen der Physiklehrkräfte mit den beobachteten Unterrichtsmustern sowie mit dem inhaltlichen Wissen und dem Interesse der Schülerinnen und Schüler. Mögliche Ursachen für die unterschiedlichen Ergebnisse können sowohl in der Erhebungsmethodik als auch in den Fachspezifika der Lehrerüberzeugungen liegen. Auf letzteres deuten die Ergebnisse der Frankfurter Referendarstudie pro $\Phi$  (Lamprecht et al., 2011) hin: Während sowohl die IPN- als auch die COACTIV-Studie von den beiden Überzeugungsmerkmalen „constructivist view“ und „transmission view“ ausgehen, fand die Referendarstudie zusätzlich ein Überzeugungsmuster, das konstruktivistische und transmissive Überzeugungsmerkmale vereint.

Forschungsfragen: Auf der Basis dieser Befunde untersucht die Studie  $\Phi$ actio den Zusammenhang zwischen den Überzeugungen zukünftiger Lehrkräfte und deren Unterrichtshandeln aus mehreren Perspektiven (Schüler- und Studierendenfragebögen sowie Videokodierung).

Um die Videostudie ökonomischer zu gestalten, wird die Unterrichtsqualität im Rahmen von komplexitätsreduzierten Physikunterrichtssequenzen (12 bis 15 Minuten Unterrichtszeit, halbe Klassenstärke, thematische Festlegung auf Freihandexperimente zur Mechanik) erhoben. Deshalb steht zunächst die forschungsmethodische Frage im Vordergrund, ob Unterrichtsqualitätsmerkmale auch in komplexitätsreduzierten Sequenzen nachgewiesen werden können.

Design und Methoden: Die Erhebungen erfolgen im Rahmen der phasen- und schulformübergreifenden Veranstaltung „Unterrichtsversuche mit Videofeedback“, deren zentrales Element die komplexitätsreduzierten Physikunterrichtssequenzen der jeweils 18 Teilnehmer ist. Jede dieser Sequenzen wird zweimal mit je einer Klassenhälfte unterrichtet und direkt nach dem Unterricht sowohl von hospitierenden Peers als auch von den Schülerinnen und Schülern bezüglich der Unterrichtsqualität (Basisdimensionen „Klassenführung“, „Kognitive Aktivierung“ und „Konstruktive Unterstützung“) geratet. Dazu wurde der Schülerfragebogen der COACTIV-Studie für den komplexitätsreduzierten Physikunterricht und zwei Minuten Bearbeitungszeit überarbeitet.

Die Analyse der videografierten Unterrichtssequenzen erfolgt mit Hilfe des Kodiermanuals der IPN-Videostudie (Seidel et al., 2003).

Vor Beginn der Veranstaltung findet eine Fragebogenerhebung zu Komponenten der professionellen Lehrerkompetenz (Lamprecht, 2011; Riese, 2009), zu Personenmerkmalen und Lerngelegenheiten sowie zu physikdidaktischen Selbstwirksamkeitserwartungen (Rabe et al., 2013) statt. Der Nachtest wird drei Monate nach der Veranstaltung durchgeführt und ist auf Überzeugungen und Selbstwirksamkeitserwartungen beschränkt.

Ergebnisse zur forschungsmethodischen Fragestellung: Die Ergebnisse der Mehrebenenstrukturgleichungsmodelle und den daraus resultierenden Intraklassenkorrelationen der Daten der Schüler- und Peerfragebögen erfüllen auch für komplexitätsreduzierte Unterrichtssequenzen und einer reduzierten Erhebungsdauer von nur zwei Minuten die Gütekriterien (Krüger et al., 2013).

Erste Videoanalysen zeigen in Sichtstukturanalysen sehr gute Übereinstimmungswerte der Beobachter und in der Tiefenstrukturanalyse gute, mit den Ergebnissen der IPN-Videostudie vergleichbare, Übereinstimmungswerte. Im hoch-inferenten Rating der gesamten Sequenzen erreichen die Items, vergleichbar mit der IPN-Videostudie, noch nicht die Kriteriumswerte (Korneck et al., 2013). Damit weisen diese ersten Analysen darauf hin, dass die eingesetzten Instrumente zur Erfassung der Unterrichtsqualität auch für kurze Sequenzen und Videos von 12 Minuten Unterrichtsdauer einsetzbar sind und vergleichbare Ergebnisse liefern.

Dubberke, T., Kunter, M., McElvaby, N., Brunner, M., Baumert, J. (2008): Lerntheoretische Überzeugungen von Mathematiklehrkräften. Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. Zeitschrift für Pädagogische Psychologie, 22, S. 193-206.

Korneck, F., Kohlenberger, M., Oettinghaus, L., Kunter, M.,

Lamprecht, J. (2013):  $\Phi$ actio - Überzeugungen und Unterrichtshandeln von Physiklehrkräften, phydid b, eingereicht.

Krüger, M., Oettinghaus, L., Korneck, F., Drobnak, L., Kunter, M. (2013): Schüler- und Peerbefragungen zu komplexitätsreduzierten Sequenzen. In Bernholt, S. (Hg.): "Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht", Jahrestagung der GDPC in München 2013, <http://www.gdcp.de> (Stand 11-2013), eingereicht.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., Neubrand, M. (2011): Die professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., Hachfeld, A. (2013): Professional competence of teachers: Effects on instructional quality and student development. Journal of Educational Psychology, 105(3), 805-820. doi: 10.1037/a0032583

Lamprecht, J. (2011): Ausbildungswege und Komponenten professioneller Handlungskompetenz. Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik. Dissertationsarbeit an der Universität Frankfurt a.M., Fachbereich Physik. Berlin: Logos Verlag.

Oettinghaus, L., Lamprecht, J., Korneck, F. (2013): Analyse der professionellen Kompetenz von Referendaren. In Bernholt, S. (Hg.): "Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht", Jahrestagung der GDPC in München 2013, <http://www.gdcp.de> (Stand 11-2013), eingereicht.

Rabe, T., Krey, O., Meinhardt, C. (2013): Physikdidaktische Selbstwirksamkeitserwartungen zukünftiger Physiklehrkräfte. In Bernholt, S. (Hrsg.): Inquiry-based Learning – Forschendes Lernen. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Hannover 2012. Münster: LIT, S. 635-637.

Riese, J. (2009): Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Lehrkräften. Dissertation. Berlin: Logos Verlag.

Seidel T., Schwindt, K., Rimmel, R. & Prenzel, M. (2008): Konstruktivistische Überzeugungen von Lehrpersonen: Was bedeuten sie für den Unterricht? In: Zeitschrift für Erziehungswissenschaften, 10, Sonderheft 8/2008, S. 259-276.

Seidel, T., Prenzel, M., Duit, R., Lehrke, M. (2003): Technischer Bericht zur Videostudie, Lehr- und Lernprozesse in der Physik, IPN-Materialien, IPN Kiel.





## Diagnostik des Erklärungswissens von Physiklehramtsstudierenden

Motivation: In der physikdidaktischen Literatur wird verständliches Erklären als eine wichtige Fähigkeit von Lehrkräften beschrieben. Gleichzeitig wird ein Defizit an empirischer Forschung beklagt (Geelan, 2013).

Theoretischer Hintergrund: Das Erklärungswissen kann als Teil des Professionswissens von Lehrkräften verstanden werden. Üblicherweise wird Professionswissen in Fachwissen, fachdidaktisches und pädagogisches Wissen unterteilt (z. B. Shulman, 1986). Das Erklärungswissen kann wie bei COACTIV (Kunter et al., 2011) als Teil des fachdidaktischen Wissens verstanden werden.

Erklärungswissen ist dazu notwendig, primär verbal einen Sachverhalt so darzustellen, dass dieser für ein Gegenüber wahrscheinlicher verstehbar wird (Kulgemeyer, im Druck; Gage, 1968). Um es erfassbar zu machen, wurde ein Modell verbalen Erklärens erstellt, das sich an das physikbezogene Kommunikationsmodell von Kulgemeyer und Schecker (2013) anlehnt. Hier wird davon ausgegangen, dass ein Erklärer auf vier Ebenen (Sprachebene, Mathematisierung, Darstellungsformen und Beispiele) Entscheidungen trifft, um eine Erklärung sachgerecht und adressatengemäß zu gestalten. Auf der Sprachebene entscheidet der Erklärer beispielsweise über Fachbegriffe, die er nutzen will. Hohes Erklärungswissen macht sich unter anderem daran fest, diese vier Ebenen so zu variieren, dass die Bedürfnisse eines Adressaten erfüllt werden. Zudem nutzen gute Erklärer Strategien, die allgemein verständnisförderlich sind, z.B. eine strukturierte Darstellung des Sachverhalts (z. B. Kiel, 1999; Brown 2006).

Ziel: Ziel des hier vorgestellten Forschungsprojekts EWis (Erklärungswissen von Physikstudierenden) ist die Modellierung des Erklärungswissens von Lehramtsstudierenden der Physik als Teil des Professionswissens. Dazu sollen Zusammenhänge zwischen dem Erklärungswissen, dem Fachwissen sowie dem (deklarativen) fachdidaktischen Wissen untersucht werden. Die beiden letzten Wissensbereiche werden mit schriftlichen Tests erhoben, die parallel von Projektpartnern des Verbundprojekts ProfiLe-P entwickelt werden. Bei der Messung des Professionswissens wird häufig kritisiert, dass mit Paper-Pencil-Test viele Aspekte nicht genau und ohne Handlungsbezug erfasst werden (von Aufschnaiter & Blömeke, 2010). Ein weiteres Ziel ist deshalb die Entwicklung einer Erhebungsmethode, mit der das Erklärungswissen von Physikstudierenden im Prozess des Erklärens erhoben werden kann.

Methode: Zur Diagnose des handlungsnahen Erklärungswissens wurde eine qualitative Erhebungsmethode aus der Forschung zur Kommunikationskompetenz bei Schülerinnen und Schülern adaptiert (Kulgemeyer & Schecker, 2013). Darin wird eine reale Vermittlungssituation nachgestellt und videografiert. Die Studierenden erhalten die Aufgabe, einem Schüler bzw. einer Schülerin der 10. Klasse (Adressaten) einen vorgegebenen physikalischen Sachverhalt zu erklären. Die Adressaten sind trainiert, bestimmte Fragen zu stellen und sich bei jedem Probanden vergleichbar zu verhalten. Mit der Methode gelingt es, eine standardisierte und unterrichtsnahe Erklärungssituation zu schaffen.

Zur u.a. kognitiven Validierung (Messick, 1995) dieser Methode wurden vier Experteninterviews mit in Schule und Hochschule erfahrenen Lehrkräften sowie sieben Stimulated Recall Interviews mit den Probandinnen und Probanden durchgeführt.

Die Auswertung der Erklärungsvideos geschieht anhand eines durch qualitative Inhaltsanalyse erzeugten Kodiermanuals, das u.a. auf dem Erklärungsmodell beruht. Aus dem Kategoriensystem wurde eine ordinale Skala für die Ausprägung von Erklärungswissen erzeugt, indem das Auftauchen von nach Expertenmeinung verständnisförderlichen Kategorien summiert wurde.

In einem weiteren Expertenrating wurde diese Skala validiert. Fünf weitere Experten wurden gebeten, die Qualität zweier Erklärungsvideos im direkten Vergleich hochinferent einzuschätzen. Insgesamt wurden bislang 64 Videos von Erklärensituationen à 10 Minuten Dauer aufgenommen. Stichprobe sind dabei Studierende des Physik-Lehramts und des Vollfachs Physik aus allen Semestern. In der gerade angelaufenen Hauptstudie werden etwa 120 weitere Videos aufgenommen.

Ergebnisse: Die Entwicklung verschiedener hinsichtlich Adressatenverhalten standardisierter Settings ist abgeschlossen. Dazu existiert ein Kodiermanual, das Kriterien guten Erklärens modellgestützt beschreibt. Aus den Kategorien lässt sich ein reliables Maß für Erklärungswissen errechnen (Cronbachs Alpha  $\alpha = 0,735$ ). Dieses Maß ist zudem in der Lage, ein Expertenurteil über die bessere von zwei Erklärung vorherzusagen. Zudem liegen qualitative Ergebnisse vor, die die Annahme stützten, dass die Settings authentisch und valide sind. Die Konzeption der Methode soll in diesem Vortrag vorgestellt werden.

Brown, G. (2006). Explaining. In: O. Hargie (Hrsg.). The Handbook of Communication Skills. London, New York: Routledge, 195-228.

Gage, N. L., Belgard, M., Dell, D., Hiller, J. E., Rosenshine, B. & Unruh, W. R. (1968). Exploration of the Teacher's Effectiveness in Explaining. Technical Report No.4. Stanford Center for Research and Development in Teaching.

Geelan, D. (2013). Teacher Explanation of Physics Concepts: a Video Study. Research in Science Education 43(5), 1751-1762.

Kiel, E. (1999). Erklären als didaktisches Handeln. Würzburg: Ergon.

Kulgemeyer, C. (2010). Physikalische Kommunikationskompetenz. Modellierung und Diagnostik. Berlin: Logos.

Kulgemeyer, C. (im Druck). Gelingensbedingungen physikalischer Erklärungen - Zu einer konstruktivistischen Auffassung des Erklärens. PhyDidB - Beiträge zur DPG-Frühjahrtagung.

Kulgemeyer, C. & Schecker, H. (2013). Students explaining science - assessment of science communication competence. Research in Science Education, DOI: 10.1007/s11165-013-9354-1

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. und Neubrand, M. (Hrsg.) (2011). Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.

Shulman, L. (1986). Those Who Understand: Knowledge Growth in Teaching. Educational Researcher, 15 (2), 4-14.

Treagust, D. & Harrison, A. (1999). The genesis of effective science explanations for the classroom. In J. Loughran (Eds.), Researching Teaching: Methodologies and Practices for Understanding Pedagogy. Abingdon: Routledge, 28-43.

von Aufschnaiter, C. & Blömeke, S. (2010). Professionelle Kompetenz von (angehenden) Lehrkräften erfassen – Desiderata. ZfDN, 16, 361-367.





## **Pilotierungsergebnisse des Projekts DaWis: Fachdidaktisches Wissen von Physiklehramtstudierenden**

**Fragestellung und Zielsetzung:** Im Verbundprojekt Profile-P (vgl. Kulgemeyer et al., 2012) sollen die für Lehrer relevanten Wissensbereiche des Erklärungswissens, des Fachdidaktischen Wissens und des Fachwissens modelliert, operationalisiert und deren Zusammenhänge untersucht werden. Insbesondere für das Fachdidaktische Wissen mangelt es jedoch sowohl an einem Konsens über die relevanten Inhalte als auch an Testinstrumenten, die auch die innere Struktur dieses Wissens empirisch abbilden könnten. Ausgehend von dieser Problematik wurde im Teilprojekt DaWis ein möglichst breit angelegtes Modell der inneren Struktur des Fachdidaktischen Wissens, das Lehramtsstudierende an der Universität erwerben können, erstellt und auf dessen Grundlage ein Testinstrument entwickelt und validiert.

**Theoretischer Hintergrund:** Ausgehend von einer Analyse verschiedener Konzeptualisierungen von Fachdidaktischem Wissen (FDW) und pedagogical content knowledge (PCK) wurde ein eigenes Modell Fachdidaktischen Wissens mit dem Schwerpunkt des an der Universität erwerbenden Wissens entwickelt. Dieses ordnet sich in die gängigen heuristischen Modelle professioneller Handlungskompetenz in Anlehnung an Baumert und Kunter (2006) ein, wobei das Fachdidaktische Wissen als Teilfacette des Professionswissens betrachtet wird. Für die Modellentwicklung wurden unterschiedliche Ansätze analysiert und miteinander verglichen (vgl. Gramzow, Riese & Reinhold, 2013). Zunächst wurden bereits existierende Konzeptualisierungen wie beispielsweise die Projekte ProWin (Borowski, Tepner, Fischer & Sumfleth, 2011), KIL (Kröger, Euler, Neumann, Härtig & Petersen, 2012), COACTIV (Kunter, Baumert, Blum, Klusmann, Krauss & Neubrand, 2011) und MT21 (Blömeke, Kaiser & Lehmann, 2008) herangezogen. Theoretische Ansätze zum FDW und PCK wie beispielsweise von Shulman (1987), Magnusson, Krajcik & Borko (1999) oder Park & Oliver (2008) lieferten zusätzlich hilfreiche Anhaltspunkte für die eigene Modellierung. Auch normative Setzungen wie beispielsweise die Bildungsstandards (KMK, 2008) oder das Quereinsteiger-Programm PD-Q (Korneck, Lamprecht, Wodzinski & Schecker, 2010) wurden miteinbezogen. Eine ausführlichere Beschreibung des Modells findet sich bei Gramzow, Riese und Reinhold (2013).

Das entwickelte Modell musste aufgrund von testtheoretischen Gründen gekürzt und auf den physikalischen Inhalt der Mechanik fokussiert werden. Es enthält daher die vier fachdidaktischen Facetten „Instruktionsstrategien“, „Schülvorstellungen“, „Experimente und Vermittlung eines angemessenen Wissenschaftsverständnisses“ sowie „Fachdidaktische Konzepte“. Des Weiteren wurde dem Modell für die Itementwicklung als dritte Dimension die Kognitive Aktivität hinzugefügt, die die Anforderungen der jeweiligen Items in die Bereiche „Reproduzieren“, „Analysieren“ und „Anwenden“ unterteilt. Anhand dieses Modells für die Itementwicklung wurden 91 Items konzipiert, validiert und anschließend pilotiert.

**Ergebnisse:** Im Rahmen der Pilotstudie konnten 216 Personen aus 12 Universitäten und Fachhochschulen befragt werden. Die Wrightmap auf Grundlage einer Raschanalyse zeigt, dass der Test zu schwer und die Varianz zu gering ist (.33). Der Itemfit ist bei 89 von 91 Items jedoch gut ( $0.8 < \text{MNSQ} < 1.2$  und  $-1.9 < T < 1.9$ ), ebenso die EAP-Reliabilität für den gesamten Test (.84).

Der Modellvergleich zwischen einem 1D-Rasch-Modell und einem 4D-Rasch-Modell, das in die fachdidaktischen Facetten „Instruktionsstrategien“, „Schülvorstellungen“, „Experimente und Vermittlung eines angemessenen Wissenschaftsverständnisses“ sowie „Fachdidaktische Konzepte“ aufspaltet, zeigt, dass das 4D-Modell hochsignifikant besser passt ( $\chi^2$ -Test). Eine Trennung der vier Facetten ist somit auch statistisch vertretbar.

**Zusammenfassung und Ausblick:** Im Projekt wurde ein Modell sowie ein Test entwickelt und validiert, der eine modellkonforme innere Struktur des Fachdidaktischen Wissens in Physik empirisch abbilden kann. Das Instrument ist jedoch zu schwer und in seiner ursprünglichen Form zu lang. Für die Haupterhebung im WS13/14 wurde eine Testkürzung auf insgesamt 65 Minuten vorgenommen. Die Itemauswahl erfolgte aufgrund von Erkenntnissen aus den Validierungsstudien und anhand statistischer Merkmale. Im Vortrag werden die Ergebnisse der Pilotierung und erste Analysen aus der Haupterhebung vorgestellt.

Baumert, J., Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9(4), 469–520.

Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.) (2008). Professionelle Kompetenz angehender Lehrerinnen und Lehrer – Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und –referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung. Münster: Waxmann Verlag.

Borowski, A., Tepner, P., Fischer, H.E., Sumfleth, E. (2011). Professionswissen in den Naturwissenschaften (ProwiN). Untersuchungsdesign und verwendete Aufgabenmodelle. In Hötter, Dietmar (Hg.), Naturwissenschaftliche Bildung als Beitrag zur Gestaltung partizipativer Demokratie. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Potsdam 2010. Berlin: LIT Verlag.

Gramzow, Y.; Riese, J.; Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. ZfDN (Zeitschrift für Didaktik der Naturwissenschaften), 19, S. 7-30

Helmke (2009). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze: Klett-Kallmeyer.

Korneck, F.; Lamprecht, J.; Wodzinski, R. & Schecker, H. (2010). Quereinsteiger in das Lehramt Physik - Lage und Perspektiven der Physiklehrerausbildung in Deutschland. Deutsche Physikalische Gesellschaft (DPG).

Kröger, J.; Euler, M.; Neumann, K.; Härtig, H. & Petersen, S. (2012). Messung Professioneller Kompetenz im Fach Physik. In Bernholt, S. (Hg.), Konzepte fachdidaktischer Strukturierung für den Unterricht. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Oldenburg 2011. (616–618). Berlin: LIT Verlag.

Kultusministerkonferenz (KMK) (2008). Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerinnen- und Lehrerbildung (Beschluss der Kultusministerkonferenz vom 16. Oktober 2008). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., Neubrand, M. (Hrsg.) (2011). Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann-Verlag.

Kulgemeyer, C.; Borowski, A.; Fischer, H.; Gramzow, Y.; Reinhold, P.; Riese, J.; Schecker, H.; Tomczynski, E.; Walzer, M. (2012). Profile-P – Professionswissen in der Lehramtsausbildung Physik. Vorstellung des Forschungsprojekts. PhyDid B – Beiträge zur DPG-Frühjahrstagung.

Lee, E. & Luft, J.A. (2008). Secondary Science Teachers' Representation of Pedagogical Content Knowledge. International Journal of Science Education, 30 (10), pp. 1343-1363.

Magnusson, S., Krajcik, L., & Borko, H. (1999). J. Nature, Sources, and Development of Pedagogical Content Knowledge for Science Teaching. In J. Gess-Newsome & N.G. Lederman (Eds.), Examining pedagogical content knowledge (pp. 95-132). Dordrecht, Netherlands: Kluwer Academic Publishers.

Park, S. & Oliver, J.S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. Research in Science Education, 38 (3), 261-284.

Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. Dissertation. Berlin: Logos Verlag.

Schmelzing (2010). Das fachdidaktische Wissen von Biologielehrkräften: Konzeptualisierung, Diagnostik, Struktur und Entwicklung im Rahmen der Biologielehrerbildung. Dissertation. Berlin: Logos Verlag.

Shulman, L. (1987). Knowledge and teaching: foundations of the new reform. Harvard Education Review, 57, 1-22.

Vogelsang, C. & Reinhold, P. (2010). Handlungsvalidierung eines Instruments zur Kompetenzdiagnose. In Hötter, Dietmar (Hg.), Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Dresden 2009. Berlin: LIT Verlag.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G150



## Kognitive und motivationale Determinanten schulischer Leistungen: Ein personen-zentrierter Ansatz

Zur Erklärung interindividueller Unterschiede in akademischen Leistungen werden neben bewährten kognitiven Konstrukten wie der Intelligenz häufig motivationale Faktoren untersucht. Die wenigen Studien, die sich dabei der relativen Bedeutsamkeit kognitiver und motivationaler Variablen widmeten (z.B. Helmke, 1992; Spinath, Spinath, Harlaar & Plomin, 2006; Steinmayr & Spinath, 2009), zeigten für verschiedene motivationale Konstrukte einen von der Intelligenz unabhängigen Anteil erklärter Varianz in der Schulleistung auf. Trotz des unbestrittenen Wertes dieser Studien, unterliegen die Studien in diesem Feld der Einschränkung eines variablen-zentrierten Ansatzes. Wodurch die Frage offen bleibt, wie sich die unterschiedlichen Ausprägungen kognitiver und motivationaler Variablen je Schüler zu verschiedenen persönlichen Profilen zusammensetzen. Anhand eines personen-zentrierten Ansatzes kann das Zusammenspiel kognitiver und motivationaler Variablen betrachtet werden. Während die prognostische Bedeutsamkeit einzelner kognitiver und motivationaler Konstrukte für die Schulleistung vielfach aufgezeigt wurde, mangelt es an Erkenntnissen zur Erklärung interindividueller Leistungsunterschiede durch verschiedene Profile kognitiver und motivationaler Variablen. Unserem Wissen nach existiert nur eine Studie (Reimann et al., 2013), die sich dieser Fragestellung widmete. Der Fokus von Reimann und Kollegen (2013) lag auf Profilanalysen über verschiedene Subtests der Intelligenz bei gleichzeitiger Berücksichtigung von zwei motivationalen Konstrukten (Persistenz, intrinsischer Motivation). Dabei wurde auf eine vorselektierte Stichprobe von Vor- und Grundschulern zurückgegriffen (i.e. Auswahl von Schülern mit hoher intraindividuelle Varianz in den Subtestwerten). In der vorliegenden Arbeit soll die Fragestellung an einer unausgewählten Stichprobe von Gymnasiasten und Realschülern überprüft werden, wobei neben kognitiven Variablen eine Reihe von bewährten motivationalen Konstrukten herangezogen wird. Ziel der vorliegenden Arbeit ist zu untersuchen a) wie viele und welche verschiedenen Profile kognitiver und motivationaler Schülermerkmale zu unterscheiden sind und b) wie diese Profile mit schulischen Leistungen zusammenhängen.

Bei 463 Schülern ( $M = 13.64$  Jahre,  $SD = .62$ , 226 Mädchen) der 8ten Jahrgangsstufe von 4 bayrischen Gymnasien und 3 Realschulen wurden allgemeine Intelligenz (CFT 20-R; Weiß, 2006), Konzentrationsfähigkeit (d2-R; Brickenkamp, Schmidt-Atzert & Liepmann, 2010), Leistungsmotive (AMS; Göttert & Kuhl, 1980) sowie domänen-spezifisches Fähigkeitsselbstkonzept (SESSKO; Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2002) und subjektive Werte in Mathematik (SESSW; Steinmayr & Spinath, 2010) als Indikatoren der Profile herangezogen. Die Leistung in einem standardisierten Mathematikleistungstest (TIMSS; Baumert et al., 1998) und die Mathematiknote wurden als Kriterien erfasst.

Bei den Latent-Profil-Analysen (MPLUS) wurde unter Berücksichtigung des besten Modell-Fits und der Interpretierbarkeit der Ergebnisse eine 3-Klassenlösung aufgezeigt. Die größte Subgruppe von Schülern (52%) zeigte durchschnittliche Werte in allen kognitiven und motivationalen Variablen (bezeichnet als „durchschnittliche“). Die „hochleistungsmotivierte“ Profil-Gruppe (21%) zeigte ebenfalls durchschnittliche Werte in den kognitiven Leistungen (Intelligenz und Konzentrationsfähigkeit) und kennzeichnet sich zudem durch allgemein hohe Werte in der Leistungsmotivtendenz Hoffnung auf Erfolg, im Fähigkeitsselbstkonzept und den subjektiven Werten in Mathematik, demgegenüber lagen in dem Motiv „Furcht vor Misserfolg“ niedrige Werte vor. Für die Gruppe mit dem Profil „niedrigleistungsmotivierte“ (27%) bestand das entgegengesetzte Bild: neben durchschnittlichen kognitiven Leistungen (bzw. einem leicht reduzierten Wert für Intelligenz) zeigte die Subgruppe allgemein niedrige Werte in Hoffnung auf Erfolg, Fähigkeitsselbstkonzept und subjektiven Werten, während bei Furcht vor Misserfolg hohe Werte bestanden. Im Hinblick auf den Zusammenhang zur Schulleistung bestanden bedeutsame Unterschiede zwischen den verschiedenen Profilen. Während „durchschnittliche“ Schüler durchschnittliche Leistungen in einem standardisierten Mathematikleistungstest und Mathematiknoten aufzeigten, konnten für „hochleistungsmotivierte“ Schüler bedeutsam höhere sowie für „niedrigleistungsmotivierte“ Schüler bedeutsam niedrigere Testleistung und Noten aufgezeigt werden. Die Ergebnisse werden in Hinblick auf die Bedeutsamkeit des Zusammenspiels von Determinanten schulischen Leistungsverhaltens diskutiert.

Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J. et al. (1998). Testaufgaben Mathematik TIMSS 7./8. Klasse (Population 2). Berlin: Max-Planck-Institut für Bildungsforschung. Zugriff am 26.11.2012 von [http://library.mpib-berlin.mpg.de/dl/Materialien/Materialien\\_060/pdf/Materialien\\_Bildungsforschung\\_MPIB\\_060.pdf](http://library.mpib-berlin.mpg.de/dl/Materialien/Materialien_060/pdf/Materialien_Bildungsforschung_MPIB_060.pdf)

Brickenkamp, R., Schmidt-Atzert, L. & Liepmann, D. (2010). Aufmerksamkeits- und Konzentrationstest : d2 – R. Göttingen: Hogrefe.

Göttert, R. & Kuhl, J. (1980). AMS — Achievement Motives Scale von Gjesme und Nygard - Deutsche Fassung. In F. Rheinberg & S. Krug (Hrsg.), Motivationsförderung im Schulalltag (S. 194–200). Göttingen: Hogrefe.

Helmke, A. (1992). Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe.

Reimann, G., Stoecklin, M., Lavallee, K., Gut, J., Frischknecht, M.-C. & Grob, A. (2013). Cognitive and motivational profile shape predicts mathematical skills over and above profile level. *Psychology in the Schools*, 50, 37–56.

Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2002). Skalen zur Erfassung des schulischen Selbstkonzepts (SESSKO). Göttingen: Hogrefe.

Spinath, B., Spinath, F. M., Harlaar, N. & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34, 363–374.

Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, 19, 80–90.

Steinmayr, R. & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica*, 56, 195–211.

Weiß, R. H. (2006). Grundintelligenztest Skala 2 – Revision (CFT 20-R). Göttingen: Hogrefe.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G150

Chair(s): Julia Warwas (Otto-Friedrich-Universität Bamberg), Jochen Wissinger (Justus-Liebig-Universität Gießen), Stefan Brauckmann (Universität Erfurt)

Im Rahmen einer outputorientierten Steuerung, die operative Gestaltungsfreiheiten der Einzelschulen mit dem Kontrollelement dezentraler Ergebnisverantwortung verbindet, wurden Entscheidungsbefugnisse und Funktionsumfang der Schulleitung erheblich gestärkt (z.B. Brauckmann, 2012). Empirische Überprüfungen der bildungspolitischen intendierten ‚Wirksamkeit‘ der Schulleitung im Sinne einer Steigerung einzelschulischer Leistungsparameter sind im deutschen Sprachraum noch selten (zsf. Wissinger, 2011; Warwas, 2013). Im angloamerikanischen Bereich existiert hierzu zwar eine lange Forschungstradition, jedoch verweisen die Befunde konsistent auf indirekte Effekte, die über arbeitsorganisatorische Gestaltungselemente der Einzelschule sowie über die pädagogische Praxis ihrer Lehrkräfte vermittelt werden (zsf. Hallinger, 2011). Dementsprechend scheint es geboten, empirische Analysen zum ‚Erfolgsfaktor Schulleitung‘ nicht auf die Betrachtung direkter Zusammenhänge mit indikatorisierten Schülerleistungen zu beschränken. Ergänzend sollte dem innerschulischen ‚Wirken‘ der Schulleitung differenziert nachgespürt werden, wobei theoriegeleitet Merkmale der Einzelschulen als Organisationen, als professionelle Gemeinschaften und soziale Erfahrungsräume zu betrachten sind. Zu fragen ist, worin sich schulisches Leitungshandeln konkretisiert, welche (den Schülerleistungen vorgelagerte) Merkmale der Lehrerarbeit und Schulqualität es gestaltend beeinflusst, und unter welchen individuellen und organisationalen Bedingungen bestimmte Handlungsformen bevorzugt auftreten. Das Symposium setzt in diesem Bemühen folgende Akzente:

1. Mittels arbeits- und organisationspsychologischer Erklärungsansätze und Instrumente eruieren Windlinger und Hostettler, welchen Einfluss Leitungskräfte auf leistungsrelevante Merkmale des Kollegiums (z.B. School Organizational Citizenship Behaviour) sowie einzelner Lehrkräfte (z.B. Arbeitszufriedenheit) ausüben. Daten einer repräsentativen Befragung an 181 Schulen des Schweizer Kantons Bern wurden hierzu mit Strukturgleichungs- und Mehrebenenmodellen ausgewertet.
2. Gestützt auf Konzepte der Führungs- und Werteforschung überprüft Warwas die konkurrierenden Thesen der Kontingenz- und Wertegebundenheit schulischen Führungshandelns. Durch Triangulation fragebogenbasierter Daten des Leitungs- und Lehrpersonals von 56 Schulen in Bayern analysiert sie die separaten und gemeinsamen Beiträge professioneller Werte und organisationaler Handlungsbedingungen für die Vorhersage grundlegender Dimensionen der Führungspraxis.
3. Internationale Forschungsarbeiten zu School Leadership Effects und Distributed Leadership liefern die konzeptionellen Grundlagen der Studie von Graf u.a. Angesiedelt im Projekt VERA 8 erlauben Längsschnittdaten Einblicke in einzelschulische Implementierungsstrategien qualitätssichernder Maßnahmen im Anschluss an Lernstandserhebungen sowie in dabei realisierte Verlagerungen von Führungsaufgaben in die Fachbereiche. Im Beitrag steht die teststatistische Güte der Erhebungsinstrumente im Vordergrund.
4. Aus dem Blickwinkel der Professionalisierungsforschung rekonstruieren Wagner und Rückmann in einer Mixed Methods-Studie an 14 beruflichen Schulen in Berlin Ausprägungsvarianten des professionellen Selbstverständnisses der Leitungskräfte. Die in Gruppendiskussionen gewonnenen Ergebnisse werden denjenigen Funktionsanforderungen gegenübergestellt, die sich inhaltsanalytisch aus Dokumenten der innerschulischen Qualitätssteuerung erschließen lassen. Zudem werden sie mit fragebogenbasierten Einschätzungen der Lehrkräfte zu wahrgenommenen Leitungsimpulsen für Qualitätsentwicklungsprozesse kontrastiert.

Brauckmann, S. (2012). Schulleitungshandeln zwischen erweiterten Rechten und Pflichten (SHaRP) im internationalen Forschungskontext. In S. Hornberg & M. Parreira do Amaral (Hrsg.), *Deregulierung im Bildungswesen* Münster (S. 223-248). Münster u.a.: Waxmann.

Hallinger, P. (2011). Leadership for learning: lessons from 40 years of empirical research. *Journal of Educational Administration*, 49(2), 125-142.

Warwas, J., & Tenberg, R. (2013). Assessing Research on School Leadership in Germany from an International Perspective. In K. Beck & O. Zlatkin-Troitschanskaia (Eds.), *From Diagnostics to Learning Success. Proceedings in Vocational Education and Training* (pp. 269-280). Rotterdam, Boston, Taipei: Sense Publishers.

Wissinger, J. (2011). Schulleitung und Schulleitungshandeln. In E. Terhart, H. Bennewitz & M. Rothland (Eds.), *Handbuch der Forschung zum Lehrerberuf* (S. 98-115). Münster u.a.: Waxmann.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 0.251

**Aspekte der Wirkung von Schulleitungshandeln am Beispiel eines großen Schweizer Kantons**

Theoretischer Hintergrund: Die Schulleitung ist eine entscheidende Einflussgröße im System (z.B. Huber und Muijs, 2010; Robinson et al., 2008; Leithwood et al., 2004, S. 70). Ihr Handeln beeinflusst die Qualität der Schule indirekt, schafft gute Arbeitsbedingungen für Lehrpersonen und fokussiert auf guten Unterricht. Zur Wirksamkeit von Schulleitung, Effektivität und Qualität von Schule existiert eine Fülle an wissenschaftlicher Literatur, zumeist aus der angelsächsischen Welt. Im deutschsprachigen Raum ist Forschung noch wenig ausgeprägt (Bonsen, 2006, S. 195). Als Indikator von Qualität steht in den meisten Untersuchungen die Verbesserung der Lernleistung von Schüler/-innen im Vordergrund (Huber und Muijs, 2010). Es besteht Einigkeit, dass die Wirkung von Schulleitungshandeln dabei indirekt über die Lehrpersonen als Einzelpersonen und als Mitglieder professioneller Gemeinschaften erfolgt und das Klima, die Arbeitszufriedenheit oder die individuelle und kollektive Selbstwirksamkeitserwartung beeinflusst. Diese Variablen sind gleichzeitig Outcomes und Mediatoren (Day & Leithwood, 2007). Wichtige Erkenntnisse dazu, wie diese Variablen durch die Arbeitsgestaltung und das Erleben von Arbeit und deren Rahmenbedingungen beeinflusst werden und wie sie mit weiteren Outcomes wie Gesundheit oder Leistung zusammenhängen, stammen aus der Arbeits- und Organisationspsychologie (z.B. Ulich, 2005). Dabei geht es auch um die Frage, wie Führungspersonen direkt (z.B. durch individuelle Unterstützung) oder indirekt (über die Veränderung von Rahmenbedingungen) darauf Einfluss nehmen können (z.B. Felfe, 2007). Für die Beschreibung der Einflussnahme der Schulleitung werden verschiedene Modelle diskutiert, z.B. „instructional leadership“, „transformational leadership“ und zunehmend „distributed leadership“ (Huber, 2009; Leithwood & Duke, 1999). Die Forschung zeigt, dass gerade eine Kombination dieser unterschiedlichen Formen wirksam ist (Leithwood et al., 2006). Schulleitende sind mit ihren Handlungen besonders erfolgreich, wenn sie den Zustand ihrer Schule „lesen“, also situations- und kontextspezifisch handeln (Bonsen, 2003; Johanson & Moos, 2009).

Fragestellung: Die Schulwirklichkeit der rund 500 Schulen des Kantons Bern (deutschsprachiger Teil), die seit 1996 „geleitet“ sind, unterscheidet sich strukturell, historisch und soziokulturell und bezüglich Umsetzungsgrad und -qualität der geleiteten Schule, den daraus erfolgten Veränderungen des Schulalltags sowie der Wirkung von Schulleitungshandeln. Die vom Schweizerischen Nationalfonds geförderte Studie (<http://p3.snf.ch/project-136877>) untersucht im Kanton Bern mit einer umfassenden Fragestellung, wie Schulleitungshandeln, Schulqualität und Schulkontext zusammenhängen. In diesem Beitrag sollen davon nun insbesondere folgende (Teil-) Fragen beleuchtet werden: Welche Wirkung hat ein bestimmtes Schulleitungshandeln auf kollegiumsbezogene Variablen (Kollektive Selbstwirksamkeit, Klima, School Organizational Citizenship Behavior) und auf individuelle Variablen (Individuelle Selbstwirksamkeit, Arbeitszufriedenheit)? Sind diese Wirkungen direkt oder durch andere Variablen vermittelt (Mediation)? Methode: In einer repräsentativen Stichprobe von 181 Schulen wurden Schulleitende und Lehrpersonen mittels quantitativer Erhebungsinstrumente schriftlich befragt. Die Schulleitenden machten Angaben zu persönlichen Merkmalen (Alter, Erfahrung, Pensum, Ausbildung etc.), ggf. zur Zusammenarbeit im Schulleitungsteam, zum internen Kontext der Schule (Größe, Ausstattung etc.), zum externen Kontext (Ausmaß der Schulautonomie). Weiter machten sie Einschätzungen zu ihrem beruflichen Selbstverständnis, der beruflichen Belastung, der Arbeitszufriedenheit (vgl. Warwas, 2012) und ihrer Arbeitszeit. Die Lehrpersonen schätzten das Führungsverhalten ihrer Schulleitung ein, die Zusammenarbeit im Schulleitungsteam, ihre eigene Selbstwirksamkeit in Bezug auf ihre Tätigkeit als Lehrperson, ihre Arbeitszufriedenheit sowie das Klima der Schule und kollegiumsbezogene Faktoren (Kollektive Selbstwirksamkeit, School Organizational Citizenship Behavior). Die Daten aus der eigenen Befragung wurden mit bestehenden Kontextinformationen (Sekundärdaten) verknüpft. Neben deskriptiven Auswertungen wurden multivariate Analysen durchgeführt (Regressionen, Strukturgleichungsmodelle, Mehrebenenanalysen).

Ergebnisse: Die Auswertungen zeigen positive Zusammenhänge zwischen transformationalem und partizipativem Führungsverhalten der Schulleitung und der Kollektiven Selbstwirksamkeit, dem School Organizational Citizenship Behavior, dem Klima, der Arbeitszufriedenheit der Lehrpersonen sowie deren individueller Selbstwirksamkeit. Erste Analysen deuten darauf hin, dass der Einfluss des Führungsverhaltens der Schulleitung auf die individuelle Selbstwirksamkeit der Lehrpersonen indirekt ist, nämlich vermittelt über die Kollektive Selbstwirksamkeit.

Bonsen, M. (2003): Schule, Führung, Organisation. Eine empirische Studie zum Organisations- und Führungsverständnis von Schulleiterinnen und Schulleiter. Münster: Waxmann.

Bonsen, M. (2006): Wirksame Schulleitung. Forschungsergebnisse. In: H. Buchen und H.-G. Rolff (Hg.): Professionswissen Schulleitung. Weinheim und Basel: Beltz, S. 193–228.

Bonsen, M.; van der Gathen, J.; Iglhaut, C.; Pfeiffer, H. (2002): Die Wirksamkeit von Schulleitung. Empirische Annäherungen an ein Gesamtmodell schulischen Leitungshandeln. Weinheim, München: Juventa.

Day, C.; Leithwood, K. (Hg.) (2007): Successful school principal leadership in times of change: International Perspectives. Dordrecht: Springer.

Felfe, J. (2006). Transformationale und charismatische Führung - Stand der Forschung und aktuelle Entwicklungen. Zeitschrift für Personalpsychologie, 5, S. 163–176.

Huber, S. G. (2009): Schulleitung. In: S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtasik und W. Scher (Hg.): Handbuch Schule. Theorie - Organisation - Entwicklung. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 502–511.

Huber, S. G.; Muijs, D. (2010): School Leadership Effectiveness. The Growing Insight in the Importance of School Leadership for the Quality and Development of Schools and Their Pupils. In: S. G. Huber, R. Saravanabhavan und S. Hader-Popp (Hg.): School Leadership - International Perspectives. Dordrecht: Springer, S. 57–77.

Johanson, O.; Moos, L. (Hg.) (2009): Special Issue. Sustaining Successful School Leadership. Journal of Educational Administration 47(6).

Leithwood, K.; Day, C.; Sammons, P.; Harris, A.; Hopkins, D. (2006): Seven strong claims about successful school leadership. Hg. v. National College for School Leadership.

Leithwood, K.; Duke, D. L. (1999): A Century's Quest for a Knowledge Base, 1976–1998. In: J. Murphy und K. S. Louis (Hg.): Handbook of Research on Educational Policy. 2nd Edition. San Francisco: Jossey-Bass, S. 45–72.

Leithwood, K.; Seashore Louis, K.; Anderson, S.; Wahlstrom, K. (2004): Review of Research. How Leadership Influences Student learning. University of Minnesota, Center for Applied Research and Educational Improvement. Minneapolis.

Robinson, V.; Lloyd, C.; Rowe, K. (2008): The Impact of Leadership on Student Outcomes. An Analysis of the Differential Effects of Leadership Types. Educational Administration Quarterly 44, S. 635–674.

Ulich, E. (2005). Arbeitspsychologie (6. Aufl). Stuttgart: Schäffer-Poeschel.

Warwas, J. (2012): Berufliches Selbstverständnis, Beanspruchung und Bewältigung in der Schulleitung. Dissertation Otto-Friedrich-Universität Bamberg, 2011. Wiesbaden: Springer.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 0.251





## ***Führungshandeln von Schulleiter(inne)n – wertegeleitet oder kontingent?***

Vor allem in konzeptionellen Beiträgen wird die konsequente Ausrichtung schulischen Führungshandelns an den professionellen Werten der jeweiligen Leitungsperson wiederkehrend betont (z.B. Hodgkinson, 1978; Harris & Johnston, 2010). Begründet wird dies u.a. mit unablässigen Entscheidungserfordernissen in der Leitungspraxis angesichts administrativer Regulierungslücken und der psychologischen Filterfunktion individueller Werthaltungen als abstrakte Vorstellungen wünschenswerter Wege, Mittel und Ergebnisse des Handelns. Empirisch unterfüttert ist die These der Wertebundenheit schulischer Führung bislang jedoch primär durch wenige ethnografische Einzelfallstudien und interviewbasierte Selbstauskünfte der Leitungskräfte. Vergleichsweise häufiger und mit mehr generalisierbaren Befunden erforscht ist die konkurrierende These einer hochgradig adaptiven, an den organisationalen Handlungskontext angepassten Führungspraxis. Zu den untersuchten Kontingenzfaktoren gehören etwa die Schul- bzw. Kollegiumsgröße sowie die Existenz formaler Steuerungsinstrumente (z.B. Clarke & Wildy, 2004). Nur selten wurden aber mögliche Ausprägungen einer ‚Values-based contingency leadership‘ (Day, Harris, & Hadfield 2001, S. 52) ergründet, in denen Zusammenhänge zwischen Werten und Handeln durch organisationale Rahmenbedingungen moderiert werden (vgl. auch Sagie, Elizur, & Koslowsky, 1996).

Der vorliegende Beitrag widmet sich deshalb der Frage, welche eigenständigen und gemeinsamen Erklärungsbeiträge verbalisierte professionelle Werte von Schulleiter(inne)n und ausgewählte Organisationsmerkmale ihrer Schulen für das vom jeweiligen Kollegium wahrgenommene Führungshandeln liefern. Den statistischen Analysen liegen Daten von 56 allgemein- und berufsbildenden Schulen in Bayern zugrunde. Von den Führungskräften wurden (jeweils drei) frei formulierte Begriffe, die wesentliche Leitwerte ihrer professionellen Tätigkeit charakterisieren, erbeten. Zudem machten sie geschlossene Angaben zu diversen Organisationsmerkmalen (Schulgröße, Unterrichtsverpflichtung der Leitung, personelle Stärke des erweiterten Leitungsteams, Einsatz von Instrumenten des schulischen Qualitätsmanagements, Ausmaß schülerbezogener Probleme an der Schule). Die Lehrkräfte der an der Studie beteiligten Schulen (durchschnittliche Rücklaufquote 52%, SD = 11) beurteilten vier generische Dimensionen des Führungshandelns in einem geschlossenen, sechsstufigen Antwortformat (strukturelle, interpersonelle, politisch-symbolische und pädagogische Führung; vgl. z.B. Fidler 1997). Intraklassenkoeffizienten lassen für jede Skala substantielle Schulunterschiede des beurteilten Führungshandelns sowie gute Beobachterübereinstimmungen erkennen. Zu Analyse Zwecken eignen sich sowohl dimensionsorientierte Verfahren (mit dem Ziel, generelle Effekte inhaltlich trennbarer Wertekategorien abzubilden; z.B. Bardi & Schwartz, 2003) als auch typologische Verfahren (um Zusammenhänge zwischen unterschiedlichen, personengebundenen Werteprofilen und dem Führungshandeln aufzudecken; z.B. Law, Dimmock & Walker, 2003). Zunächst wurden daher die Aussagen der Leitungskräfte inhaltsanalytisch zu Wertekategorien verdichtet. Hierdurch entstanden gestützt auf konzeptionelle Beiträge zur wertebundenen Schulführung vier Hauptkategorien: (1) an Schulergebnissen orientierte Werte, (2) an der Qualität des Schullebens orientierte Werte, (3) an der Qualität pädagogischer und operativer Arbeitsprozesse orientierte Werte, (4) an operativen Voraussetzungen und persönlichen Eigenschaften orientierte Werte. Ihnen konnten insgesamt 16 Unterkategorien zugeordnet werden.

Erste Ergebnisse auf Basis hierarchischer Regressionsanalysen mit Interaktionseffekten, in die sowohl die ermittelten Wertekategorien als auch die erhobenen Kontextfaktoren eingespeist wurden, verweisen auf eine nur marginale prädiktive Kraft professioneller Werte. Sie stützen für die Mehrheit der erfassten Kontextfaktoren deutlich die Kontingenzthese schulischen Führungshandelns. Ergebnisse der typologischen Auswertungsstrategie stehen zum Zeitpunkt der Abstract-Einreichung noch aus.

Bardi, A., & Schwartz, S. H. (2003). Values and Behavior: Strength and Structure of Relations. *Personality and Social Psychology Bulletin*, 29(10), 1207-1220.

Clarke, S., & Wildy, H. (2004). Context counts. Viewing small school leadership from the inside out. *Journal of Educational Administration*, 42(5), 555-572.

Day, C., Harris, A., & Hadfield, M. (2001). Challenging the orthodoxy of effective school leadership. *International Journal of Leadership in Education*, 4(1), 39-56.

Fidler, B. (1997). School Leadership: some key ideas. *School Leadership & Management*, 17(1), 23-37.

Harris, M., & Johnston, M. A. (2010). Values-Based Leadership: At the Heart of Sustaining School Reform and Improvement. In S.G.H. Huber, R.C.-H. Saravanabhavan & S.C.-H. Hader-Popp (Eds.), *School Leadership – International Perspectives* (pp. 145-159). Dordrecht, Heidelberg, London, New York: Springer.

Hodgkinson, C. (1978). *Towards a Philosophy of Administration*. Oxford: Basil Blackwell.

Law, L. Y. S., Walker, A., & Dimmock, C. (2003). The influence of principals' values on their perception and management of school problems. *Journal of Educational Administration*, 41(5), 498-523.

Sagie, A., Elizur, D., & Koslowsky, M. (1996). Work values: a theoretical overview and a model of their effects. *Journal of Organizational Behavior*, 17, 503-514.



Tanja Graf (Freie Universität Berlin), Barbara Muslic (Freie Universität Berlin), Harm Kuper (Freie Universität Berlin), Stefan Brauckmann (Universität Erfurt)

## **Zusammenhang von Schulleitungshandeln und Schulleistung – Entwicklung eines Instruments zur Erfassung von Schulleitungshandeln in Folge zentraler Lernstandsergebnisse**

Zentrale Lernstandserhebungen als Element eines neuen Steuerungsparadigmas, das durch Standard-, Evidenz- und Outputorientierung gekennzeichnet ist (Altrichter & Maag Merki, 2010), richten sich direkt an Schulen sowie schulische Akteure und sollen zur Schul- und Unterrichtsentwicklung beitragen (Diemer & Kuper, 2011). Für Schulleitungen – als wichtigste Akteure für die Implementation von Reformen sowie evidenzbasierte Steuerung auf Einzelschulebene (u.a. Rosenbusch, 2005, Fend, 2008, May & Supovitz, 2010) – resultieren aus den Modellen der neuen Steuerung zusätzliche Kompetenzanforderungen und neue Verantwortungsbereiche (Böttcher, 2002, Buchen & Rolff, 2009). Allerdings stellt die Verwendung von Ergebnissen aus zentralen Lernstandserhebungen bislang für die Schulleitungsforschung ebenso ein Desiderat dar wie die besondere Verantwortung der Schulleitungen in Maßnahmen der testbasierten Schulreform für die Rezeptionsforschung.

Im Rahmen der hier vorgestellten Wirkungsstudie wird der Zusammenhang zwischen Schulleitungsaktivitäten in Folge zentraler Lernstandserhebungen und Schulleistungen untersucht. Theoretisch wird dabei auf die Forschung zur Effizienz von Schulleitungen zurückgegriffen, die den Zusammenhang zwischen den Fachleistungen der Schülerschaft einer Schule und der Schulleitung herausstellt. Hierbei können diesem als school-leadership-effects (zusammenfassend: Scheerens, 2012) innerhalb der internationalen school-effectiveness-research bezeichneten Strang der Schulleitungsforschung wichtige Anknüpfungspunkte entnommen werden. Empirische Befunde aus vorwiegend internationalen Studien weisen auf einen indirekten Effekt von Schulleitungshandeln auf die Leistungsentwicklung von Schüler/-innen hin (Leitner, 1994, Creemers & Reezigt, 1996, Hill, 1998, Bryk et al., 2010). Darüber hinaus wird zur Erfassung der Wirkungen von Schulleitungsaktivitäten auf organisationaler Ebene auf das Modell der distributed leadership zurückgegriffen (u.a. Spillane et al., 2001, Gronn, 2002, Harris, 2004), in dem die Verantwortung der Schulleitung in Relation zu einer organisatorischen Verantwortung der Schule und des Lehrerkollegiums konzipiert wird.

Vor diesem Hintergrund wird im Beitrag der Frage nachgegangen, inwieweit Maßnahmen der Qualitätssicherung infolge der VERA-Ergebnisse initiiert werden und – im Sinne einer distributed leadership – an welche Funktionsstellen der Organisation dabei die Verantwortung verteilt wird.

Eine Fragebogenerhebung mit Schulleiter/-Innen aus den Bundesländern Niedersachsen, Berlin und Brandenburg in einem Längsschnittdesign über zwei Messzeitpunkte bildet die empirische Grundlage der geplanten Studie „Maßnahmen von Schulleitungen in Folge zentraler Lernstandserhebungen und ihre Wirkung auf Schulleistungen“. Dabei stehen die Prozessvariablen (Verantwortungs- und Entscheidungsbereiche von Schulleitungen) im Mittelpunkt, über die Schulleitungen einen mittelbaren Einfluss auf die Unterrichtsqualität und die Leistungsentwicklung der Schüler/-Innen nehmen.

Für die Auswertung der Hauptuntersuchung wird ein auf Schulebene aggregierter Datensatz aus den vorliegenden VERA-8 Daten aus den Schuljahren 2010/11 ff. konstruiert und mit Daten aus Befragungen der Schulleiter/-innen und zusätzlich der Fachbereichsleiter/innen (zur Überprüfung der tatsächlich umgesetzten Maßnahmen durch das Lehrerkollegium) in zwei Wellen verbunden. Die Daten erlauben eine Prüfung der Zusammenhänge mit Regressionsanalysen und Pfadmodellen. Die zum Messzeitpunkt 2013 bei den Schulleitern/innen erhobenen Daten gehen hier als unabhängige Variablen in Regressionsanalysen ein, um die prädiktiven Wirkungen auf die aus Sicht der Fachbereichsleitungen implementierten Maßnahmen (erhoben zum Messzeitpunkt 2014) zu überprüfen.

Das Instrument zur Erfassung der Schulleitungsaktivitäten im Anschluss an VERA-Ergebnisse soll in einem Pretest an Schulen in Niedersachsen auf seine Güte untersucht werden. Hierzu ist eine Zufallsstichprobe (N=100) geplant, die alle öffentlichen Schulen mit einer Sekundarstufe I und Erfahrungen mit VERA 8 einschließt. Im Beitrag werden Ergebnisse des Pretests vorgestellt. Es sollen dabei nach Möglichkeit Aussagen zur Güte und Reichweite des eingesetzten Instruments getroffen werden.

Die Ergebnisse der Schulleiterbefragung sollen zum einen Erkenntnisse darüber liefern, inwieweit Maßnahmen der Qualitätssicherung infolge der VERA-Ergebnisse initiiert werden und – im Sinne einer distributed leadership –, inwiefern Verantwortungsbereiche auf schulische Funktionsstellen und das Lehrerkollegium verlagert werden. In der Hauptuntersuchung sollen unter Hinzuziehen der Schulleistungsdaten zusätzliche Erkenntnisse darüber gewonnen werden, welche Folgen diese Schulleitungsaktivitäten in der Schule zeitigen und welche Effekte auf die Ergebnisse einer Schule in anschließenden Vergleichsarbeiten zu verzeichnen sind.

Altrichter, H. & Maag Merki, K. (Hrsg.) (2010): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften.

Böttcher, W. (2002): Kann eine ökonomische Schule auch eine pädagogische sein? Schulentwicklung zwischen Neuer Steuerung, Organisation, Leistungsevaluation und Bildung. Weinheim/München: Juventa.

Bryk, A. S., Bender Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). Organizing schools for improvement. Lessons from Chicago. Chicago: The University of Chicago Press.

Buchen, H. & Rolff, H.-G. (Hrsg.) (2009): Professionswissen Schulleitung. Weinheim: Beltz.

Creemers, B.P.M., & Reezigt, G.J. (1996). School level conditions affecting the effectiveness of instruction. School Effectiveness and School Improvement, 7 (3), 197-228.

Diemer, T. & Kuper, H. (2011): Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen. In: Zeitschrift für Pädagogik, 56 (4), S. 554-571.

Fend, H. (2008): Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität. Wiesbaden: VS-Verlag.

Gronn, P. (2002): Distributed leadership. In: Leithwood, K. & Hallinger, P. (Eds.): Second international handbook of educational leadership and administration. The Netherlands: Kluwer, 653-696.

Harris, A. (2004): Distributed Leadership and School Improvement. Leading or Misleading? Educational Management Administration & Leadership, 32 (1), 11-24.

Hill, P. (1998): Shaking the foundations: Research-driven school reform. School Effectiveness and School Improvement, 9 (4), 419-436.

Leitner, D. (1994): Do Principals affect student outcomes: An organizational Perspective. School Effectiveness and School Improvement, 5 (3), 219-38.

May, H. & Supovitz, J. A. (2010): The Scope of Principal Efforts to Improve Instruction. Educational Administration Quarterly, 47(2) 332-352.

Rosenbusch, H. S. (2005): Organisationspädagogik der Schule. Grundlagen pädagogischen Führungshandelns. Neuwied: Luchterhand.

Scheerens, J. (2012): School Leadership Effects Revisited. Review and Meta-Analysis of Empirical Studies. Heidelberg, London, New York: Springer.

Spillane, J.P., Halverson R. & Diamond J.B. (2001): Investigating School Leadership Practice: A Distributed Perspective. Educational Researcher 30 (23). 23-28.

GEBF TAGUNG  
2014

G11 c

Symposiums  
beitrag

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 0.251

Wirken und Wirksamkeit der Schulleitung



## **Professionelles Selbstverständnis schulischer Führungskräfte und Impulse für Schulqualität in der „outputorientierten Steuerung“**

Hintergrund: Die schulischen Führungskräfte werden sowohl in der nationalen als auch in der internationalen wissenschaftlichen Diskussion als Schlüsselgrößen für die Schulqualität und Schulentwicklung angesehen. Die aktuelle Ausrichtung der bildungspolitischen Steuerungsansätze und deren verstärkte Orientierung entlang festgelegter Leistungs- und Qualitätsparameter bildet den Rahmen für eine Veränderung von innerschulischen Führungsfunktionen (vgl. z. B. Pont, Nusche, & Moorman, 2008). Dies geschieht in Deutschland vor dem Hintergrund einer nach wie vor defizitären Situation systematischer Qualifizierungsangebote und -verpflichtungen für das Leitungspersonal und führt dazu, dass eher eine unbewusste Übernahme von Führungskonzepten bzw. deren Teilelementen bei den Akteuren stattfindet, die durch erfahrungsgestütztes sowie situations- und kontextgebundenes Wissen zu Handlungsstrategien ergänzt werden. Als ehemalige Lehrkräfte greifen die Führungskräfte nicht selten auf ein kontextspezifisch auf ihre Schule ausgerichtetes professionelles Selbstverständnis zurück. Hieraus entsteht der Anspruch an die Akteure nach einer Veränderung bzw. Erweiterung der eigenen Professionalität. Sie sind gefordert, sich in einem Spannungsfeld von einzelschulischer Interessenwahrnehmung und schulübergreifender Qualitätssteuerung zu positionieren. Dabei ist zu erwarten, dass die auf kontextualisierende Interpretation ausgerichteten Verstehensmuster und epistemologischen Überzeugungen von Leitungskräften mit den auf Dekontextualisierung gründenden standardisierten Steuerungsinstrumenten der Qualitätssicherung und -steigerung, die in der Mehrheit der Bundesländer verwendet werden, konfliktieren.

Fragestellungen: Im Rahmen der vorgelegten Studie sollen deshalb die folgenden Fragen analysiert werden:

(1) Welche Anforderungen an das professionelle Handeln schulischer Führungskräfte werden im Rahmen der schulübergreifenden Qualitätssteuerung formuliert? (2) Wie stellt sich demgegenüber das professionelle Selbstverständnis der Akteure dar und welche übereinstimmenden bzw. abweichenden Orientierungsmuster sind zu identifizieren? (3) Welche Wirkungen entfaltet das Führungshandeln auf den Aufbau schulischer Qualitätsstrukturen und die Beteiligung von Lehrkräften in Qualitätsprozessen und wie sind diese im oben beschriebenen Spannungsfeld von einzelschulischer Interessenwahrnehmung und schulübergreifender Qualitätssteuerung zu bewerten?

Methode: Die Studie ist als Triangulationsstudie mit einer Kombination qualitativer und quantitativer Erhebungsverfahren angelegt. Im Rahmen qualitativer Dokumentenanalysen (Aufgaben- und Funktionsbeschreibungen, Verfahrensvorschriften und Erläuterungen zur schulübergreifenden Qualitätssteuerung) erfolgt eine kriterienorientierte Beschreibung von Anforderungen an das professionelle Handeln schulischer Führungskräfte. Diese werden als analytisches Raster für eine qualitativ-vergleichende, inhaltsanalytische Auswertung der Gruppendiskussionen genutzt. Die Gruppendiskussionen erfassen die Selbstwahrnehmung der schulischen Führungskräfte auf ihr professionelles Selbstverständnis und auf die wahrgenommenen Wirkungen ihres Handelns auf die schulischen Qualitätsprozesse. Die quantitative Befragung der Lehrkräfte ermöglicht weiterhin einen mehrperspektivischen Blick auf die innerschulischen Wirkungen des qualitätsbezogenen Führungshandelns. An den Gruppendiskussionen beteiligten sich N=72 Führungskräfte und an der Fragebogenerhebung N=955 Lehrkräfte aus 14 beruflichen Schulen in Berlin.

Ergebnisse: Zum Einreichungszeitpunkt sind die Auswertungen des Datenmaterials noch nicht abgeschlossen. Erste Befunde verweisen darauf, dass die befragten Führungskräfte im Kontext der aktuell vorherrschenden Steuerungsstrategien Veränderungen in ihrem professionellen Selbstverständnis dahingehend wahrnehmen, dass sie die einzelschulischer Qualitätssicherung und -entwicklung als eine bedeutsame Führungsaufgabe stärker gewichten. Sie sehen in diesem Zusammenhang die Instrumente und Maßnahmen der schulübergreifenden Qualitätssteuerung jedoch eher als formal-organisatorischen Aufwand denn als Orientierungs- oder Umsetzungshilfen, die einer inhaltlichen Auseinandersetzung mit qualitätsbezogenen Themen in der Schule entgegenstehen. Zum Teil werden diese sogar als die Schulqualitätsbemühungen behindernd bzw. sogar gefährdend wahrgenommen. Im Gegensatz zu Befunden aus empirischen Untersuchungen der jüngeren Vergangenheit richten die schulischen Führungskräfte ihr professionelles Selbstverständnis stärker auf ihre direkte Beteiligung und Einflussnahme auf die individuelle Entwicklung und Profilierung der Einzelschule aus; den Managementfunktionen kommt eine eher untergeordnete Bedeutung zu (vgl. Wagner, 2012).

## ***Das Arbeitsgedächtnis in der Grundschule: Variabilität zwischen Kindern und Variabilität über die Zeit verstehen und nutzen***

Chair(s): Judith Dirk (DIPF, IDeA-Zentrum), Tanja Könen (DIPF, IDeA-Zentrum)

Diskutant/in: Claudia Mähler (Universität Hildesheim)

Kognitive Fähigkeiten sind eine zentrale Grundlage für das Lernen und Erbringen von Leistungen in der Schule. Besonders das Arbeitsgedächtnis ist ein etablierter Prädiktor von Schulleistungen im mathematischen (z.B. Swanson, 2011; Swanson, Jerman, & Zheng, 2008) und sprachlichen Bereich (z.B. Alloway & Alloway, 2010; Loosli, Buschkuehl, Perrig, & Jaeggi, 2012) und wird hierbei sogar als eventuell besserer Prädiktor als allgemeine Intelligenz diskutiert (Alloway & Alloway, 2010; Preßler, Könen, Hasselhorn, & Krajewski, 2013). Auch Lern- und Leistungsstörungen gehen oftmals mit Arbeitsgedächtnisdefiziten einher (Schuchard, Mähler, & Hasselhorn, 2008). Arbeitsgedächtnisleistungen beruhen auf dem simultanen und kontrollierten Speichern und Verarbeiten von Informationen (Baddeley & Hitch, 1994).

Da in der Grundschule die Basis für die gesamte schulische Entwicklung gelegt wird, ist die Rolle des Arbeitsgedächtnisses in dieser Zeit von besonderem Interesse und wird von unterschiedlichen Forschungstraditionen betrachtet. In unserem Symposium möchten wir die Breite und Vielfalt des Forschungsfeldes anhand eigener empirischer Arbeiten aufzeigen. Wir ergründen die Variabilität von Arbeitsgedächtnisleistungen aus zwei verschiedenen Perspektiven.

Einerseits betrachten wir Unterschiede innerhalb von Kindern über die Zeit. Schwanken Arbeitsgedächtnisleistungen bedeutsam im Schulalltag? Mit welchen Variablen gehen die Schwankungen einher und unterscheiden sich Kinder in diesen Zusammenhängen? Mit diesen Fragen beschäftigen sich die ersten beiden Vorträge: Dirk und Schmiedek betrachten dabei die Rolle von wahrgenommener Ablenkung für tagtägliche Arbeitsgedächtnisleistungen und Könen, Dirk und Schmiedek die Rolle von Schlafverhalten.

Andererseits untersuchen wir Unterschiede in Arbeitsgedächtnisleistungen zwischen Kindern mit und ohne Lern- und Leistungsstörungen. In welchen Bereichen zeigen Kinder mit Lern- und Leistungsstörungen Defizite? Zeigen sich bei unterschiedlichen Störungen spezifische Defizite und gibt es auch wichtige Gemeinsamkeiten? Diesen Fragen werden im Beitrag von Schuchard und Mähler anhand der Lese-Rechtschreibstörung und der spezifischen Sprachentwicklungsstörung untersucht. Der Beitrag von Göthe, Gendt, Esser und Kliegl hinterfragt, welche Mechanismen Arbeitsgedächtnisleistungen bei Kindern begrenzen. Können diese Mechanismen auch als Erklärung für Defizite bei Kindern mit Lese-Rechtschreibstörungen herangezogen werden?

Unser Symposium kombiniert experimentelle, querschnittliche und mikro-längsschnittliche Arbeiten und verbindet unterschiedliche Traditionen der pädagogischen, kognitiven und entwicklungspsychologischen Forschung. Wir betrachten Unterschiede in der Verfügbarkeit kognitiver Ressourcen zwischen Grundschulkindern sowie innerhalb von Kindern über die Zeit. Dies kann als Grundlage für spezifische Interventionen im frühen Schulkontext nutzbar gemacht werden, die insbesondere auch benachteiligte Kinder durch differentielle Förderempfehlungen in ihrer Lern- und Leistungsfähigkeit unterstützen können.

Alloway, T.P. & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106, 20–29.

Baddeley, A. D. & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485–493.

Loosli, S. V., Buschkuehl, M., Perrig, W. J. & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology*, 18, 62–78.

Preßler, A.-L., Könen, T., Hasselhorn, M. & Krajewski, K. (2013). Cognitive preconditions of early reading and spelling: a latent-variable approach with longitudinal data. *Reading and Writing*. Advance online publication.

Schuchard, K., Mähler, C., & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities*, 41, 514–523.

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, 103, 821–837.

Swanson, H.L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematic problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, 100, 343–379.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 0.254



## ***Beeinträchtigt wahrgenommene Ablenkung die Leistung?***

### ***Die Rolle von Störeinflüssen auf Arbeitsgedächtnisleistung im Schulkontext***

Im Alltag werden kognitive Aufgaben oft unter dem Einfluss von Störeinflüssen wie Ablenkung durch Geräusche und Personen ausgeführt. Zahlreiche Studien haben den negativen Einfluss von Lärm auf Leistung bestätigt und gezeigt, dass Kinder besonders stark von Beeinträchtigungen durch Umgebungsgeräusche betroffen sind, da ihre kognitive Fähigkeiten weniger stark automatisiert und somit anfälliger für Störungen sind als bei Erwachsenen (cf. Elliott, 2002; Klatter, Bergström, & Lachman, 2013). Wenige Studien haben diesen Effekt jedoch in der natürlichen Umgebung von Kindern untersucht und Schwankungen im Störerleben sowie ihren Einfluss auf Leistung betrachtet. Natürliche Lernumgebungen sind von zeitlich variierenden Hintergrundgeräuschen gekennzeichnet. In diesen Umgebungen werden besonders für Kinder starke Einbußen in der Arbeitsgedächtnisleistung erwartet (Elliott, 2002). Darüber hinaus hat eine aktuelle Feldstudie auch den negativen Einfluss von Hintergrundgeräuschen im Klassenraum auf die Motivation und das Wohlbefinden im Schulkontext gezeigt (Klatte, Hellbrück, Seidel, & Leistner, 2010). Unter starken Störeinflüssen wird das Lernen anstrengender und es müssen mehr Ressourcen für das Unterdrücken von Störeinflüssen aufgewendet werden. Dies kann sich mittelfristig negativ auf die Leistungsmotivation und das Wohlbefinden im Klassenkontext auswirken. Basierend auf diesen Befunden sollen im aktuellen Beitrag tagtägliche Schwankungen in Arbeitsgedächtnisleistung und das Erleben von Ablenkung durch Störeinflüsse bei Grundschulkindern im Alltagskontext anhand einer Studie des Projektes FLUX mit ambulanten Messungen auf Smartphones betrachtet werden. Dabei lösten 110 Kinder (8-11 Jahre) vier Wochen lang zu Beginn und am Ende des Unterrichts sowie am Nachmittag Arbeitsgedächtnisaufgaben und beantworteten Fragen. Die erlebte Ablenkung wurde mit vier kurzen Fragen zu Ablenkung durch Geräusche und Personen erfasst. Als weitere mögliche Einflussfaktoren auf Arbeitsgedächtnisleistungen wurden Motivation und Wohlbefinden erhoben. Die Motivation wurde mit sechs kurzen Fragen zu aktueller Anstrengung und Spaß bei der Aufgabenbearbeitung erfasst. Des Weiteren beurteilten die Kinder ihr Wohlbefinden mittels Einschätzungen der aktuellen positiven und negativen Stimmung auf sechs Adjektiven. Die Arbeitsgedächtnisleistungen, die wahrgenommene Ablenkung bei der Aufgabenbearbeitung sowie die Motivation und die Stimmung zeigten systematische tagtägliche Schwankungen. Erste Ergebnisse deuten des Weiteren darauf hin, dass im Mittel an Tagen, an denen Kinder sich stärker abgelenkt fühlen, ihre Arbeitsgedächtnisleistungen schlechter sind. Dieser Beitrag diskutiert diese Ergebnisse im Rahmen existierender Arbeiten zum Einfluss von Störeinflüssen auf kognitive Leistungen und versucht mögliche Mechanismen, die die beobachteten Zusammenhänge erklären können, durch weitere Prädiktoren der Arbeitsgedächtnisleistung im Schulkontext (z.B. Motivation, Stimmung) aufzuklären.

Elliott, E. M. (2002). The irrelevant speech effect and children: Theoretical implications of developmental change. *Memory and Cognition*, 30, 478-487.

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review of noise effects on cognitive performance in children. *Frontiers in Developmental Psychology*, 4:578.

Klatte, M., Hellbrück, J., Seidel, J., & Leistner, P. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment & Behavior*, 42, 659-692.

Tanja Könen (DIPF, IDeA-Zentrum), Judith Dirk (DIPF, IDeA-Zentrum),  
Florian Schmiedek (DIPF, IDeA-Zentrum, Goethe-Universität Frankfurt)

## ***Kognitive Leistungsfähigkeit von Grundschulkindern im Alltag: Welche Rolle spielt das Schlafverhalten?***

Grundschul Kinder begegnen in ihrem Alltag einer Fülle von kognitiven Anforderungen. Sie erlernen zentrale Grundlagen für ihr späteres Leben im Unterricht, bei den Hausaufgaben und auch in der Freizeit. Dabei erscheint es plausibel, dass Kinder bessere und schlechtere Momente und Tage in ihrer kognitiven Leistungsfähigkeit erleben und dass diese Variation nicht für alle Kinder gleich stark sein muss. Neuere Studien zeigen Leistungsschwankungen im Bereich des Arbeitsgedächtnisses (AG) bei Jugendlichen und Erwachsenen (Riediger, Wrzus, Schmiedek, Wagner, & Lindenberger, 2011; Schmiedek, Lövdén, & Lindenberger, 2013). Da sich Grundschul Kinder jedoch in ihrem alltäglichen Lebenskontext und in ihren kognitiven Voraussetzungen maßgeblich von Jugendlichen und Erwachsenen unterscheiden, ist unklar, inwiefern Leistungsschwankungen bei ihnen auftreten und womit diese bei ihnen einhergehen. Ziel des Projektes FLUX ist es, das Ausmaß tagtäglicher Variabilität in AG-Leistungen bei Grundschulkindern zu quantifizieren und in Zusammenhang zu möglichen Mechanismen und Konsequenzen zu bringen. Dafür lösten 110 Grundschul Kinder (8-11 Jahre) vier Wochen lang täglich mehrfach AG-Aufgaben auf Smartphones und beantworteten Fragen zu ihrem aktuellen Erleben und Verhalten. Im Mittel zeigten die Kinder systematische und bedeutsame Leistungsschwankungen zwischen Tagen und auch im Tagesverlauf.

In unserem Beitrag fokussieren wir nun darauf, inwiefern das Schlafverhalten und die Müdigkeit der Kinder im Alltag mit den Leistungsschwankungen assoziiert sind. Experimentelle Studien zeigten, dass bereits eine Schlafreduktion von einer Stunde die AG-Leistung von Kindern reduzieren kann (Sadeh, Gruber, & Raviv, 2003; Vriend et al., 2013). Wenn entsprechende Variationen des Schlafverhaltens auch im Lebensalltag der Kinder vorkommen, kann ein ähnlicher Einfluss auf die AG-Leistungsschwankungen angenommen werden. Im Mittel zeigten die Kinder bedeutsame Schwankungen in der Schlafdauer, der eingeschätzten Schlafqualität und der aktuellen Müdigkeit. Alle drei Aspekte waren prädiktiv für die kognitive Leistungsfähigkeit. Die Schlafdauer und Schlafqualität der letzten Nacht waren prädiktiv für die Leistungsfähigkeit am Morgen danach. Die Kinder erbrachten bessere Leistungen, wenn sie im Vergleich zu ihrer mittleren Schlafqualität überdurchschnittlich gut geschlafen hatten. Dieser Zusammenhang war besonders ausgeprägt für Kinder mit generell niedrigem Leistungsniveau. Zudem zeigten die Kinder bessere Leistungen, wenn sie zuvor so lange wie üblich geschlafen hatten. Weniger oder deutlich mehr Schlaf als für sie persönlich üblich war nachteilig. Überraschenderweise zeigten sich diese Effekte jedoch nur für die Leistung am Morgen und nicht stabil über den Tag hinweg. Eventuell beeinflusst der Schlaf der letzten Nacht den Start in den Tag, während die Leistungsfähigkeit am Mittag und Nachmittag eher vom Tagesgeschehen abhängt. Nachmittags zeigten die Kinder schlechtere Leistungen an Tagen, an denen sie müder waren als für sie zu dieser Zeit üblich. Vermutlich sind Kinder nachmittags müde, wenn sie einen anstrengenden Tag hatten. Die Zusammenhänge sind besonders aufschlussreich, da sie direkt im Lebenskontext der Kinder gefunden wurden. Die Leistungsfähigkeit der Kinder wurde dort gemessen, wo üblicherweise kognitive Anforderungen an sie gestellt werden: in der Schule und zu Hause. Schlechter Schlaf ging mit weniger kognitiven Ressourcen einher und das bei Kindern, die sich in einer zentralen Lernphase befinden. Als nächsten Schritt werden wir daher Prädiktoren des Schlafverhaltens im Alltag betrachten.

Riediger, M., Wrzus, C., Schmiedek, F., Wagner, G., & Lindenberger, U. (2011). Is seeking bad mood cognitively demanding? Contra-hedonic orientation and working-memory capacity in everyday life. *Emotion*, 11, 656–665.

Sadeh, A., Gruber, R., & Raviv, A. (2003). The effects of sleep restriction and extension on school-age children: what a difference an hour makes. *Child Development*, 74, 444–455.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2013). Keeping it steady: Older adults perform more consistently on cognitive tasks than younger adults. *Psychological Science*. Advance online publication.

Vriend, J. L., Davidson, F., Corkum, P., Rusak, B., McLaughlin, E., & Chambers, C. T. (2012). Sleep quantity and quality in relation to daytime functioning in children. *Children's Health Care*, 41, 204–222.

**Das Arbeitsgedächtnis in der Grundschule: Variabilität verstehen und nutzen**

**GEBF** TAGUNG  
2014

G12 b

Symposiums  
beitrag

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 0.254





## ***Auffälligkeiten im Arbeitsgedächtnis bei Kindern mit Lese-Rechtschreibstörung und Spezifischer Sprachentwicklungsstörung***

Störungen der Sprachentwicklung (SSES) gehen häufig mit späteren Beeinträchtigungen im Lese- und Rechtschreiberwerb einher. Schätzungen zu Folge entwickeln etwa 25-75 % der Kinder mit SSES in der Folge auch Schwierigkeiten im Schriftsprachbereich (Catts et al., 2005; McArthur et al., 2000). Umgekehrt weist jedes fünfte Kind mit einer LRS in der Vorgeschichte auch Beeinträchtigungen im Spracherlernen auf (Catts et al., 2005). Trotz der hohen Komorbiditätsraten liegen bislang nur wenige Forschungsarbeiten zu den Verbindungen und gemeinsamen Verursachungsfaktoren von Sprachstörungen und Lese-Rechtschreibproblemen vor. So ist besonders interessant, ob es sich bei diesen Störungsbildern um zwei voneinander unterscheidbare Störungen mit unterschiedlicher Verursachung oder aber zwei Ausprägungen eines gemeinsam zugrunde liegenden kognitiven Defizits handelt. Als mögliche Ursachenfaktoren werden in diesem Zusammenhang insbesondere Defizite im Arbeitsgedächtnis diskutiert. So konnte in einer Vielzahl von Studien sowohl für Kinder mit LRS aber auch für Kinder mit SSES ein massives Funktionsdefizit in der phonologischen Schleife ausgemacht werden. Darüber hinaus werden aber auch für beide Störungsbilder zentral-exekutive Beeinträchtigungen berichtet.

In der vorliegenden Studie soll die Frage geklärt werden, ob Kinder mit LRS die gleichen Arbeitsgedächtnisdefizite aufweisen im Vergleich zu Kindern mit LRS und kombinierter SSES.

Hierzu werden die Leistungen von 30 Kindern mit LRS (108.57 Monate, IQ = 99.67), 16 Kindern mit LRS und SSES (111.56 Monate, IQ = 93.87) und 30 Kindern einer unauffälligen Kontrollgruppe (108.77 Monate, IQ = 100.53) miteinander verglichen. Die Kinder mit LRS wurden an einer universitären Beratungsstelle für Lernschwierigkeiten rekrutiert. Die Kinder mit LRS und SSES besuchen eine Sprachheilschule. Die Kontrollgruppe sind Kinder der zweiten bis vierten Klassenstufe einer Grundschule. Alle Kinder bearbeiteten in Einzeltestungen eine umfangreiche Testbatterie, die verschiedene Indikatoren zu den drei Subkomponenten des Arbeitsgedächtnisses nach Baddeley (1986) enthält: Phonologische Schleife: Bilderspanne, Ziffernschleife, ein- und dreisilbige Wort- und Nonwordspanne, Kunstwörter-Nachsprechen; Visuell-räumlicher Notizblock: Lokationsspanne, einfache und komplexe Corsi-Block- und Matrixspanne; Zentrale Exekutive: Ziffer- und Wortrückwärtsspanne, Doppelspanne, Zählspanne.

Es zeigt sich, dass sich die Kinder mit LRS im Vergleich zur Kontrollgruppe in der phonologischen Schleife und der zentralen Exekutive unterscheiden, wobei nicht in allen untersuchten Maßen Leistungsunterschiede deutlich wurden. Die Gruppe LRS und SSES zeigt im Vergleich zur Kontrollgruppe ebenso Auffälligkeiten im phonologischen und zentral-exekutiven Arbeitsgedächtnis, allerdings ist hier von einer breiten alle Aufgaben umfassenden Beeinträchtigung auszugehen. Diese fällt insbesondere in der phonologischen Schleife, zum Teil aber auch in der zentralen Exekutive noch massiver aus als bei den Kindern mit isolierter LRS. Das Ergebnismuster erwies sich auch unter Kontrolle des IQ als stabil, so dass die Leistungen nicht aufgrund intellektueller Unterschiede zwischen den Gruppen erklärt werden können. Im Bereich des visuell-räumlichen Notizblocks finden sich wie erwartet keine Unterschiede zwischen den betrachteten Gruppen.

Zusammengefasst lässt sich festhalten, dass durch Gegenüberstellung der Leistungsprofile beider klinischer Gruppen störungsspezifische Besonderheiten aufgedeckt werden konnten. Eine genaue Funktionsanalyse des Arbeitsgedächtnisses stellt somit einen gewinnbringenden Ansatz dar, aufgrund zugrundeliegender kognitiver Defizite zwischen den verschiedenen Störungsbildern zu differenzieren. Außerdem könnte die Abgrenzung der Störungsbilder aufgrund unterschiedlicher zugrundeliegender kognitiver Defizite zu unterschiedlichen Förderempfehlungen führen, die die jeweiligen Stärken oder Schwächen der Kinder differentiell berücksichtigen.

Catts, H., Adolf, S., Hogan, T., & Ellis Weismer, S. (2005). Are specific language impairments and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, 48, 1378-1396.

McArthur, G. M., Hogben, J. H., Edward, V. T., Heath, S. M., & Mengler, E. D. (2000). On the „specifics“ of specific reading disability and specific language impairment. *Journal of Child Psychological Psychiatry*, 41, 869-874.

Katrin Göthe (Universität Potsdam), Anja Gendt (Universität Potsdam), Günther Esser (Universität Potsdam),  
Reinhold Kliegl (Universität Potsdam)

## ***Interferenz als begrenzender Mechanismus des Arbeitsgedächtnisses bei Kindern mit und ohne Lese-Rechtschreibstörung***

In der Forschung der letzten 20 Jahre werden Arbeitsgedächtnisdefizite als mögliche Ursache für verschiedene Lernschwierigkeiten (Gathercole & Pickering, 2001) darunter auch Lese-Rechtschreibstörung (LRS) diskutiert (Schuchard, Mähler, & Hasselhorn, 2008; Siegel, 1994). Ein wichtiges Ziel der Kognitionspsychologie ist es daher herauszufinden, welcher Mechanismus unser Arbeitsgedächtnis begrenzt und ob dieser Mechanismus auch als Erklärung für Defizite bei spezifischen Lern- und Leistungsstörungen herangezogen werden kann. Ein vielversprechender Kandidat die generelle Arbeitsgedächtnisbegrenzung zu erklären, ist die Interferenzhypothese. Sie nimmt an, dass sich aktiv gehaltene Repräsentationen im Arbeitsgedächtnis gegenseitig auf verschiedenen Ebenen stören. Oberauer und Kliegl (Oberauer & Kliegl, 2006) haben in ihrem Interferenzmodell zwei unterschiedliche Mechanismen formalisiert. Zum einen Interferenz durch das partielle Überschreiben einzelner Merkmale von aktiv gehaltenen Repräsentationen, die diese Merkmale miteinander teilen. Zum anderen Interferenz durch Verwechslung ganzer Repräsentationen. Das Modell konnte bereits Daten von jungen und alten Erwachsenen erfolgreich erklären.

In der aktuellen Studie wurde das Interferenzmodell auf die Arbeitsgedächtnisleistung von Kindern mit und ohne LRS angewendet. Die Kinder beider Gruppen wurden hinsichtlich ihres Alters (8 Jahre), der nonverbalen Intelligenz, der Konzentration und der Rechenleistung vergleichbar und im Durchschnittsbereich gehalten. Lediglich die Lese- und Rechtschreibleistungen der Kinder mit LRS lagen 1,5 Standardabweichungen unter dem Erwartungswert ihrer Klassenstufe.

In einer verbalen Gedächtnisaktualisierungsaufgabe haben Kinder beider Gruppen kleine Rechenaufgaben gelöst. Dabei wurde die Anzahl von Früchten einer oder zweier Sorte(n) innerhalb eines Durchganges mehrmals aktualisiert, d.h. erhöht oder verringert. Am Ende eines jeden Durchganges sollte das Endergebnis (eine Fruchtart)/ die Endergebnisse (zwei Fruchtarten) wiedergegeben werden.

Die Genauigkeiten der beiden Gruppen in der Gedächtnisaktualisierungsaufgabe wurden auf die theoretischen Parameter des Interferenzmodells abgebildet und so miteinander verglichen. Ziel war es herauszufinden, ob 1. die Arbeitsgedächtnisleistungen der Kinder mit LRS hinter denen ohne LRS zurück bleiben, 2. das Interferenzmodell Daten von Kindern adäquat erklären kann und 3. die Parameter des Modells die zu erwartenden Arbeitsgedächtnisnachteile der dyslexischen Kinder sinnvoll in den Interferenzparametern abbildet.

Die Modellierungsergebnisse zeigen, dass das Interferenzmodell in der Lage ist, die Genauigkeitsdaten von Kindern abzubilden, wobei Kinder mit einer LRS höhere Interferenz durch Verwechslung ganzer Repräsentationen zeigen als Kinder ohne LRS. Unerwarteterweise zeigen Kinder mit einer LRS jedoch eine geringere Interferenz durch partielles Überschreiben einzelner Merkmale als Kinder ohne diese Störung.

Ein Nachfolgeexperiment mit einem breiteren Altersrange soll testen, ob diese konträren Befunde bezüglich der Interferenzparameter repliziert werden und auf andere Interferenz erzeugende Bedingungen ausgeweitet werden können.

Gathercole, S. E., & Pickering, S. J. (2001). Working memory deficits in children with special educational needs. *British Journal of Special Education*, 28, 89-97.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55(4), 601-626.

Schuchard, K., Mähler, C., & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities*, 41(6), 514-523.

Siegel, L. S. (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development*, 17(1), 109-124.

## **Warten zwischen Hauptschulabschluss und Berufsausbildung im historischen Vergleich. Einfluss von Veränderungen im Bildungssystem, demographischer Entwicklung und dem Konjunkturzyklus auf die Übergangszeiten: Eine Analyse der Schulabgangskohorten 1958-2004**

**Chair(s):** Stefan Zimmermann (Otto-Friedrich-Universität Bamberg)

Besonders Jugendlichen mit Hauptschulabschluss oder keinem Schulabschluss gelingt es oftmals nicht, zeitnahe nach Schulabschluss mit einer Berufsausbildung zu beginnen. In der Folge können sich prekäre Erwerbsbiographien mit einem erhöhten Arbeitslosigkeitsrisiko ergeben (Caspi, Wright, Moffitt & Silva, 1998). Zu Erklärung der historisch veränderten Chancen von schlecht ausgebildeten Schulabgängern stellt sich die Frage, welche Erklärungskraft dem displacement argument zukommt, nachdem über die Zeit gestiegene Erwartungen an die Auszubildenden für längere Übergangszeiten bei der Lehrstellensuche gesorgt haben, und welche Erklärungskraft dem stigmatization by negative selection argument zukommt, nachdem eine Negativ-Selektion dazu führt, dass sich in der Gruppe der verbleibenden Hauptschüler immer mehr Risikofaktoren angehäuft haben (vgl. Solga, 2002).

Bei der Betrachtung des historisch veränderten Übergangsprozesses muss man jedoch auch Veränderungen im Bildungssystem und die Auswirkungen des demographischen Wandels berücksichtigen, der das Verhältnis von Lehrstellenangebot und -nachfrage entscheidend mit beeinflusst. Die in den letzten Jahren gestiegene Bedeutung des Übergangssystems (vgl. Baethge, Solga & Wieck, 2007) ist teilweise hierauf zurückzuführen. Während in den 50er- und 60er-Jahren der Hauptschulbesuch zur Normalbiographie gehörte, ist der Hauptschulabschluss als Folge der Bildungsexpansion in den 70er Jahren immer mehr zum Ausnahmephänomen geworden (Becker, 2006). Bei genauerer Betrachtung zeigt sich, dass der Rückgang der Schülerzahlen an den Hauptschulen in den verschiedenen Bundesländern unterschiedlich schnell vorangeschritten ist. Gleichzeitig wurde wiederholt versucht, die Attraktivität der Hauptschule wiederherzustellen, zunächst mit der Volksschuloberstufenreform Ende der 60er-Jahre, oder später mit der Einführung des freiwilligen 10. Schuljahres, oder der Einführung des Unterrichtsfachs Arbeitslehre, das die Lernenden mit „Grundzügen des Arbeitens und der Produktionsweise“ vertraut machen und somit einen Beitrag zur „Berufswahlreife“ leisten sollte (Dederling, 1996, S. 260). Der Trend der sinkenden Schülerzahlen an den Hauptschulen konnte durch diese Maßnahmen jedoch nicht umgekehrt werden. Ziel der vorliegenden Arbeit ist es zu untersuchen, unter welchen regionalen und historischen Kontextfaktoren der Einstieg in eine vollqualifizierende Berufsausbildung noch gelingt und somit die Hauptschule ihre Sozialisations- und Integrationsfunktion (vgl. hierzu Popp, 1998; Wiater, 2009; Fend, 2006) noch erfüllt. Dazu wird mit der Ereignis-Analyse zunächst untersucht, wie die Entwicklungsaufgabe des Berufseinstiegs von Hauptschülern in unterschiedlichen regionalen und historischen Kontexten bewältigt wurde. Als Datengrundlage wird die Erwachsenenkohorte des Nationalen Bildungspanels (Blossfeld, Roßbach & von Maurice, 2011) verwendet, in der 2887 Erwachsene, die einen Volksschul- oder Hauptschulabschluss in den Jahren 1958-2004 erlangt haben, retrospektiv zu ihrem Lebenslauf befragt wurden. Auf Grundlage dieser Lebensverläufe werden die Wartezeiten zwischen Hauptschulabschluss und Aufnahme einer vollqualifizierenden Berufsausbildung im Rahmen einer Kaplan-Meier Schätzung analysiert. In einem zweiten Schritt wird in einem Piecewise Constant Exponential Model (PCEM) (Blossfeld, Golsch & Rohwer, 2007) für 1290 NEPS-Erstbefragte der Einfluss historischer Kontextbedingungen auf die Hazardrate bzw. Transitionsrate untersucht. Zur Erfassung der historischen Kontextbedingungen wurden auf Grundlage der statistischen Jahrbüchern für jedes Jahr auf Ebene der Bundesländer folgende Indikatoren berechnet: Anteil der Hauptschüler an allen Schülern, Anteil der 15-18-Jährigen, die Arbeitslosenquote, zusätzlich wird eine Zeitreihe zu gesamtwirtschaftlichen Entwicklung (Statistisches Bundesamt, 2013) mit berücksichtigt. Es wird im Rückgriff auf die Typologie von Von Below (2009) zwischen Bildungssystemen unterschieden, die in Bezug auf die Schulstruktur oder in Bezug auf die zu vermittelnden Inhalte strenger oder liberaler sind.

In der empirischen Analyse zeigt sich, dass in den jüngeren Schulabgangskohorten der Übergang in eine Berufsausbildung deutlich langsamer erfolgt. Zur Erklärung der Übergangszeiten tragen signifikant bei: die Altersstruktur der Gesellschaft, die relative Bildungsbeteiligung und die verlängerten Schulzeiten. Gleichzeitig muss festgestellt werden, dass die beobachteten signifikanten Unterschiede zwischen den Schulabgangskohorten sich nicht auf die im Modell berücksichtigten Merkmale zurückführen lassen, da diese bestehen bleiben, wenn Merkmale des Bildungssystems, des Arbeitsmarktes und des demographischen Wandels mit in das Modell aufgenommen werden. Dies kann als Indiz für eine qualitative Veränderung des Übergangsprozesses gedeutet werden, in dem die Hauptschüler zunehmend stigmatisiert werden.

Baethge, M., Solga, H. & Wieck, M. (2007). Berufsbildung im Umbruch: Signale eines überfälligen Aufbruchs, Berlin: Friedrich-Ebert-Stiftung.

Becker, R. (2006). Dauerhafte Bildungsungleichheiten als unerwartete Folge der Bildungsexpansion? In A. Hadjar & R. Becker (Hrsg.) Die Bildungsexpansion: Erwartete und unerwartete Folgen, Wiesbaden: VS Verlag für Sozialwissenschaften, 27-61.

Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (Hrsg.) (2011). Education as a Lifelong Process – The German National Educational Panel Study (NEPS), Zeitschrift für Erziehungswissenschaft: Special Issue 14.

Blossfeld, H.-P. & Golsch, K., Rohwer, G. (2007). Event History Analysis with STATA, London: LEA.

Caspi, A., Wright, B. R. E., Moffitt, T. E. & Silva, P. A. (1998). Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood, American Sociological Review, 63 (3), 424-451.

Dederling, H. (1996). Arbeitslehre in der Sekundarstufe I. In H. Dederling (Hrsg.), Handbuch zur arbeitsorientierten Bildung, München: Oldenbourg.

Fend, H. (2006). Neue Theorie der Schule, Wiesbaden: VS Verlag für Sozialwissenschaften.

Popp, U. (1998). Allgemeinbildende Schulen: Die soziale Funktionen schulischer Bildung. In C. Führ & C.-L. Furck (Hrsg.) Handbuch der deutschen Bildungsgeschichte: 1945 bis zur Gegenwart, München: C. H. Beck, 265-276.

Solga, H. (2002). Stigmatization by Negative Selection: Explaining Less-Educated People's Decreasing Employment Opportunities, European Sociological Review, 18, 159-178.

Statistisches Bundesamt (2013). Volkswirtschaftliche Gesamtrechnungen: Bruttoinlandsprodukt, Bruttonationaleinkommen, Volkseinkommen. Lange Reihen ab 1950.

Von Below, S. (2009). Bildungssysteme im historischen und internationalen Vergleich. In R. Becker, Lehrbuch der Bildungssoziologie, Wiesbaden: VS Verlag für Sozialwissenschaften, 131-153.

Wiater, W. (2009). Theorie der Schule, Donauwörth: Auer.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 1.411



## **Bewerberselektion beim Übergang in eine betriebliche Erstausbildung – Zum Einfluss von Migrationshintergrund und berufsfeldspezifischen Unterschieden**

Theoretischer Hintergrund: Formale Ausbildungsabschlüsse sind in Deutschland von langfristiger Bedeutung für Arbeitsmarktintegration und Erwerbsverlauf (Konietzka, 2010). Die mit dem Ausbildungsabschluss erworbene berufliche Qualifikation hat zumeist eine strukturierende Wirkung für den Übergang in eine erste Erwerbstätigkeit und begründet somit etwa Unterschiede in Einkommen, Berufsprestige sowie Arbeitslosigkeitsrisiko (Weil & Lauterbach, 2011). Daher ist diese Schwelle besonders kritisch für die (Re-)Produktion sozialer Ungleichheiten.

In bisherigen Untersuchungen zu Fremdselektionsprozessen beim Übergang in eine betriebliche Ausbildung ergab sich bisher bereits Evidenz für schlechtere Einmündungschancen von Jugendlichen mit Migrationshintergrund (Diehl et al., 2009; Beicht & Granato, 2011).

Ausgehend von einer ökonomischen Konzeptualisierung des Übergangs in die berufliche Erstausbildung nehmen Betriebe als Gatekeeper (Kohlrausch, 2012) für die Erklärung des Entstehens sozialer Disparitäten eine zentrale Rolle ein. Mittels der Signaling-Theory (Spence, 1973) lässt sich sodann die betriebsseitige Bewerberselektion als Rückgriff auf die Signale und Indizes von Bewerbern verstehen (Solga, 2005, S. 64ff): Beide dienen der (sozialen) Kategorisierung von Bewerbern und werden häufig zu deren Bewertung bezüglich der potenziellen Eignung für (Auszubildenden-)Stellen herangezogen. So kann es bspw. zu einer ethnisch modifizierten Bewerberselektion kommen.

Fragestellung: Aufgrund der Angebot-Nachfrage-Relation sollte sich dabei ein Überangebot (an Bewerbern) tendenziell zu Ungunsten von Bewerbern mit negativ konnotierten Indizes auswirken. Dementgegen kann ein Unterangebot als eine notwendige Bedingung gelten, um den Zugang entsprechender Bewerber zu offenen Stellen zu begünstigen. Die hinreichende Bedingung hierfür ist jedoch, dass sie von Beschäftigern für ‚hinlänglich beschäftigungsfähig‘ gehalten werden (Solga, 2005).

Vor dem Hintergrund eines sich abzeichnenden Mangels an beruflich qualifizierten Fachkräften im MINT-Segment (MINT=Mathematik, Informatik, Naturwissenschaften & Technik) (Anger et al., 2013) ergeben sich die folgenden Fragestellungen:

Sind bei der Vergabe von Ausbildungsplätzen die Bewerbungschancen für Jugendliche mit Migrationshintergrund signifikant schlechter als für Jugendliche ohne Migrationshintergrund? Trifft dies auch auf die MINT-Ausbildungsberufe zu?

Methode:Für die Analyse wird der Datensatz der Teilstudie TOSCA-10 (2007) des Forschungsprojekts „Transformation des Sekundarschulsystems und akademische Karrieren“ (TOSCA) verwendet. Die untersuchte Teilstichprobe dieser Übergangsstudie umfasst n=945 baden-württembergische Realschüler, die sich auf eine duale Ausbildung nach Ende der Sekundarstufe I beworben haben.

Die Anzahl an Bewerbungen, die ein Proband maximal bis zur ersten Einladung zum Bewerbungsgespräch benötigt, dient als Indikator für seine Bewerbungschancen.

Die Angaben zum Geburtsland von Schülern und ihren Eltern werden auf Grundlage einer breiten Definition des Begriffs Migrationshintergrund in einer dichotomisierten Variablen zusammengefasst. Eine Binnendifferenzierung erscheint aufgrund der Zellbesetzung nicht sinnvoll.

Die Unterteilung in MINT- und Büro-Berufe erfolgt auf Grundlage einer vom IAB veröffentlichten Liste von MINT-Ausbildungsberufen sowie der Berufsfelddefinitionen des BIBB (Tiemann et al., 2008).

Der Einfluss von Migrationshintergrund auf Bewerbungserfolg, wird mittels einer Multilevel Mixed-Effects Negative Binomial Regression geprüft, da der Bewerbungserfolg durch eine Zählvariablen quantifiziert wird und mit der Arbeitslosenquote eine Kontrollvariable Schul-Ebene ins Modell einbezogen wird. Als weitere Kontrollvariablen auf Schüler-Ebene werden etwa die Vermittlungshilfe durch Mitglieder der jeweiligen sozialen Netzwerke und Notenprofile (vgl. Nagy, Trautwein & Maaz, 2012) verwendet.

Die Analysen werden getrennt für die Berufsfelder (MINT-, Büro- und sonstige Berufe) gerechnet und anschließend hinsichtlich signifikanter Unterschiede in der Ausprägung der Prädiktoren getestet.

Ergebnisse: Die getrennt für die verschiedenen Berufsfelder gerechneten Analysen haben Unterschiede bezüglich der Prädiktoren des Bewerbungserfolgs ergeben. So lässt sich auch nach umfassender Kontrolle für andere Erklärungsfaktoren in den MINT-Berufen eine signifikante Benachteiligung von Jugendlichen mit Migrationshintergrund feststellen. Ihre Chancen zu einem Bewerbungsgespräch eingeladen zu werden, sind darüber hinaus in den MINT-Berufen signifikant schlechter als in der Gruppe sonstiger Berufe.

Diskussion: Die Ergebnisse deuten auf eine strukturelle Benachteiligung von Jugendlichen mit Migrationshintergrund bei der Auszubildendenselektion hin. Diese zeigt sich bereits im ersten Schritt des Bewerbungsprozesses. Jedoch lässt sich die sozialelektive Wirkung an dieser Schwelle, in letzter Konsequenz, nur unter zusätzlicher Berücksichtigung der folgenden Schritte des Bewerbungsprozesses erfassen. Zudem wäre längsschnittliche Replikation der Ergebnisse auf Bundesebene wünschenswert.

Anger, C., Demary, V, Koppel, O. & Plünnecke, A. (2013). MINT-Frühjahrsreport 2013. Innovationskraft, Aufstiegschancen und demografische Herausforderung. Köln:Institut der deutschen Wirtschaft Köln.

Beicht, U. & Granato, M. (2009). Übergänge in eine berufliche Ausbildung. Geringere Chancen und schwierige Wege für junge Menschen mit Migrationshintergrund. Bonn: Friedrich-Ebert-Stiftung.

Diehl, C., Friedrich, M. & Hall, A. (2009). Jugendliche ausländischer Herkunft beim Übergang in die Berufsausbildung: Vom Wollen, Können und Dürfen. Zeitschrift für Soziologie. 38 (1), 48-67.

Kohlrausch, B. (2012). Betriebe als Gatekeeper. Sozialer Fortschritt, 61 (10), 257-265.

Konietzka, D. (2010): Berufliche Ausbildung und der Übergang in den Arbeitsmarkt. In Becker, R., Lauterbach, W. (Hrsg.), Bildung als Privileg: Erklärungen und Befunde zu den Ursachen von Bildungsungleichheit (S. 277-304). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Nagy, G., Trautwein, U., & Maaz, Kai (2012). Fähigkeits- und Interessenprofile am Ende der Sekundarstufe I: Struktur, Spezifikation und der Zusammenhang mit der Gymnasialzweigwahl. Zeitschrift für Pädagogische Psychologie, 26 (2), 79-99.

Solga, H. (2005). Ohne Abschluss in die Bildungsgesellschaft: Die Erwerbschancen gering qualifizierter Personen aus soziologischer und ökonomischer Perspektive. Opladen: Budrich.

Spence, M. (1973). Job Market Signaling. The Quarterly Journal of Economics, 87 (3), 355-374.

Tiemann, M., Schade, H.-J., Helmrich, R., Hall, A., Braun, U. & Bott, P. (2008). Berufsfeld-Definitionen des BIBB. Bonn: Bundesinstitut für Berufsbildung.

Weil, M. & Lauterbach, W. (2011). In Becker, R. (Hrsg.), Lehrbuch der Bildungssoziologie (S. 329-36). Wiesbaden: VS-Verlag für Sozialwissenschaften.



## Die Bedeutung der beruflichen Interessenkongruenz für den Ausbildungserfolg in mathematisch-naturwissenschaftlich intensiven Ausbildungsgängen

Die Aufnahme einer beruflichen Erstausbildung ist ein einschneidendes Lebensereignis mit wichtigen Folgen für den weiteren Lebenslauf. Die erfolgreiche Absolvierung der Berufsausbildung ist eine wichtige Voraussetzung für die Eingliederung in das spätere Berufsleben (Beicht & Walden, 2013). Der Ausbildungserfolg ist auch aus betrieblicher und volkswirtschaftlicher Perspektive ein wichtiges Zielkriterium. Aus dieser Sicht gelten geringe Abschlussquoten als Indikatoren fehlgeleiteter Personal- und Bildungsinvestitionen (Stamm, 2012).

Im Bereich der mathematisch-naturwissenschaftlich orientierten Ausbildungsgänge stellen sich besondere Herausforderungen. Einerseits gilt es den seit längerem beklagten Fachkräftemangel durch eine Erhöhung der entsprechenden Ausbildungszahlen zu begegnen. Andererseits muss der Ausbildungserfolg der jungen Erwachsenen gefördert werden. So betragen die Abbruchquoten in mathematisch-naturwissenschaftlich orientierten Berufen immerhin 8 bis 10 Prozent, wobei rund 60 Prozent der Auszubildenden die Ausbildung im ersten Jahr abbrechen (Beicht & Walden, 2013).

Gegenstand des vorliegenden Beitrags ist die Untersuchung individueller Determinanten des Ausbildungserfolgs zu Beginn der dualen Ausbildung in ausgewählten mathematisch-naturwissenschaftlich orientierten Ausbildungsberufen. Die Untersuchung personenseitiger Determinanten des Berufs- und Studienerfolgs hat eine lange Tradition (z.B. Apenburg, 1980), während vergleichsweise wenig über die Korrelate des Erfolgs in der dualen Ausbildung bekannt ist. Bisherige Arbeiten haben die Bedeutung von kognitiven und schulisch erworbenen Fähigkeiten (z.B. Lubinski & Dawis, 1992), von Persönlichkeitsmerkmalen (z.B. DeNeve & Cooper, 1998) und von Interessenkongruenz (Passung individueller Interessen zu den umweltseitigen Anforderungen; Holland, 1997) hervorgehoben. Diese Arbeiten zeigen, dass individuelle Fähigkeiten vorwiegend mit dem Leistungsaspekt des Arbeits- und Studienerfolgs sowie im geringeren Maß mit der Arbeits- und Studienstabilität assoziiert sind (z.B. Lubinski & Dawis, 1992). Persönlichkeitsvariablen, wie die „Big Five“, erwiesen sich als prädiktiv für den Leistungs-, den Zufriedenheits- und den Stabilitätsaspekt des Berufs- bzw. Studienerfolgs (z.B. Judge, Heller & Mount, 2002). Interessenkongruenz erwies sich schließlich als besonders prädiktiv für die Berufs-/Studienzufriedenheit und der Berufs-/Studienstabilität (Holland, 1997).

In diesem Beitrag werden die Beziehungen kognitiver und schulischer Leistungen, Persönlichkeitsvariablen (Big Five) und der Interessenkongruenz für den Erfolg in dualen Ausbildungsgängen vorgestellt. Die Datengrundlage bildet die laufende Studie „Mathematisch-naturwissenschaftliche Kompetenzen in der beruflichen Erstausbildung“ (ManKobE). Untersucht wurden N = 1967 Auszubildende im ersten Lehrjahr (Durchschnittsalter 18.20 (SD = 2.38), 23.6% weiblich).

Im Fokus standen zwei Indikatoren des Ausbildungserfolgs, nämlich die Ausbildungszufriedenheit (Westermann, Heise, Spies & Trautwein, 1996;  $\alpha = .87$ ) und die Intention eines vorzeitigen Ausbildungsabbruchs (Ditton, 1998;  $\alpha = .79$ ). Die individuelle Leistungsfähigkeit wurde durch den Kognitiven Fähigkeitstest (Heller & Perleth, 2000; WLE-Reliabilität = .87) und Tests zu den nationalen Bildungsstandards (BISTA) in den Bereichen Mathematik, Physik, Chemie und Biologie erfasst (Pant, Stanat, Schroeders, Roppelt, Siegle & Pöhlmann, 2013; WLE-Reliabilitäten = .75 bis .81). Die Big Five wurden durch das Neo-Fünf-Faktoren Inventar (Borkenau & Ostendorf, 1993) erfasst ( $\alpha = .67$  bis .82). Die Interessenkongruenz wurde mittels des Allgemeinen Interessen und Umwelt Strukturtest (Bergman & Eder, 2005;  $\alpha = .83$  bis .92) erfasst. Konkret wurde die individuelle erreichte Interessenkongruenz durch die intraindividuelle Korrelation der Angaben auf den Interessenitems und der entsprechenden Items zur Erfassung der wahrgenommenen Tätigkeitsanforderungen erfasst (M = .47; SD = .25).

Die Ergebnisse der nachfolgend durchgeführten korrelativen und regressionsanalytischen Auswertungen lassen sich wie folgt zusammenfassen. Alle untersuchten Erfolgsmerkmale waren mit den Kriterien hochsignifikant korreliert. Die Interessenkongruenz erwies sich als leistungsstarker Prädiktor der Ausbildungszufriedenheit ( $\beta = .26$ ,  $p < .001$ ) und der Abbruchsintention ( $\beta = -.63$ ,  $p < .001$ ). Starke Zusammenhänge mit beiden Indikatoren des Ausbildungserfolgs zeigten weiter die Persönlichkeitsvariablen. Von der individuellen Leistungsfähigkeit steuerten lediglich BISTA-Tests einen hochsignifikanten Beitrag zur Prädiktion der Ausbildungszufriedenheit ( $\beta = .11$ ) bei.

Die ermittelten Befunde dokumentieren, dass individuellen Testleistungen, Persönlichkeitsdispositionen und der individuellen Ausprägung der Interessenkongruenz auch im Kontext der beruflichen Ausbildung eine wichtige Rolle zukommt. Im Hinblick auf die untersuchten Erfolgsmerkmale erwiesen sich vor allem die Interessenkongruenz und die Persönlichkeitsvariablen als zentrale Prädiktoren.

Apenburg, E. (1980). Untersuchungen zur Studienzufriedenheit in der heutigen Massenuniversität. Frankfurt a.M.: Lang.

Beicht, U. & Walden, G. (2013). Duale Berufsausbildung ohne Abschluss – Ursachen und weiterer bildungsbiografischer Verlauf. BIBB Report, 21, 1-15.

Bergmann, C. & Eder, F. (2005). Allgemeiner Interessen-Struktur-Test (AIST)/Umwelt-Struktur-Test (UST). Manual (2. Auflage). Göttingen: Beltz.

Borkenau, P. & Ostendorf, F. (1993). Neo-Fünf-Faktoren Inventar nach Costa und McCrae. Handanweisung. Göttingen: Hogrefe.

DeNeve, K.M. & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. Psychological Bulletin, 124, 197-229.

Ditton, H. (1998). Studieninteresse, kognitive Fähigkeiten und Studienerfolg. In J. Abel & C. Tarnai (Hrsg.), Pädagogisch-psychologische Interessenforschung in Studium und Beruf (S. 45-61). Münster: Waxmann.

Heller, K. A. & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz.

Holland, J. L. (1997). Making vocational choices: A theory of vocational personalities and work environments (3rd edition). Odessa, FL: Psychological Assessment Resources.

Judge, T. A., Heller, D. & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. Journal of Applied Psychology, 87, 530-541.

Lubinski, D. & Dawis, R. V. (1992). Aptitudes, skills, and proficiencies. In M. D. Dunette & M. Hough (Eds), Handbook of industrial and organizational psychology. 2nd Edition, Vol. 3 (pp. 1-59). Palo Alto, CA: Consulting Psychologist Press.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Berlin: Waxmann.

Stamm, M. (2012). Zur Rolle des Betriebs beim Ausbildungsabbruch. Zeitschrift für Berufs- und Wirtschaftspädagogik, 108, 18-27.

Westermann, R., Heise, E., Spies, K. & Trautwein, U. (1996). Identifikation und Erfassung von Komponenten der Studienzufriedenheit. Psychologie in Erziehung und Unterricht, 43, 1-22.





Samuel Krattenmacher (Pädagogische Hochschule St.Gallen), Patrizia Salzmann Gischig (Pädagogische Hochschule St.Gallen), Simone Berweger (Pädagogische Hochschule St.Gallen)

## **Prädiktoren intrinsischer Lernmotivation von Berufslernenden im Lehrbetrieb – Ergebnisse einer Schweizer Längsschnittstudie**

In der Schweiz beginnen mehr als zwei von drei Jugendlichen nach der obligatorischen Schulzeit eine Berufslehre. Das Erreichen einer beruflichen Qualifizierung ist eine wichtige Entwicklungsaufgabe im Übergang von der Kindheit zum Erwachsenenalter (z.B. Fend, 2001; Flammer & Alsaker, 2002) und bedeutsam für den Einstieg und die Integration in den Arbeitsmarkt (OECD, 2000; Ryan, 2001; Bynner & Parsons, 2002). Für die Entwicklung berufspraktischer Fähigkeiten spielt die intrinsische Lernmotivation eine wichtige Rolle (Lewalter, Wild & Krapp, 2001). Gemäss der Selbstbestimmungstheorie (Deci & Ryan, 1985; Ryan & Deci, 2000) sind das Kompetenzerleben, das Autonomieerleben sowie die soziale Eingebundenheit wichtige Prädiktoren der intrinsischen Lernmotivation. Diese werden im theoretischen Modell von Krapp und Lewalter (2001) von individuellen Merkmalen der Person sowie von Merkmalen der (betrieblichen) Ausbildung beeinflusst.

In diesem Beitrag gehen wir der Frage nach, welche Ausbildungsmerkmale die intrinsische Motivation von Lernenden, insbesondere ihre Lernmotivation im Lehrbetrieb, beeinflussen. Dabei stützen wir uns auf Daten von 843 Lernenden des Gesundheits- und Sozialwesens sowie der Bauberufe und konzentrieren uns auf Daten aus dem ersten und zweiten Lehrjahr. Ein Grossteil der Lernenden hat die Lehre im Jahr 2012 erfolgreich abgeschlossen.

Anhand eines Strukturgleichungsmodells kann aufgezeigt werden, dass sich die Passung mit der Ausbildung (Beta = .46\*\*\*) direkt, sowie indirekt über das Kompetenzerleben (Beta = .29\*\*\* / Beta = .24\*\*\*) auf die intrinsische Lernmotivation im Betrieb auswirkt. Zudem besitzt ein positives Arbeitsklima, vor allem indirekt über das Kompetenzerleben (Beta = .38\*\*\* / Beta = .24\*\*\*), auf die intrinsische Lernmotivation im Betrieb ebenfalls einen Einfluss. Mit dem getesteten Modell können 46% der Varianz der intrinsischen Lernmotivation im Lehrbetrieb aufgeklärt werden. Das berechnete Modell kann anhand der Gütekriterien (Modell-Fit und Anpassungsgüte) [ $\chi^2 = 84.432$ ;  $p = .000$ ;  $\chi^2/d.f. = 2.222$ ; SRMR = .034; RMSEA = .045; CFI = .981] als gültig angesehen werden.

Bynner, J. & S. Parsons (2002). Social Exclusion and the Transition from School to Work: The Case of Young People Not in Education, Employment, or Training (NEET). *Journal of Vocational Behavior* 60, 289-309.

Deci, E. L. & Ryan, R. M. (1985). The general causality orientations scale: self-determination in personality. *Journal of Research in Personality*, 19, 109-134.

Deci, E. L. & Ryan, R. M. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54-67.

Fend, H. (2001). *Entwicklungspsychologie des Jugendalters*. Ein Lehrbuch für pädagogische und psychologische Berufe. Opladen: Leske + Budrich.

Flammer, A. & Alsaker, F. (2002). *Entwicklungspsychologie der Adoleszenz*. Die Erschliessung innere und äussere Welten im Jugendalter. Bern: Hans Huber.

Krapp, A. & Lewalter, D. (2001). Development of interests and interest-based motivational orientations: A longitudinal study in vocational school and work settings. In S. Volet & S. Järvelä (Eds.), *Motivation in learning contexts: Theoretical advances and methodological implications* (S. 209-232). Amsterdam: Pergamon.

Lewalter, D., Wild, K.-P. & Krapp, A. (2001). Interessenentwicklung in der beruflichen Ausbildung. In K. Beck & V. Krumm (Hrsg.), *Lehren und Lernen in der beruflichen Erstausbildung*. Grundlagen einer modernen kaufmännischen Berufsqualifizierung. Opladen: Leske + Budrich.

OECD (2000). *From Initial Education to Working Life - Making Transitions Work*. OECD. Paris.

Ryan, P. (2001). The School-to-Work Transition: A Cross-National Perspective. *Journal of Economic Literature*, 39(1), 34-92.

## **Entwicklung aufgabenspezifischer Selbstwirksamkeit in Abhängigkeit von erlebtem Erfolg und experimentellem Feedback**

Theoretischer Hintergrund und Fragestellungen: Aufgrund der speziell in Berufsschulen vorherrschenden Heterogenität der Schülerschaft kommt es bei Lernschwächeren im Rahmen selbstgesteuerter Lernprozesse oftmals zu Leistungsproblemen und negativen Beurteilungen. Die damit verbundenen Misserfolgserfahrungen schwächen auf Dauer das Selbstwirksamkeitserleben der Lernenden (z.B. Schwarzer & Jerusalem, 2002; Bandura, 1997) - vor allem dann, wenn die erlebten Misserfolge auf einen Mangel an vorhandenen Fähigkeiten zurückgeführt werden (Ziegler & Schober, 2001). Niedrige Selbstwirksamkeitserwartungen wirken sich wiederum negativ auf Motivation, künftige Leistungen und Erfolgsaussichten der Lernenden aus (Smith, Kass, Rotunda & Schneider, 2006).

Verschiedene Studien zeigten, dass Misserfolge zwar zu einem Absinken der aufgabenspezifischen, nicht aber der (relativ stabilen) allgemeinen Selbstwirksamkeit führen. Aus diesem Grund stand vor allem die aufgabenspezifische Selbstwirksamkeit im Fokus der vorliegenden Untersuchung. Da Effekte auf die nächsthöhere mittlere Ebene - die bereichsspezifische bzw. in diesem Fall die schulische Selbstwirksamkeit - allerdings nicht auszuschließen waren, wurde diese ebenfalls berücksichtigt. Die schulische Selbstwirksamkeit bezeichnet die Überzeugung einer Person, mit den vorhandenen Fähigkeiten eine geforderte Leistung in schulischen Settings erbringen zu können (Jerusalem & Satow, 1999), die aufgabenspezifische Selbstwirksamkeit bezieht sich hingegen auf Leistungen im Rahmen konkreter Aufgaben (Bandura, 2006).

Einen Ansatzpunkt für die Förderung von Selbstwirksamkeitserwartungen bildet Leistungsattribution: Durch die Vermittlung von günstigen Attribuierungen erbrachter Leistungen (z.B. mithilfe von Feedback) können - vor allem im Falle eines niedrigen Ausgangsniveaus - positive Effekte auf die Selbstwirksamkeit erzielt werden (Jain, Bruce, Stellern & Srivastava, 2007). Und je höher die Selbstwirksamkeitserwartungen einer Person sind, desto motivierter wird diese auch sein, bestimmte Handlungen auszuführen (Krapp & Ryan, 2002), sich dabei höhere Ziele zu setzen (Schwarzer & Jerusalem, 2002) und mehr Anstrengung zu investieren (Jain et al., 2007). Dies sind wiederum günstige Voraussetzungen für gute Leistungen und Erfolge, welche langfristig das künftige Selbstwirksamkeitserleben steigern (Bandura, 1997).

Im Rahmen der vorliegenden Studie wurden folgende Fragestellungen untersucht:

- 1) Welche Verläufe der aufgabenspezifischen Selbstwirksamkeit bilden sich im Rahmen latenter Wachstumsmodelle ab?
  - 2) Wie lassen sich Veränderungen der aufgabenspezifischen Selbstwirksamkeit über die Zeit durch den erlebten Erfolg der Lernenden sowie den Erhalt spezieller Feedbacks vorhersagen?
- Methode: Die Stichprobe setzt sich zusammen aus N=171 Schüler/innen (90,6% männlich) einer gewerblichen Berufsschule für Metall- und Elektrotechnik. Das durchschnittliche Alter der Versuchspersonen lag bei 17,19 Jahren. Zur Beantwortung der zentralen Fragestellungen wurde im Rahmen einer pädagogischen Maßnahme an oben genannter Berufsschule eine experimentelle Feldstudie durchgeführt.

Teil der Maßnahme war das Setzen persönlich relevanter Lernziele. Diese individuell festgelegten Ziele sollten von allen Versuchspersonen innerhalb von sechs Wochen schrittweise erreicht werden. Auf Basis wöchentlicher Online-Lernberichte erhielten die Proband/innen während dieses Zielerreichungsprozesses kontinuierliches lernbegleitendes E-Mail-Feedback. Dabei wurden zwei experimentelle Feedbackvarianten implementiert: Die Experimentalgruppe erhielt attributionales E-Mail-Feedback, das speziell auf die Unterstützung von (insbesondere eher niedrigen) Selbstwirksamkeitserwartungen ausgerichtet war. Die Kontrollgruppe erhielt neutrales E-Mail-Feedback, das möglichst frei von jeglichen Attribuierungen formuliert wurde. Die Einteilung der Proband/innen in die Versuchsgruppen erfolgte randomisiert.

Mithilfe der wöchentlichen Online-Lernberichte sowie anhand von Pre-/Post-Tests wurden zahlreiche Eingangs-, Prozess- und Produktvariablen erhoben. Dies schließt unter anderem die Messung der aufgabenspezifischen Selbstwirksamkeit (in Anlehnung an Bandura, 2006), der schulischen Selbstwirksamkeit (Skala "WIRKSCHUL" von Jerusalem & Satow, 1999), der Lernmotivation, der Akzeptanz bezüglich der Lernbegleitung, die Messung von Indikatoren für das Erleben von Erfolg bzw. für Fortschritte im Zielerreichungsprozess der Versuchspersonen sowie die Messung sozio-demographischer Daten ein.

Ergebnisse: Bisherige Analysen (mehrfaktorielle Varianzanalysen mit Messwiederholungsdesign) zeigen über die Dauer der Maßnahme hinweg eine positive Entwicklung der aufgabenspezifischen und der schulischen Selbstwirksamkeit bei Proband/innen mit höherem Erfolg, während die Selbstwirksamkeit der Personen mit niedrigerem Erfolg stagniert.

Im Rahmen des Tagungsbeitrags werden die Verläufe der aufgabenspezifischen Selbstwirksamkeit in Form eines latenten Wachstumsmodells präsentiert. Des Weiteren wird der Einfluss verschiedener Prädiktoren auf das Wachstumsmodell anhand von Strukturgleichungsmodellen überprüft und vorgestellt.

Bandura, A. (1997). Self-efficacy. The exercise of control. New York, NY: Freeman.

Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In F. Pajares & T. C. Urdan (Hrsg.), Self-efficacy beliefs of adolescents (S. 307-337). Greenwich, Conn: IAP - Information Age Pub.

Jain, S., Bruce, M. A., Stellern, J. & Srivastava, N. (2007). Self-efficacy as a function of attributional feedback. In Journal of School Counseling, 5 (4), 1-22.

Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Hrsg.), Skalen zur Erfassung von Lehrer- und Schülermerkmalen (S. 15-16). Berlin: Freie Universität Berlin.

Krapp, A. & Ryan, R. M. (2002). Selbstwirksamkeit und Lernmotivation. In M. Jerusalem & D. Hopf (Hrsg.), Zeitschrift für Pädagogik, 44. Beiheft. Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen (S. 54-82). Weinheim: Beltz.

Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), Zeitschrift für Pädagogik, 44. Beiheft. Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen (S. 28-53). Weinheim: Beltz.

Smith, S. A., Kass, S. J., Rotunda, R. J. & Schneider, S. K. (2006). If at first you don't succeed: effects of failure on general and task-specific self-efficacy and performance. In North American Journal of Psychology, 8 (1), 171-182.

Ziegler, A. & Schober, B. (2001). Theoretische Grundlagen und praktische Anwendung von Reattributionstrainings. Regensburg: Roderer.



## **Forschungskompetenz in bildungswissenschaftlichen Studiengängen – kognitive und motivationale Kompetenzstrukturen**

Chair(s): Jana Groß Ophoff (Pädagogische Hochschule Freiburg), Sandra Schladitz (Pädagogische Hochschule Freiburg), Markus Wirtz (Pädagogische Hochschule Freiburg)

Diskutant/in: Robin Stark (Universität des Saarlandes)

In der Hochschullandschaft werden unter akademischen Kompetenzen schon seit einigen Jahren die Fähigkeiten Studierender thematisiert – sei es normativ in Form allgemeiner Standards für Studienabschlüsse (Kultusministerkonferenz, 2004, 2005) oder aber auch empirisch wie z.B. in dem BMBF-Forschungsprogramm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor“ (Blömeke & Zlatkin-Troitschanskaja, 2013). Den Bildungswissenschaften kommt diesbezüglich eine besondere Rolle zu, da deren Absolvent/-innen im späteren Beruf Verantwortung für den Aufbau grundlegender Kompetenzen künftiger Generationen tragen. In der empirischen Bildungsforschung richtet sich der Blick insbesondere auf die Anforderung, Probleme aus der beruflichen Praxis evidenzorientiert reflektieren und lösen zu können. Dies wird von Shank und Brown (2007) als Educational Research Literacy bezeichnet, also als die Fähigkeit, zielgerichtet Erkenntnisse aus der Bildungsforschung zu erschließen, kritisch zu reflektieren und daraus handlungsleitende Schlussfolgerungen zu ziehen. Die vom Wissenschaftsrat (2006) neben der Berufsqualifizierung geforderte Qualifizierung für die eigene Forschung geht über die hier untersuchte Forschungskompetenz als Rezeption und Anwendung von Forschungserkenntnissen hinaus und umfasst auch die selbstständige Durchführung eigener Forschung. Der gängigen Kompetenzdefinition nach Weinert (2001) lassen sich neben kognitiven auch affektiv-motivationale Aspekte von (Bildungs-)Forschungskompetenz beschreiben. Während Wissensfacetten in objektiven Leistungstests untersucht werden können, lassen sich Überzeugungen und Motivation eher in subjektiven Einschätzungen erfassen. Die Möglichkeiten differenzierter Zugänge zur Konzeptualisierung und Messung der genannten Forschungskompetenzen in bildungswissenschaftlichen Studiengängen werden in diesem Symposium aufgezeigt und abschließend in einem Diskussionsbeitrag von Prof. Dr. Robin Stark kritisch reflektiert werden:

Zwei Beiträge beschäftigen sich mit der kognitiven Facette von Forschungskompetenz: Der Beitrag von Hetmanek, Kiesewetter, Wecker, Trempler, Fischer, Gräsel und Fischer (Projekt KOMPARE) greift die im Rahmen eines bildungswissenschaftlichen Studiums erworbene Kompetenz zur Beurteilung empirischer Evidenz auf. Im Beitrag von Groß Ophoff, Schladitz, Hosenfeld und Wirtz (Projekt LeScEd) wird der Vorhersage der in einem Leistungstest erfassten Forschungskompetenz Studierender durch die subjektive Selbsteinschätzung analysiert und damit der Zusammenhang zwischen objektivem und subjektivem Kompetenzmaß in den Mittelpunkt gestellt.

Zudem untersuchen zwei Beiträge Überzeugungen und affektiv-motivationale Facetten von Forschungskompetenz: In dem Beitrag von Rott, Leuders und Stahl (Projekt LeScEd) werden epistemologische Überzeugungen als handlungsleitende, Aspekte von Forschungskompetenz und ihr Zusammenhang mit kritischem Denken diskutiert. Der Beitrag von Gess und Rueß (bologna lab) befasst sich darüber hinaus mit der Frage, wie das Forschungsinteresse und die forschungsbezogene Selbstwirksamkeit im Studium gefördert werden können.

Ganz im Sinne des Tagungsthemas vereint das Symposium die Perspektiven verschiedener Disziplinen wie der Erziehungswissenschaft, Psychologie, Mathematikdidaktik, Medizindidaktik sowie der Hochschuldidaktik, wodurch eine mehrperspektivische Betrachtung des Konzepts der Forschungskompetenz in den Bildungswissenschaften ermöglicht wird.

Blömeke, S. & Zlatkin-Troitschanskaja, O. (2013). Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor: Ziele, theoretischer Rahmen, Design und Herausforderungen des BMBF-Forschungsprogramms KoKoHs. Humboldt-Universität & Johannes Gutenberg-Universität. Information: [http://www.kompetenzen-im-hochschulsektor.de/Dateien/KoKoHs\\_WP1\\_Bloemeke\\_Zlatkin-Troitschanskaia\\_2013\\_.pdf](http://www.kompetenzen-im-hochschulsektor.de/Dateien/KoKoHs_WP1_Bloemeke_Zlatkin-Troitschanskaia_2013_.pdf) [2013, 30. September].

Kultusministerkonferenz. (2004). Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004. Information: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf), 15.09.2010].

Kultusministerkonferenz. (2005). Qualifikationsrahmen für Deutsche Hochschulabschlüsse. Information: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2005/2005\\_04\\_21-Qualifikationsrahmen-HS-Abschluesse.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2005/2005_04_21-Qualifikationsrahmen-HS-Abschluesse.pdf).

Shank, G. & Brown, L. (2007). Exploring Educational Research Literacy. New York, NY: Routledge.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), Defining and selecting key competences (S. 45–65). Göttingen: Hogrefe.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
IG-Farben-Haus  
Raum 251

Andreas Hetmanek (LMU München), Jan Kiesewetter (LMU München), Christof Wecker (LMU München), Kati Trempler (Universität Wuppertal), Martin R. Fischer (LMU München), Cornelia Gräsel (Universität Wuppertal), Frank Fischer (LMU München)

## **Argumentieren mit Evidenzen im Bildungsbereich. Kompetenzen der Bewertung von Studien**

Theoretischer Hintergrund: Im Bereich der Medizin hat sich die Idee der evidenzbasierten Praxis bereits seit zwei Jahrzehnten etabliert. In jüngster Zeit ist diese Idee auch im Bildungsbereich zum Thema lebhafter Debatten geworden. Bildungswissenschaftliche Evidenz kann für bildungspolitische ebenso wie für didaktische Entscheidungen im Unterrichtsgeschehen herangezogen werden. Voraussetzung für die einträgliche Nutzung von bildungswissenschaftlicher Evidenz als Baustein in komplexen Entscheidungsfindungsprozessen ist die Befähigung zur kompetenten Rezeption und Verwertung von vorhandener Evidenz. Diese Kompetenz umfasst mehrere Teilkompetenzen: i) Suche und Auswahl von relevanter Evidenz, ii) Bewertung von Evidenz mit Bezug auf eine Anwendungssituation, iii) Verwertung der Evidenz in argumentativen Abwägungsprozessen.

Die Beurteilung der Relevanz und Qualität von vorhandener Evidenz im Hinblick auf eine konkrete Entscheidungssituation ist eine wichtige Voraussetzung für fundierte praktische Entscheidungen. Bislang fehlt ein standardisiertes Bewertungsraster zur Beurteilung von vorhandener Evidenz. Zur Entwicklung eines entsprechenden Bewertungsrasters wurden verschiedene Ansätze aus dem psychologischen und medizinischen Bereich kombiniert: So wurden bei der Zusammenstellung und Adaption für den speziellen Bereich der Bewertung von empirischer Evidenz für alltägliche Entscheidungen in der Unterrichtsplanung Aspekte von bereits existierenden Ansätzen zur systematischen Studienbewertung (Cronbach, 1982; Cook & Campbell, 1979 & Harden, Crosby, Davis, & Friedman, 1999) berücksichtigt. Fragestellung: Im vorliegenden Beitrag befassen wir uns mit folgender Frage: Inwiefern lässt sich die Teilkompetenzen der Bewertung von wissenschaftlicher Evidenz in Bezug auf eine Entscheidungssituation unter Verwendung des neu entwickelten Bewertungsrasters valide messen?

Methode: Die Testung der Teilkompetenz der Bewertung von wissenschaftlicher Evidenz in Bezug auf eine Entscheidungssituation (BeStuE) wurde im Rahmen eines ca. zweistündigen Kompetenztests mit 222 Studierenden und Promovierenden aus lehramts- oder bildungswissenschaftlichen Studiengängen an den Universitäten München und Wuppertal erhoben. 82 % der Teilnehmenden waren weiblich, 18 % männlich. Das durchschnittliche Alter betrug 22.4 (SD = 5.2).

Die Testung wurde in einer online-Umgebung szenariobasiert im Labor durchgeführt. In das Szenario waren zwei Entscheidungen integriert; in der ersten sollte zwischen der Arbeit mit gemischten (positive und fehlerhafte) oder rein positiven Beispielen entschieden werden, in der zweiten zwischen Gruppenpuzzle und individueller Textarbeit.

Der Test verlief in drei Phasen: Einführung in das Szenario, Trainingsfall und die beiden Testentscheidungen. Das Bewertungsraster, das acht spezifische und ein zusammenfassendes Item umfasst, wurde von den Probanden auf die Bewertung von acht (je vier pro Entscheidung) ca. einseitigen strukturierten Zusammenfassungen von originalen Forschungsartikeln angewandt. Ein Beispiel-Item (spezifisch) lautet: „Aus dem Text geht hervor, dass die für die Entscheidung relevanten Ergebnisse statistisch belastbar sind.“ Das Zusammenfassungsitem lautet: „Unter Berücksichtigung der zuvor bewerteten Aspekte spielt die Studie für die aktuelle Entscheidung eine wichtige Rolle“. Die Bewertungen wurden auf einer 9-stufigen Likertskala erhoben. Die Reliabilitäten sind durchweg akzeptabel (zwischen .65 und .83).

Das Zusammenfassungsitem wird in der Analyse als abhängige Variable, die spezifischen als Prädiktoren verwendet. Unter Berücksichtigung der hierarchischen Struktur der Daten wurden Mehrebenenanalysen (vgl. u.a. Goldstein, 1995) mit Bewertungen auf Ebene 1 und Versuchspersonen auf Ebene 2 berechnet. Zur besseren Interpretierbarkeit der Regressionskoeffizienten wurden alle Prädiktoren z-standardisiert.

Ergebnisse: Vorläufige Auswertungen der Mehrebenenanalysen zeigen, dass sieben der acht spezifischen Bewertungsdimensionen einen signifikanten Beitrag zur Gesamtbeurteilung einer Studie leisten. Die p-Werte für die sieben signifikanten Prädiktoren sind alle  $p < .01$ . Die Regressionskoeffizienten bewegen sich zwischen .19 und .29. Der Anteil der insgesamt aufgeklärten Varianz ist substantiell (44%). Auf dieser Grundlage gehen wir davon aus, dass das Bewertungsraster zur Beurteilung der Relevanz und Qualität von empirischer Evidenz im Hinblick auf konkrete Entscheidungssituationen ein geeignetes Instrument zu Bestimmung der Nützlichkeit von Studien darstellt. In einem nächsten Schritt wird die Eignung als Testinstrument zur Kompetenztestung zu prüfen sein.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: design & analysis issues for field settings. Boston: Houghton Mifflin Company.

Cronbach, L. J. (1982). Designing evaluations of educational and social programs. San Francisco, CA: Jossey-Bass.

Goldstein, H. (1995). Multilevel Statistical Models. London: Edward Arnold.

Harden, R. M., Crosby, J. R., Davis, M. H., & Friedman, M. (1999). AMEE Guide No. 14: Outcome-based education: Part 5 - From competency to meta-competency: a model for the specification of learning outcomes. Medical Teacher, 21(6), 546-552.

## **Forschungskompetenz in bildungswissenschaftlichen Studiengängen.**

### **Lässt sich die Leistung in objektiven Tests durch die Selbsteinschätzung vorhersagen?**

Theoretischer Hintergrund: An Studiengänge, die für pädagogische Praxis qualifizieren (z.B. Lehramt, Erziehungswissenschaften) wird zunehmend die Anforderung gestellt, Absolvent/-innen dahingehend zu qualifizieren, dass wissenschaftliche Erkenntnisse auf konkrete Probleme angewendet werden können. Der damit verbundene Begriff Evidenzorientierung hebt hervor, dass verfügbare Befunde aus systematischer, valider Forschung mit individuellem Erfahrungswissen verbunden und daraus konkrete Handlungsentscheidungen abgeleitet werden können sollten. Noch findet diesbezüglich in den Bildungswissenschaften eine vornehmlich normativ-konzeptuelle Auseinandersetzung statt. Empirische Erkenntnisse zu bildungswissenschaftlicher Forschungskompetenz lassen sich in angrenzenden Forschungsbereichen finden, deren Schwerpunktsetzung jedoch durch curriculare Inhalte geprägt ist: In der Ausbildung Psychologiestudierender wird beispielsweise Forschungskompetenz mit überwiegend methodisch-kritisch reflektierendem Fokus untersucht (Schweizer, Steinwascher, Moosbrugger, & Reiss, 2011). Dagegen konzentrieren sich Studien zur sog. Information Literacy auf die Fähigkeit, Informationen zu recherchieren, zu evaluieren und zu nutzen (Blixrud, 2003) und zur sog. Statistical Literacy auf die Fähigkeit, Daten zu organisieren, zu konstruieren und darzustellen (Lovett & Shah, 2007). Das hochschulübergreifende Projekt Learning the Science of Education (LeScEd) widmet sich vor diesem Hintergrund der Entwicklung und empirischen Überprüfung eines Kompetenzstrukturmodells, das mit Fokus auf die Bildungswissenschaften die genannten Forschungsbereiche zusammenführt und wesentliche Schritte des Forschungszyklus integriert

Neben standardisierten Leistungstests finden sich im Hochschulsektor auch Verfahren zur Selbsteinschätzung akademischer Kompetenzen – und damit zumindest von einzelnen Aspekten der Forschungskompetenz (z.B. Braun & Hannover, 2008). Dabei ermöglicht speziell multimethodale Diagnostik, also die Kombination von z.B. subjektiven und objektiven Messverfahren, eine Kontrolle des sog. Methodenbias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) und ist grundlegend für die Validierung von Fähigkeitsschätzungen. Die Befundlage und damit der Stand der Diskussion zu objektiven Leistungstests vs. subjektiven Selbsteinschätzungen akademischer Kompetenzen ist jedoch kontrovers, v.a. da nur kleine bis mittlere Zusammenhänge zwischen den Untersuchungsverfahren berichtet werden (z.B. Lowman & Williams, 1987). Insbesondere neigen Studierende häufig zu einer Überschätzung ihrer Kompetenzen (Chevalier, Gibbons, Thorpe, Snell, & Hoskins, 2009).

Fragestellung: Auf Basis von Daten aus dem Projekt LeScEd soll daher in diesem Vortrag der Frage nachgegangen werden, ob die objektiv gemessene Forschungskompetenz bei Studierenden der Bildungswissenschaften durch die subjektive Kompetenzeinschätzung vorhergesagt werden kann.

Methode: In der vorgestellten Large-Scale-Studie wurde an mehreren deutschen Hochschulen die bildungswissenschaftliche Forschungskompetenz von ca. 1300 Studierenden (v.a. Lehramt, Erziehungswissenschaft) erfasst. Der verwendete Itempool wurde basierend auf einschlägiger Literatur und bestehender Instrumente entwickelt und anhand der Erkenntnisse aus Vorstudien optimiert (u.a. Experteninterviews). Um den Aufwand für die Proband/-innen möglichst gering zu halten, trotzdem aber eine hinreichende Menge an Items zu testen, wurde ein Testheftdesign verwendet, wodurch 150 Items verschiedener Komplexität einbezogen werden konnten. Ergänzend wurden alle teilnehmenden Studierenden gebeten, ihre Forschungskompetenz auf einer mehrstufigen Skala (8 Items) selbst einzuschätzen. Die Daten werden mithilfe mehrdimensionaler IRT-Modelle und Strukturgleichungsmodellen ausgewertet.

Ergebnisse: Die Kompetenztestitems werden derzeit ausgewertet und auf Basis von IRT-Modellen analysiert. Erste Befunde sprechen für ein eindimensionales (AIC = 64478.5268; BIC = 66136.9731) statt einem zweidimensionalen Kompetenzstrukturmodell (AIC = 64529.2310; BIC = 66198.1078). Dagegen lassen sich für die Selbsteinschätzung der Forschungskompetenz die Faktoren statistische Kompetenz und kritische Reflexion unterscheiden (Messmodell:  $\chi^2 = 46.720$ ;  $df = 8$ ; CFI = .982; RMSEA = .060). Hierfür erweist sich jedoch nur ersterer als signifikanter Prädiktor mit jedoch geringer Effektstärke ( $\beta = .16$ ;  $p < .05$ , zweiseitige Testung) der Fähigkeitsschätzung durch den standardisierten Kompetenztest. Korrespondierend mit der Literatur weisen vertiefende Analysen darauf hin, dass auch bildungswissenschaftliche Studierende systematisch ihre Forschungskompetenz überschätzen.

Blixrud, J. C. (2003). Project SAILS: Standardized Assessment of Information Literacy Skills. Information: <http://www.arl.org/bm~doc/sails.pdf230/231%5D>.

Braun, E. & Hannover, B. (2008). Kompetenzmessung und Evaluation von Studienerfolg. In N. Jude, J. Hartig & E. Klieme (Hrsg.), *Kompetenzerfassung in pädagogischen Handlungsfeldern* (Vol. 26, S. 153–160). Bonn, Berlin: Bundesministerium für Bildung und Forschung.

Cassidy, S. & Eachus, P. (2000). Learning Style, Academic Belief Systems, Self-report Student Proficiency and Academic Achievement in Higher Education. *Educational Psychology*, 20 (3), 307-322.

Chevalier, A., Gibbons, S., Thorpe, A., Snell, M. & Hoskins, S. (2009). Students' academic self-perception. *Economics of Education Review*, 28 (6), 716-727.

Lovett, M. C. & Shah, P. (Hrsg.). (2007). *Thinking with data*. New York, NY: Lawrence Erlbaum Associates.

Lowman, R. L. & Williams, R. E. (1987). Validity of self-ratings of abilities and competencies. *Journal of Vocational Behavior*, 31 (1), 1-13.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. 88, 879-903.

Schweizer, K., Steinwascher, M., Moosbrugger, H. & Reiss, S. (2011). The structure of research methodology competency in higher education and the role of teaching teams and course temporal distance. *Learning and Instruction*, 21 (1), 68–76.





Benjamin Rott (Pädagogische Hochschule Freiburg), Timo Leuders (Pädagogische Hochschule Freiburg),  
Elmar Stahl (Pädagogische Hochschule Freiburg)

## **Zusammenhang von epistemologischen Überzeugungen und kritischem Denken bei Studierenden des Lehramts Mathematik**

Theoretischer Hintergrund: Studierende des Lehramtes müssen wissenschaftliche Kompetenzen in verschiedenen, heterogenen Bezugswissenschaften erwerben, sowohl in den Bildungswissenschaften als auch in den Disziplinen, welche den gewählten Schulfächern zugeordnet sind. Wissenschaftsbezogene Kompetenzen solcher Art können einerseits nach ihrem Domänenbezug und andererseits nach den Facetten Wissen und Können sowie Überzeugungen unterschieden werden. Innerhalb dieses komplexen Kompetenzbündels, welches im Projekt LeScEd (Learning the Science of Education) in bildungswissenschaftlichen Studiengängen untersucht wird, nehmen wir im hier berichteten Teilprojekt zunächst (1) das kritische Denken und (2) Überzeugungsstrukturen von Lehramtsstudierenden in den Blick. Wir konzentrieren uns hierbei auf die Wissenschaft Mathematik, da Überzeugungen von Lernenden oft wenig reflektiert sind und in starkem Gegensatz zur reichhaltigen Diskussion in der Mathematikphilosophie stehen.

(1) Als eine Facette von Forschungskompetenz bezogen auf die Mathematik, wie sie auch für ein Studium des Lehramtes bedeutsam ist, fokussieren wir auf das kritische Denken in einer geeigneten domänenspezifischen Form. Stanovich und Stanovich (2010, S. 210 ff.) verorten kritisches Denken in einem dreistufigen Modell des Denkens als Prozess der kritischen und reflektierten Überwachung von Problemlöseprozessen. Mit diesem Konstrukt lässt sich beispielsweise erklären, warum Personen mit den gleichen Voraussetzungen auf der Stufe des algorithmischen Denkens insbesondere bei komplexeren Aufgaben unterschiedlich gut abschneiden (vgl. Stanovich & Stanovich, 2010, S. 212 ff.).

(2) Im Bereich der Einstellungen und Überzeugungen fokussieren wir auf epistemologische Überzeugungen, d.h. Überzeugungen zur Natur des Wissens und zum Wissenserwerb, als weitere Facette bildungswissenschaftlicher Forschungskompetenz. Unterschieden werden hier meist mehrere Dimensionen (vgl. Hofer & Pintrich, 1997), wobei Lernende im Verlauf ihrer schulischen und universitären Ausbildung in jeder Dimension Entwicklungen durchlaufen können von eher starren, absolutivistischen Annahmen hin zu flexiblen, vernetzten und evaluativistischen Überzeugungen. Diese Überzeugungen sind für die Lehr-Lernforschung von besonderer Bedeutung, da ihnen handlungsleitende und handlungssteuernde Funktionen zugeschrieben werden und es sich bei ihnen um Prädiktoren für den Erfolg von Lernprozessen handelt (vgl. Hofer & Pintrich, 1997; Schommer, 1998).

Beispielsweise könnte jemand, der Wissen als sicher und in isolierten Stücken strukturiert auffasst, bei mathematischen Problemen nach „der einen“ richtigen Methode suchen; wohingegen jemand, der Wissen als provisorisch und komplex begreift, vermutlich eher nach unterschiedlichen Zugängen und Lösungswegen sucht (Schommer, 1998, S. 133 f.). Fragestellung: Daher ist zu vermuten, dass sophistische epistemologische Überzeugungen und kritisches Denken miteinander einhergehen. Eine Forschungsfrage unseres Projektes, die in diesem Vortrag elaboriert wird, ist es, die Ausprägungen und Zusammenhänge dieser beiden Facetten in Abhängigkeit von Studiengang, Studiendauer und Domänenbezug zu untersuchen.

Methode: In der hier vorgestellten Studie wurden zunächst (1) Kompetenzen des kritischen Denkens bezogen auf die Domäne der Mathematik über eine eigens konstruierte Skala erfasst (unter Verwendung von Aufgaben wie die „bat-and-ball“-Aufgabe nach Kahneman & Frederick, 2002) und in Bezug gesetzt zu (2) epistemologischen Überzeugungen zum Bereich der Mathematik, die wiederum über Interviews und mithilfe eines selbst entwickelten Fragebogens (Rott, Leuders & Stahl, im Druck) erfasst wurden. In der Gruppe von Studierenden des Lehramts Mathematik an der PH Freiburg (N = 128; davon 87 Studienbeginner und 41 Studierende des 3. bis 5. Semesters) wurden die beiden beschriebenen Facetten erfasst.

Ergebnisse: Wir erwarten, dass sich innerhalb der zwei Teilgruppen starke Zusammenhänge zwischen sophistizierten Überzeugungen einerseits und ausgeprägten Fähigkeiten zum kritischen Denken zeigen. Beide Kompetenzfacetten sollten sich ebenfalls über mehrere Semester entwickeln, was durch den Vergleich des Pseudolängsschnittes mithilfe der beiden Teilgruppen belegt werden soll (Mortalität < 10%). In dem Vortrag werden die Ergebnisse in Bezug auf die beiden Konstrukte sowie ihre Diskriminanzvalidität diskutiert und es werden Rückschlüsse auf die Bedeutung der Konstrukte innerhalb der Wissenschaftskompetenz gezogen.

Stanovich, Keith E. & Stanovich, Paula J. (2010): A framework for critical thinking, rational thinking, and intelligence. In D. Preiss & R. J. Sternberg (Hrsg.): Innovations in educational psychology: Perspectives on learning, teaching and human development. New York: Springer. S. 195 – 237.

Kahneman, Daniel & Frederick, Shane (2002): Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Hrsg.): Heuristics and biases: The psychology of intuitive judgment. New York: Cambridge University Press. S. 49 – 81.

Hofer, Barbara K. & Pintrich, Paul R. (1997): The Development of epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning. In: Review of Educational Research 1997, Vol. 67, No. 1. S. 88 – 140.

Rott, Benjamin; Leuders, Timo & Stahl, Elmar (im Druck): Epistemological Judgments in Mathematics: an Interview Study Regarding the Certainty of Mathematical Knowledge. In: Proceedings of the MAVI 2013 Conference.

Schommer, Marlene (1998): The Role of Adults' Beliefs About Knowledge in School, Work, and Everyday Life (Ch. 7). In: Smith, M. Cecil & Pourchot, Thomas (Hrsg.): Adult Learning and Development. London: Lawrence Erlbaum. S. 127 – 143.

GEBF TAGUNG  
2014

G14 c

Symposiums  
beitrag

Mittwoch,

05.03.2014,

15:15 - 17:00,

IG-Farben-Haus

Raum 251



Forschungskompetenz in bildungswissenschaftlichen Studiengängen

## Entwicklung von affektiv-motivationalen Facetten von Forschungskompetenz durch Forschendes Lernen

Theoretischer Hintergrund: Dieser Beitrag geht der hochschuldidaktischen Frage nach, wie sich Forschungskompetenz im Studium effektiv fördern lässt. Auf theoretischer Ebene liegt diesem Beitrag die Überzeugung zugrunde, dass rein kognitive Dispositionen, wie beispielsweise forschungsmethodisches Wissen nur dann Anwendung erfahren, wenn Studierende motiviert sind, dieses Wissen produktiv einzusetzen (Weinert, 2001).

Als eine der zentralen motivationalen Variablen für selbstregulierte Prozesse wird die Selbstwirksamkeitserwartung diskutiert (Schwarzer & Jerusalem, 2002; Liem et al., 2008; Köller & Möller, 2010). Eine einseitige Fixierung auf Selbstwirksamkeit scheint jedoch der komplexen Struktur von Motivation nicht gerecht zu werden; vielmehr sollte insbesondere das individuelle Interesse als weitere motivationale Komponente berücksichtigt werden (Krapp & Ryan, 2002; Marsh et al., 2005). Übertragen auf Forschungskompetenz bedeutet dies, dass das individuelle Interesse an Forschung sowie die forschungsbezogene Selbstwirksamkeit wichtige affektiv-motivationale Facetten bilden.

Die Literatur legt nahe, dass sich das Interesse an Forschung durch die eigene Auseinandersetzung mit Forschungstätigkeiten steigern ließe (vgl. bspw. Person-Objekt-Theorie nach Krapp, 2010). Auch für die forschungsbezogene Selbstwirksamkeit wird deutlich, dass zunächst die (stellvertretende) Durchführung von Forschungstätigkeiten erforderlich ist, um Erfolgserlebnisse, Erfahrungen, verbale Ermutigung oder emotionale Erregung zu ermöglichen (Bandura, 1977).

So könnte es also möglich sein, über Lehr-Lernformen, die eigenständiges studentisches Forschen vorsehen, das Forschungsinteresse der Studierenden und ihre forschungsbezogene Selbstwirksamkeit zu stärken – somit also die affektiv-motivationalen Facetten von Forschungskompetenz zu fördern. Eine dieser Lehr-Lernformen ist das „Forschende Lernen“, wo Studierende eine selbst gewählte Forschungsfrage verfolgen und dabei den kompletten Forschungsprozess durchlaufen (Huber, 2009). Die Literatur deutet darauf hin, dass sich Forschungsinteresse und forschungsbezogene Selbstwirksamkeit durch „Forschendes Lernen“ fördern lassen (Bauer & Bennett, 2003; Seymour et al., 2004; Fichten, 2010; John & Creighton, 2011). Wie genau diese Förderung erzielt wird, ist jedoch bislang empirisch nicht hinreichend untersucht.

Fragestellung: Der vorliegende Beitrag soll ein Versuch sein, diese Lücke zu schließen. Über die vergangenen Semester wurden zwei Pilotprojekte im Format des „Forschenden Lernens“, die an der Humboldt-Universität zu Berlin entwickelt und erprobt wurden, empirisch begleitet. Anhand der Ergebnisse dieser Untersuchungen sollen erste Antworten auf die Frage diskutiert werden, wie „Forschendes Lernen“ gestaltet sein muss, um die affektiv-motivationalen Facetten der Forschungskompetenz zu fördern.

Methode: Die pro Semester 100-150 Studierenden wurden zu Beginn und zum Ende des Semesters befragt, um Änderungen im Zeitverlauf identifizieren zu können. Das Instrument zur Messung des Forschungsinteresses wurde in Anlehnung an den AIST-R (Bergmann & Eder, 2005) entwickelt ( $\alpha=.84$ , 11 Items). Die forschungsbezogene Selbstwirksamkeit wurde anhand des Konzepts der situationsspezifischen Selbstwirksamkeitserwartungen (Schwarzer & Jerusalem, 2002) operationalisiert ( $\alpha=.78$ , 9 Items). Neben den beiden abhängigen Variablen, die an beiden Messzeitpunkten betrachtet werden, werden in der Untersuchung weitere Variablen erhoben, die sich im Kontext des „Forschenden Lernens“ auf theoretischer wie empirischer Basis als relevant erwiesen haben. Dazu zählen: relativ zeitstabile Personenmerkmale wie die Zielorientierung und Studienmotivation (Hänze & Moegling, 2004), Einschätzungen zu den bedürfnisbezogenen Erlebnisqualitäten (Deci & Ryan, 2000), demographische Merkmale, nähere Informationen über den Ablauf der Veranstaltungen und Kontrollvariablen zeitgleich gewählter Lehrveranstaltungen. Da ein Kontrollgruppendesign erst für 2014/15 vorgesehen ist, können bislang lediglich Vergleiche zwischen verschiedenen Arten „Forschenden Lernens“ herangezogen werden, die auf Basis der ex-post erhobenen Informationen über den Ablauf der Veranstaltungen mittels bivariater Analysen sowie einer multivariaten Gesamtbetrachtung vorgenommen werden.

Ergebnisse: Es liegen bereits die Ergebnisse zur Förderung von Forschungsinteresse vor, welche die auf Basis der Person-Objekt Theorie (Krapp, 2010) entwickelten Hypothesen bestätigen. Die Ergebnisse zeigen, dass die Wirksamkeit „Forschenden Lernens“ untrennbar damit zusammenhängt, wie intensiv die Studierenden forschen. Auf Basis der Ergebnisse werden Implikationen für die Praxis diskutiert und die iterative Optimierung der Pilotprojekte vorgestellt. Erste Ergebnisse zur Förderung der forschungsbezogenen Selbstwirksamkeit werden ebenfalls berichtet.

Bandura, a. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191–215.

Bauer, K. W., & Bennett, J. S. (2003). Alumni Perceptions Used to Assess Undergraduate Research Experience. *The Journal of Higher Education*, 74(2), 210–230.

Bergmann, C., & Eder, F. (2005). AIST-R: allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R); Revision; Manual. Beltz Test.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4), 227–268.

Fichten, W. (2010). Forschendes Lernen in der Lehrerbildung. In U. Eberhardt (Hrsg.), *Neue Impulse in der Hochschuldidaktik - Sprach- und Literaturwissenschaften* (S. 127–182). Wiesbaden: VS Verlag für Sozialwissenschaften.

Hänze, M., & Moegling, K. (2004). Forschendes Lernen als selbständigkeitsorientierte Unterrichtsform: Persönliche Voraussetzungen und motivationale Wirkmechanismen. *Psychologie in Erziehung und Unterricht*, 51, 113–125.

Huber, L. (2008). Wie studiert man in "Bologna"? Vorüberlegungen für eine notwendige Untersuchung. In B. M. Kehm (Ed.), *Hochschule im Wandel - Die Universität als Forschungsgegenstand* (S. 295–308). Frankfurt/Main: Campus Verlag

Huber, L. (2009). Warum Forschendes Lernen nötig und möglich ist. In L. Huber, J. Hellmer, & F. Schneider (Hrsg.), *Forschendes Lernen im Studium. Aktuelle Konzepte und Erfahrungen* (S. 9–35). Bielefeld: UVW UniversitätsVerlagWeblar.

John, J., & Creighton, J. (2011). Researcher development: the impact of undergraduate research opportunity programmes on students in the UK. *Studies in Higher Education*, 36(7), 781–797.

Köller, O., & Möller, J. (2010). Selbstwirksamkeit. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 767–774). Weinheim und Basel: Beltz.

Krapp, A. (2010). Interesse. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 311–324). Weinheim und Basel: Beltz.

Krapp, A., & Ryan, R. M. (2002). Selbstwirksamkeit und Lernmotivation. Eine kritische Betrachtung der Theorie von Bandura aus der Sicht der Selbstbestimmungstheorie und der pädagogisch-psychologischen Interessentheorie. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. Zeitschrift für Pädagogik*, 44. Beiheft (S. 54–82). Weinheim: Beltz.

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology*, 33(4), 486–512.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child development*, 76(2), 397–416.

Schwarzer, R., & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Eds.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. Zeitschrift für Pädagogik*, 44. Beiheft (S. 28–53). Weinheim und Basel: Beltz.

Seymour, E., Hunter, A., Laursen, S. L., Deantoni, T., & Al, S. E. T. (2004). Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study. *Science Education*, 88, 493–534.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen: Eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim und Basel: Beltz.

Wissenschaftsrat. (2006). *Empfehlungen zur künftigen Rolle der Universitäten im Wissenschaftssystem*. Berlin.

Mittwoch,  
05.03.2014,

15:15 - 17:00,

IG-Farben-Haus  
Raum 251



## Die Erfassung von Wortschatztiefe bei Vorschulkindern mit Deutsch als Zweitsprache

### Chair(s): Astrid Jurecka (Goethe-Universität Frankfurt)

Hinsichtlich bildungsbezogener Chancengleichheit von Kindern mit und ohne Migrationshintergrund in Deutschland wurde in den letzten Jahren immer häufiger die Sprachkompetenz bezüglich der Unterrichtssprache Deutsch als eine wichtige Voraussetzung für den Bildungserfolg hervorgehoben (Söhn, 2005; Baumert & Schümer, 2001).

Verschiedene empirische Studien weisen jedoch auf deutliche Unterschiede bezüglich des sprachlichen Entwicklungsstands von Kindern mit Deutsch als Zweitsprache (DaZ) und Deutsch als Muttersprache (DaM) zugunsten der einsprachig aufgewachsenen hin (Limbird, 2007; Dubowy et al., 2008; Weinert et al., 2012). Besonders Wortschatzkenntnisse gelten jedoch als wichtige Vorbedingung für den schulischen Erfolg (z.B. Eckhardt, 2008). So konnte verschiedentlich ein Zusammenhang zwischen Wortschatz und Lesekompetenz im Schulalter gezeigt werden (Tannenbaum, 2008; Oulette, 2006). Ferner existieren Hinweise, dass eine Kontextualisierung des Wortschatzes („Theory of learning from context“; Sterberg & Powell, 1983; Sternberg, 1987) beim Erlernen des zweitsprachlichen Wortschatzes von Vorteil sein könnte (bzw. fehlende Kontextualisierung von Nachteil; Eckhardt, 2009), da so die Bedeutung unbekannter Wörter leichter erschlossen werden kann. Wortschatz lässt sich in den quantitativen Aspekt der Wortschatzbreite (Anzahl bekannter Wörter) sowie den qualitativen Aspekt der Wortschatztiefe (Ausmaß semantischer Repräsentation) unterteilen (Oulette, 2006; Tannenbaum, 2008). Beide Aspekte sagen verschiedene Teilleistungsaspekte von Lesekompetenz vorher (Oulette, 2006), weshalb eine differentielle Erfassung beider Aspekte relevant ist. Während sich die Wortschatzbreite als rein quantitativer Aspekt des Wortschatzes jedoch relativ einfach erfassen lässt, ist die Erfassung der Wortschatztiefe gerade bei Kindern im Vorschulalter, die des Lesens meist noch nicht mächtig sind, diffiziler. Der Vortrag behandelt folgende Forschungsfragen:

1.: Lässt sich ein reliables Instrument zur Erfassung der im Rahmen von Kontexten (hier: Geschichten) erworbenen Wortschatztiefe bei Vorschulkindern entwickeln?

2.: Inwieweit unterscheiden sich Vorschulkinder mit und ohne DaZ hinsichtlich der Wortschatztiefe?

Basis des Beitrags stellt eine Interventionsstudie von McElvany und Hardy (2010) dar, in deren Rahmen ein auf Kontextlernen basierendes Förderprogramm zur Förderung des Wortschatzes bei türkischsprachigen Vorschulkindern (n=43) entwickelt wurde. Das Erlernen eines bildungssprachlich ausgerichteten Zielwortschatzes (60 Wörter) wurde mit Hilfe von auditiv vorgegebenen Geschichten (Hörspiele) kontextualisiert. Zusätzlich wurde eine deutsche Vergleichsstichprobe (n=45/keine Intervention) erhoben.

Zu 1: Zur Erfassung der Wortschatztiefe wurden 20 Zielwörter in Sätze eingebunden. Wortschatztiefe (flexibler Einsatz/schnelle Abrufbarkeit in vielfältigen Kontexten; Hardy & McElvany, 2010) wurde anhand einer satzgebundenen Vorgabe der Wörter in unterschiedlichen Kontextbedingungen operationalisiert (auf eine Trainingsgeschichte bezogene/alltagsbezogene Sätze, semantisch richtige/falsche Sätze = 4 Bedingungen/60 Sätze). Die Kinder sollten die vorgelesenen Sätze hinsichtlich ihrer semantischen Richtigkeit einschätzen. Eine einfache Formulierung der Sätze sollte ferner sicherstellen, dass Testwertunterschiede nicht auf konstruktirrelevante schwierigkeitsgenerierende Itemeigenschaften (z.B. Satzlänge) zurückzuführen sind. Im Mittel beantworteten die Kinder 40,6 (s=7,5) der Sätze korrekt (Cronbach's alpha: .65-.76). Ferner wurden die Items eindimensional Rasch-skaliert (ACER ConQuest), wobei lediglich 6 Items keine Modellkonformität aufwiesen (weighted MNSQ <.80 bzw. >1.20; T>2; EAP/PV-Rel.:74). Diese Ergebnisse weisen darauf hin, dass die Items das gleiche Konstrukt erfassen und es insgesamt gelungen scheint, ein reliables Instrument zur Erfassung von Wortschatztiefe bei Vorschulkindern zu entwickeln.

Zu 2: Basierend auf empirischen Studien (s.o.) wurde angenommen, dass Kinder mit DaZ insgesamt eine geringere Wortschatztiefe aufweisen als Kinder mit DaM. Aufgrund der kontextbasierten Intervention bei den DaZ-Kindern wurde jedoch erwartet, dass insbesondere bei Sätzen, die sich auf die Trainingsgeschichten beziehen, kein Gruppenunterschied besteht.

Wie erwartet konnten die Kinder mit DaM insgesamt eine signifikant größere Anzahl von Items korrekt beantworten (t-test; DaM: m=43,7; DaZ: m=36,3; p=0,0). Werden die vier Einzelskalen jedoch getrennt betrachtet zeigt sich bezüglich der geschichtsbezogenen, semantisch richtigen Sätze kein signifikanter Gruppenunterschied (t-test; DaM: m=9,0; DaZ: m=8,4;p=.39). Dies kann als ein erster Hinweis darauf gewertet werden, dass kontextuelles Lernen zu einer Verringerung der Leistungsunterschiede zwischen den Gruppen beitragen kann. Diese sowie ergänzende Ergebnisse (Itemanalysen/varianzanalytische Auswertungen) werden im Beitrag dargestellt und bezüglich ihrer Bedeutung für den Forschungsbereich diskutiert.

Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich (S. 323-410). Opladen: Leske + Budrich.

Dubowy, M., Ebert, S. von Maurice, J. & Weinert, S. (2008). Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten. Ein Vergleich von Kindern mit und ohne Migrationshintergrund. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 40(3), 124-134.

Eckhardt, A. G. (2008). Sprache als Barriere für den schulischen Erfolg. Potentielle Schwierigkeiten beim Erwerb schulbezogener Sprache für Kinder mit Migrationshintergrund. Münster: Waxmann.

Eckhardt, A. (2009). Leistungsunterschiede im Hörverstehen bei Kindern deutscher und nichtdeutscher Herkunftssprache in Abhängigkeit von der Einbettung der Sprache in soziale Handlungen. In K. Schramm (Hrsg.), Empirische Zugänge zu Spracherwerb und Sprachförderung in Deutsch als Zweitsprache (S. 91-110, Mehrsprachigkeit Bd. 23). Munster: Waxmann.

Limbird, C. (2007). Phonological processing, verbal abilities, and second language literacy development among bilingual Turkish children in Germany. Available at: [http://www.diss.fu-37.berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS\\_derivate\\_00000](http://www.diss.fu-37.berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_00000)

McElvany & Hardy (2010). Potenzial der Muttersprache zur Verringerung von Bildungsungleichheit - Wortschatzerwerb von Kindern nichtdeutscher Familiensprache vor zentralen Übergängen des Bildungssystems. Antragstext. Studie gefördert vom BMBF.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology, 98(3), 554-566.

Söhn, J. (2005). Zweisprachiger Schulunterricht für Migrantenkinder. Ergebnisse der Evaluationsforschung zu seinen Auswirkungen auf Zweitspracherwerb und Schulerfolg. Berlin: Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (AKI) Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), The nature of vocabulary instruction (pp. 89-105). Hillsdale, NJ: Erlbaum.

Sternberg, R. & Powell, J. (1983). Comprehending verbal comprehension. American Psychologist, 39, 878-893.

Tannenbaum, K. (2008). Relationship between Measures of Word Knowledge and Reading Comprehension in Third- and Seventh-Grade Children. Electronic Theses, Treatises and Dissertations. Paper 1703.

Weinert, S., Ebert, S., Lockl, K. & Kuger, S. (2012). Disparitäten im Wortschatzerwerb: Zum Einfluss des Arbeitsgedächtnisses und der Anregungsqualität in Kindergarten und Familie auf den Erwerb lexikalischen Wissens. Unterrichtswissenschaft, 40, 4-25.

Mittwoch,  
05.03.2014,  
15:15 - 17:00,  
PEG-Gebäude  
Raum 1.G092



## **Untersuchung der diagnostischen Güte von Instrumenten und Verfahren zur Bestimmung des Sprach- und Leseförderbedarfs von Grundschulkindern**

Theoretischer Hintergrund: Sprach- und Lesekompetenz sind zentrale Voraussetzungen einer erfolgreichen Schullaufbahn (Weinert, 2000). Das schulische Unterrichtsgeschehen ist in hohem Maße sprachlich geprägt, wobei bildungssprachliche Merkmale im Bildungsverlauf zunehmend an Bedeutung gewinnen (Hepppt, Dragon, Berendes, Stanat & Weinert, 2012). Damit die Schülerinnen und Schüler von der schulischen Wissens- und Kompetenzvermittlung profitieren können, müssen sie also über ein hohes Maß an Sprachkompetenz verfügen. Gleichzeitig verstärkt sich im Bildungsverlauf auch die Bedeutung der Lesekompetenz für den schulischen Lernerfolg (Stanat, Weirich & Radmann, 2012).

Da die sprachlichen Kompetenzen der Schülerinnen und Schüler für das Unterrichtsgeschehen und den schulischen Kompetenz- und Wissenserwerb von zentraler Bedeutung sind, bedarf es spezifischer und gezielter Angebote und Maßnahmen zur Förderung von Sprach- und Lesekompetenz. Eine Schlüsselrolle kommt dabei der Kompetenzdiagnostik zu: Damit Fördermaßnahmen die mit ihnen intendierten Erfolge zeitigen und zur Bildungsgerechtigkeit beitragen können, muss zuvorderst sichergestellt sein, dass diejenigen Schülerinnen und Schüler, die einen Förderbedarf aufweisen, auch tatsächlich als solche erkannt werden.

Die wenigen bislang vorliegenden Untersuchungen zur Sprachstandsdiagnostik im Elementar- und Primarbereich konnten zeigen, dass in den Bundesländern eine Vielzahl unterschiedlicher Verfahren zur Erhebung des Sprach- und Leseförderbedarfs genutzt wird (Redder et al., 2011; Fickermann et al., 2011). Diese Verfahren variieren (u. a.) hinsichtlich ihres Standardisierungsgrades – neben Tests (z. B. Hamburger Schreibprobe, MSVK, HASE, BISC) werden beispielsweise auch wenig standardisierte Beobachtungsverfahren (z. B. SISMIK, SELDAK) eingesetzt – und bezüglich ihrer empirischen Fundierung (Schneider et al., 2012). Bemerkenswert erscheint in diesem Zusammenhang, dass die Diagnosegüte vieler Verfahren nicht bzw. kaum evaluiert wurde; auch allgemeineren Fragen – wie zum Zusammenhang von Diagnosegüte und Standardisierungsgrad der eingesetzten Verfahren – wurde bislang kaum nachgegangen.

Fragestellungen: Im Fokus unserer Studie steht zunächst die Frage (1), welche standardisierten und unstandardisierten Verfahren an deutschen Grundschulen genutzt werden, um den sprachlichen Förderbedarf der Schülerinnen und Schüler zu ermitteln. Daran anknüpfend wird der Frage (2) nachgegangen, ob – und falls ja in welchem Maß – die Diagnosegüte mit dem Standardisierungsgrad der eingesetzten Verfahren variiert.

Methode: Unsere Studie basiert auf vertiefenden Analysen der Daten des IQB-Ländervergleichs 2011, an dem insgesamt 27081 Schülerinnen und Schüler der vierten Jahrgangsstufe aus 1349 Schulen teilgenommen haben. Im Rahmen des Ländervergleichs 2011 sind in einem Schulleiterfragebogen unter anderem Informationen über die Art und Weise erhoben worden, mit der an den Schulen sprachlicher Förderbedarf ermittelt wird. Der Fragebogen umfasste beispielsweise Fragen danach, welche Verfahren konkret eingesetzt wurden und welchen Standardisierungsgrad sie aufwiesen. Zur Beantwortung unserer ersten Fragestellung (s. o.) erfolgte eine deskriptive Auswertung dieser Items.

Die Diagnosegüte, die im Fokus unserer zweiten Fragestellung (s. o.) stand, wurde unter Rückgriff auf die Logik der Signalentdeckungstheorie bzw. der ROC-Analyse ermittelt: Im Zuge des Ländervergleichs 2011 war jeweils erhoben worden, inwiefern eine Schülerin bzw. ein Schüler Sprachförderung erhält. In einer 2x2-Matrix wurden diese Informationen den Leistungen der Schülerinnen und Schüler in den Kompetenzbereichen „Zuhören“ und „Lesen“ gegenübergestellt. Dabei galten diejenigen Schülerinnen und Schüler als förderbedürftig, die in mindestens einem der beiden Kompetenzbereiche die Mindeststandards verfehlten. Im Anschluss wurde mithilfe logistischer Mehrebenenanalysen untersucht, inwiefern die Häufigkeit der in der Matrix als „Treffer“, „Falsche Alarme“ usw. klassifizierten Fälle mit dem Standardisierungsgrad der genutzten diagnostischen Verfahren variiert.

Ergebnisse: Die in anderen Bestandsaufnahmen berichtete Heterogenität der eingesetzten Verfahren zur Erhebung des sprachlichen Förderbedarfs zeigte sich auch in den Daten des IQB-Ländervergleichs 2011. Insgesamt berichteten die befragten Schulleiter die Nutzung von über 100 verschiedenen Verfahren, die nicht nur hinsichtlich des Standardisierungsgrades, sondern auch bezüglich der jeweiligen foci differieren (z. B. Fokus auf Beginn der Primarstufe, Fokus auf Schülerinnen und Schüler mit Zuwanderungsgeschichte). Ferner zeigen erste Zwischenergebnisse einen statistisch signifikanten, aber als gering zu klassifizierenden Effekt des Standardisierungsgrades auf die Diagnosegüte.

Fickermann, D., Forschner, S., Kulik, M. & Redder, A. (2011). Auswertung der Länderabfrage zu Sprachstandserhebungen im Kindergarten und beim Übergang in die Schule. Hamburg: ZUSE.

Hart, B. & Risley, T. R. (1995). Meaningful differences in the everyday experiences of young American children. Baltimore, MD: Paul H. Brooks Publishing Co.

Hepppt, B., Dragon, N., Berendes, K., Stanat, P. & Weinert, S. (2012). Beherrschung von Bildungssprache bei Kindern im Grundschulalter. Diskurs Kindheits- und Jugendforschung, 3, 349-356.

McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41(3), 121–131.

Redder, A., Schwippert, K., Hasselhorn, M., Forschner, S., Fickermann, D. & Ehrlich, K. (2011). Bilanz und Konzeptualisierung von strukturierter Forschung zu "Sprachdiagnostik und Sprachförderung". ZUSE-Berichte Band 2.

Hamburg: Hamburger Zentrum zur Unterstützung der wissenschaftlichen Begleitung und Erforschung schulischer Entwicklungsprozesse (ZUSE).

Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T., Roßbach, H.-G., Roth, H.-J., Rothweiler, M. & Stanat, P. (2012). Bildung durch Sprache und Schrift (BISS). Expertise zur Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. Verfügbar unter: [http://www.bmbf.de/pubRD/BISS\\_Expertise.pdf](http://www.bmbf.de/pubRD/BISS_Expertise.pdf) [31.10.2013]

Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011. Münster: Waxmann.

Stanat, P., Weirich, S. & Radmann, S. (2012). Sprach- und Leseförderung. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik (S. 251–276). Münster: Waxmann.





## Entwicklung eines Gruppentests zur Früherkennung von Problemen beim Grammatikerwerb

Das Sprachverständnis von Kindern insgesamt – und ihr Grammatikverständnis im Besonderen – sind höchst relevant für ihren Schulerfolg. Zum einen wird die sprachliche Kompetenz direkt als schulische Leistung gefordert, zum anderen ist sie auch für den Fortschritt in allen anderen Lernbereichen notwendig. Schwierigkeiten im Spracherwerb (besonders im Bereich der Cognitive Academic Language Proficiency) erschweren das gesamte schulische Lernen. Daraus ergibt sich die Notwendigkeit Kinder mit Schwierigkeiten im Grammatikerwerb frühzeitig zu identifizieren, um sie angemessen zu unterstützen und dadurch möglichen Folgeproblemen in anderen Lernbereichen vorzubeugen. Auch im Rahmen von wissenschaftlichen Publikationen wurde mehrfach auf das Fehlen eines geeigneten Gruppentests zur (näherungsweise) Erfassung grammatischer Kompetenzen von GrundschulernInnen hingewiesen. Diese Lücken könnten mit dem hier evaluierten Gruppentest geschlossen werden. Zur Erfassung der rezeptiven grammatischen Kompetenzen liegt mit dem TROG-D (Fox, 2013), der deutschen Version des englischsprachigen Test for Reception of Grammar (TROG von Bishop, 1989), ein in Forschung und Praxis anerkanntes Verfahren vor. Der TROG-D wird allerdings im Einzelsetting durchgeführt und erfordert so einschließlich Auswertung mehr als zwanzig Minuten pro SchülerIn. Als Einzeltest ist er wegen des Aufwands nicht zur Früherkennung oder für den Einsatz in Studien mit größeren Gruppen geeignet.

In einer Forschungsgruppe an der TU Dortmund wurde daher ein Gruppentest konstruiert und pilotiert, der auf dem TROG-D basiert und der im schulischen Kontext mit einer ganzen Klasse durchgeführt werden kann. Dazu wurde das Bildauswahlverfahren des TROG-D in Form eines Testheftes, bei dem die korrekten Bilder angekreuzt werden müssen, realisiert. Die sonst vom Untersucher vorgelesenen auditiven Stimuli werden im Gruppentest von CD abgespielt. Ein solcher Gruppentest hätte für Schulpraxis und Forschung großes Potential, da hiermit ein ökonomisches Verfahren zur Früherkennung von Problemen beim Grammatikerwerb und gleichzeitig ein dringend benötigtes Forschungsinstrument zur Messung grammatischer Kompetenzen in Gruppen zur Verfügung stünde. In der vorliegenden Studie wurde der Gruppentest an einer Zufallsstichprobe von 93 SchülerInnen zweiter und dritter Klassen erstmalig evaluiert. Zunächst wurde in allen Klassen der Gruppentest durchgeführt. Wie schon in den Pilotierungen, erwies sich der Gruppentest als wenig fehleranfällig und für die TeilnehmerInnen sehr motivierendes Verfahren. Im Abstand von zwei Wochen wurden dann alle SchülerInnen mit dem TROG-D im Einzelsetting getestet. Die doppelt minderungskorrigierte Korrelation von  $r = .88$  zwischen den beiden Tests legt nahe, dass der Gruppentest ähnliche Kompetenzen abprüft, wie der Einzeltest. Weiterhin erweist sich auch der Einsatz des Gruppentests zum Screening als sinnvoll: Mit einer Sensitivität von 75,0% und einer Spezifität von 90,1% (Youden-Index = .90; RAZ-Index = .69) gelingt es mit dem Gruppentest sehr gut diejenigen Kinder zu identifizieren, die sich später auch in der Einzeltestung mit dem TROG-D als unterdurchschnittlich erwiesen (Gold-Standard TROG-D T-Wert  $\leq 40$ ). Die Objektivität des Gruppentests ist durch die deutlich stärker standardisierte Durchführung (Auditive Stimuli von CD, vorgegebene Zeitfenster zur Bearbeitung der einzelnen Items) wahrscheinlich höher als beim TROG-D. Die Reliabilität des Gruppentests konnte in dieser Studie nicht kontrolliert werden – lediglich die interne Konsistenz des Verfahrens, die mit der des TROG-D vergleichbar ist, konnte belegt werden. Allerdings liegen auch für den TROG-D selbst keine Schätzungen der Retestreliaibilität vor. Auf Grund der Veränderung des Aufgabenformats (insbesondere der Einführung der Speed-Bedingung) ist nicht auszuschließen, dass der Einfluss von anderen Fähigkeiten wie zum Beispiel der Konzentrationsfähigkeit oder der visuellen Wahrnehmung größer sind als im Einzeltest. Gemessen an den Qualitätsmerkmalen für die Güte von Sprachstandsverfahren von Becker-Mrotzek et al. (2013) könnte der Gruppentest ein für die Vorauswahldiagnostik objektives, reliables und valides Instrument sein, bei dem sich die notwendige Weiterentwicklungs- und Evaluationsarbeit auszahlen könnte. Die hier genannten und weitere Ergebnisse werden diskutiert und Desiderata abgeleitet, die die Entwicklung eines Gruppentests zur Erfassung grammatischer Kompetenzen voranbringen würden.

Becker-Mrotzek, M., Ehlich, K., Füssenich, I., Günther, H., Hasselhorn, M., Hopf, M. et al. (2013). Qualitätsmerkmale für Sprachstandsverfahren im Elementarbereich. Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache (Hrsg.), Köln.

Bishop, D. V. M. (1989). TROG. Test for Reception of Grammar. Manchester: University of Manchester.

Fox, A. V. (2013). TROG-D. Test zur Überprüfung des Grammatikverständnisses (6. Auflage). Idstein: Schulz-Kirchner.



Marion Händel (Friedrich-Alexander-Universität Erlangen-Nürnberg), Kathrin Lockl (Otto-Friedrich-Universität Bamberg),  
Jana Heydrich (Otto-Friedrich-Universität Bamberg), Sabine Weinert (Otto-Friedrich-Universität Bamberg),  
Cordula Artelt (Otto-Friedrich-Universität Bamberg)

## **Wie lässt sich metakognitives Wissen bei Förderschülern valide erfassen?**

Metakognition wird als ein möglicher Bedingungsfaktor für Lernschwierigkeiten bei Schülerinnen und Schülern mit Förderbedarf diskutiert. Lernende mit Förderbedarf haben häufig Probleme ihr eigenes Lernverhalten zu überwachen und regulieren (Schrüder, 2000) und verfügen über ein geringeres metakognitives Lesestrategiewissen als Lernende ohne Lernschwierigkeiten (Pintrich, Anderman & Klobucar, 1994). Schülerinnen und Schüler mit Förderbedarf bilden eine sehr heterogene Gruppe (insgesamt werden neun Arten von Förderbedarf unterschieden), weshalb im Weiteren auf die zahlenmäßig größte Gruppe, Lernende mit Förderbedarf Lernen, fokussiert wird (KMK, 2010).

Die Kompetenztestung von Förderschülern unterscheidet sich von der Kompetenztestung bei Regelschülern z.B. aufgrund von Aufmerksamkeitsspannen, Lesefähigkeiten oder kognitiven Fähigkeiten in vielerlei Hinsicht. Um domänenspezifische oder metakognitive Kompetenzen zu erfassen, können verschiedene Akkomodationen – bezüglich des Tests (weniger oder leichtere Testitems) oder der Durchführung des Tests (Einzel- statt Gruppentestung, längere Bearbeitungszeit) vorgenommen werden (vgl. Bolt & Ysseldyke, 2007; Pitoniak & Royer, 2001 für eine Diskussion über Testakkomodationen bei Lernenden mit Förderbedarf). Bisher jedoch noch wenig betrachtet wurden die Effekte einer Akkomodation auf die Testleistung, und zwar weder innerhalb der Gruppe der Lernenden mit Förderbedarf noch zwischen Lernenden mit und ohne Förderbedarf (Lehmann & Hoffmann, 2009; Wocken, 2000; Wocken & Gröhlich, 2009). Daher wurde in der hier vorgestellten Studie anhand der Messung metakognitiven Wissens der Fragestellung nachgegangen, ob eine ausgewählte Testakkomodation die Testbearbeitung für Schüler und Schülerinnen der Förderschule im Vergleich zur Hauptschule erleichtert und welche Rolle weitere testrelevante Fähigkeiten dabei spielen.

Um systematisch zu untersuchen, welchen Einfluss eine Akkomodation der Testdurchführung auf die Messung metakognitiver Kompetenzen nimmt, wurde anhand von Daten einer Machbarkeitsstudie des Nationalen Bildungspanels (NEPS) der Einfluss einer Testakkomodation bei Schülerinnen und Schülern der Haupt- und der Förderschule untersucht. Weiterhin wurden Lesegeschwindigkeit und schlussfolgerndes Denken als mögliche Moderatoren für die Bearbeitung des metakognitiven Wissenstests angenommen. Es wurden zwei verschiedene Bedingungen realisiert, die sich hinsichtlich der Testdurchführung unterschieden: in einer dem regulären NEPS-Setting äquivalenten Bedingung bearbeiteten die Lernenden den Test in ihrer eigenen Geschwindigkeit, indem sie die Aufgaben selbst durchnahen und beantworteten (Selbstlesebedingung). In einer zweiten Bedingung wurden den Lernenden die schriftlichen Instruktionen und Aufgaben sowohl in Papierform vorgegeben und zusätzlich vorgelesen (Vorlesebedingung). Beide Bedingungen wurden sowohl bei Testungen in der Förderschule als auch in der Hauptschule realisiert. Insgesamt 804 Lernende der sechsten Jahrgangsstufe, ca. 50% davon weiblich nahmen an der Testung teil (397 Förderschülerinnen und -schüler und 407 Lernende aus Hauptschulen).

Folgende Messinstrumente kamen zum Einsatz: das metakognitive Wissen der Teilnehmer und Teilnehmerinnen wurde mittels eines szenariobasierten Tests, der auf unterschiedliche Aspekte des Strategiewissens abzielt, erfasst (Händel, Artelt & Weinert, 2013). Zusätzlich wurden die Lesegeschwindigkeit und schlussfolgerndes Denken erfasst, um mehr Informationen darüber zu erhalten, unter welchen Bedingungen die Lernenden in der Lage sind den metakognitiven Wissenstest in zwei unterschiedlichen Testadministrationen zu bearbeiten. Zur Untersuchung der Fragestellung wurden deskriptive Statistiken, Korrelationsanalysen und Varianzanalysen durchgeführt. Zudem wurden Moderatoranalysen für die Lesegeschwindigkeit und das schlussfolgernde Denken durchgeführt.

Die Ergebnisse der Studie sind im Folgenden zusammengefasst dargestellt: Über beide Schulformen hinweg zeigten sich nur wenige fehlende Werte in der Testbearbeitung und insgesamt waren zufriedenstellende Reliabilitäten zu beobachten. Förderschüler verfügten über deutlich geringeres metakognitives Wissen als Hauptschüler. Es zeigte sich keine bessere Testleistung durch Vorlesen in der Förderschule, Unterschiede ergaben sich jedoch in der Hauptschule zugunsten der Vorlesebedingung. Durch das Vorlesen der Testitems scheint eine validere Erfassung von metakognitivem Wissen möglich gemacht zu werden (dies zeigt sich über beide Schulen hinweg in geringeren Korrelationen mit der Lesegeschwindigkeit in dieser Bedingung). Schließlich ergaben Moderatoranalysen, dass der Einfluss der Testakkomodation wiederum abhängig ist von bestimmten individuellen Voraussetzungen der Lernenden. Weitere Analysen unter Einbezug von Migrationshintergrund werden im Vortrag vorgestellt und die Implikationen für die Testung von Förderschulen abgeleitet.

Bolt, S. E. & Ysseldyke, J. (2007). Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability types. *Journal of Psychoeducational Assessment*, 26, 121–138. doi: 10.1177/0734282907307703

Händel, M., Artelt, C. & Weinert, S. (2013). Assessing metacognitive knowledge: development and evaluation of a test instrument. *Journal of Educational Research Online*, 5, 162–188.

KMK - Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010b). *Sonderpädagogische Förderung in Förderschulen (Sonderschulen) 2009/2010*.

Lehmann, R. & Hoffmann, E. (Eds.) (2009b). *BELLA. Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf „Lernen“*. Münster: Waxmann.

Pintrich, P. R., Anderman, E. M. & Klobucar, C. (1994). Intraindividual differences in motivation and cognition in students with and without learning disabilities. *Journal of Learning Disabilities*, 27, 360–370.

Pitoniak, M. J. & Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, 71, 53–104.

Schröder, U. (2000). Metakognition. In J. Borchert (Ed.), *Handbuch der Sonderpädagogischen Psychologie* (pp. 642–653). Göttingen: Hogrefe.

Wocken, H. (2000). Leistung, Intelligenz und Soziallage von Schülern mit Lernbehinderungen: Vergleichende Untersuchungen an Förderschulen in Hamburg. *Zeitschrift für Heilpädagogik*, 51, 492–503.

Wocken, H. & Gröhlich, C. (2009) Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen. In Bos, W. & Bonsen, M. (Eds.) 2009, *KES 7: Kompetenzen und Einstellungen von Schülerinnen und Schülern der Jahrgangsstufe 7* (pp. 133-142). Münster: Waxmann.

