



## Abstractband

### A01

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus S103

#### **Einflüsse der lexikalischen Textschwierigkeit auf die Lesekompetenz von Schülerinnen und Schülern am Ende der Sekundarstufe I**

**Thomas Canz, Kathrin Jonkmann**

FernUniversität in Hagen, Deutschland; [thomas.canz@fernuni-hagen.de](mailto:thomas.canz@fernuni-hagen.de)

#### Hintergrund

Lesekompetenz ist ein komplexes Konstrukt, welches Fähigkeiten auf verschiedenen Ebenen umfasst. Auf basaler Ebene sind Fähigkeiten der Wort- und Satzerkennung angesiedelt, auf hierarchiehöherer Ebene Fähigkeiten zur globalen Kohärenzbildung, zum Erkennen von Superstrukturen und somit zum inferentiellen Lesen (Lenhard & Artelt, 2009; Rosebrock & Nix, 2008). Beim Lesevorgang sind stets alle Ebenen involviert; während jedoch bei unerfahrenen Lesenden bottom-up-Prozesse dominieren und basale Lesefähigkeiten zentral sind, werden Top-Down-Prozesse und damit die Relevanz hierarchiehöherer Prozesse mit zunehmender Leseexpertise dominanter (Bell & Perfetti, 1994; Droop & Verhoeven, 2003). So zeigt sich auch ein hoher Zusammenhang zwischen Wortschatzfähigkeiten und Lesekompetenz bei Leselernern, dieser Zusammenhang wird jedoch mit zunehmendem Alter und steigender Leseexpertise schwächer (Gillum, 2006).

Aufgrund dieser Befunde wird auch im Rahmen von Assessment-Studien davon ausgegangen, dass mit zunehmender Leseexpertise basale Lesefähigkeiten wie Wortschatzfähigkeiten eine untergeordnete Rolle spielen, weshalb bei der Aufgabenentwicklung für Lesekompetenztests dem lexikalischen Niveau der Texte nur relativ wenig Beachtung geschenkt wird.

#### Fragestellungen

Die vorliegende Studie prüft die theoretische Annahme des untergeordneten Einflusses der Lexik auf die Lesekompetenz bei Lesenden mit fortgeschrittener Leseexpertise kritisch und untersucht Effekte des lexikalischen Textniveaus auf die ermittelte Lesekompetenz von Schülerinnen und Schülern am Ende der Sekundarstufe I. Darüber hinaus wird geprüft, ob der Effekt der lexikalischen Textschwierigkeit durch Faktoren, welche Lesekompetenz- und Leseexpertiseunterschiede indizieren (Geschlecht, Sprachhintergrund, Klassenstufe, Schulform bzw. Schulzweig, Lesefreude, Lesehäufigkeit, Deutschnote), moderiert wird.

#### Methode

In einem experimentellen Design wurden 147 Schülerinnen und Schülern der achten und neunten Jahrgangstufe aus sieben Klassen von insgesamt drei Gemeinschaftsschulen in Hessen vier Testaufgaben zur Ermittlung der Lesekompetenz vorgelegt. Die Aufgaben waren vor der Untersuchung bereits als Lesekompetenzaufgaben erprobt. Für jede Aufgabe wurden zwei Versionen erstellt und der Originaltext hinsichtlich des lexikalischen Niveaus manipuliert, indem ca. jedes fünfte Wort durch ein lexikalisch schwierigeres Synonym oder lexikalisch leichteres Synonym ersetzt wurde. Allen Teilnehmenden erhielten je zwei lexikalisch leichte und zwei lexikalisch schwere Aufgaben. Aufgabenschwierigkeitskombinationen und Aufgabenreihenfolge wurden ausbalanciert, die Zuweisung zu den Testheften erfolgte randomisiert.

Prozentuale Lösungshäufigkeiten für lexikalisch leichte sowie lexikalisch schwere Aufgaben wurden berechnet und anhand eines t-Tests für verbundene Stichproben verglichen. Zur Identifikation der Moderatoren wurde zunächst eine multiple Regression mit den Faktoren Geschlecht, Sprachhintergrund, Klassenstufe, Schulzweig, Lesefreude, Lesehäufigkeit und Deutschnote als Prädiktoren und der Lesekompetenz in Form der im Test erreichten Gesamtpunktzahl als abhängige Variable berechnet. Die potentiellen Moderatoreffekte wurden mittels Varianzanalysen mit Messwiederholung (prozentuale Lösungshäufigkeit für lexikalisch leichte und für lexikalisch schwere Aufgaben als die beiden relevanten Stufen) unter Aufnahme des Moderators in das Modell geprüft.

#### Ergebnisse und Diskussion

Es zeigte sich ein bedeutsamer Haupteffekt für das lexikalische Textniveau ( $M_{\text{leicht}}=37,2\%$ ;  $M_{\text{schwer}}=32,6\%$ ;  $t=2,79$ ;  $p=.006$ ;  $d=0,23$ ). Bedeutsame Haupteffekte für die Gesamtlesekompetenz lagen für Klassenstufe ( $\beta=.24$ ), Schulzweig ( $\beta=.43$ ), Sprachhintergrund ( $\beta=-.16$ ) und Lesefreude ( $\beta=.15$ ) vor. Moderierende Einflüsse auf den lexikalischen Schwierigkeitseffekt zeigten sich lediglich für den Sprachhintergrund ( $F=5,71$ ;  $p=.018$ ;  $\eta^2=.038$ ), jedoch in unerwarteter Richtung: Schülerinnen und Schüler mit nichtdeutscher Erstsprache zeigten geringere Unterschiede zwischen den Lösungshäufigkeiten für lexikalisch

leichte und lexikalisch schwere Texte als Schülerinnen und Schüler mit deutscher Erstsprache. Diese Differenz mag darin begründet sein, dass sich die lexikalische Schwierigkeit von Wörtern dominant am Sprachgebrauch von Erstsprachlern orientiert und in den schwereren Textvarianten mehr Wörter mit fremdsprachlichen Wurzeln enthalten waren, welche fremdsprachlichen Lesenden möglicherweise vertrauter sind.

Die Ergebnisse verdeutlichen einen Einfluss der lexikalischen Textschwierigkeit auf die Lesekompetenz auch bei fortgeschrittener Leseexpertise. Eine Abschwächung dieses Effekts bei höherer Leseexpertise konnte innerhalb des vorliegenden Expertisespektrums nicht festgestellt werden. Dieser Einfluss sollte bei dem Einsatz von Lesematerial sowohl im Assessmentbereich, als auch im Unterricht beachtet werden und ein gesteigertes Augenmerk auf die sprachliche Schwierigkeit von Texten gelegt werden.

## **Evidenzbasierte unterrichtliche Förderung der Schreibkompetenz: Befunde und Erfahrungen aus einem dreistufigen Forschungsprogramm**

**Joachim Grabowski<sup>1</sup>, Michael Becker-Mrotzek<sup>2</sup>**

<sup>1</sup>Leibniz Universität Hannover, Deutschland; <sup>2</sup>Universität zu Köln, Deutschland; [grabowski@psychologie.uni-hannover.de](mailto:grabowski@psychologie.uni-hannover.de)

Aus der interdisziplinären Zusammenarbeit zwischen psychologischer Schreibprozessforschung und Sprachdidaktik resultiert ein Fokus auf die Rolle textsortenübergreifender Fähigkeiten beim Aufbau von Schreibkompetenz (Becker-Mrotzek et al., 2014; Grabowski, 2014). In theoretischer Sicht erweisen sich, neben der Rolle bildungssprachlicher Mittel (Brinkhaus, 2017), vor allem die Fähigkeitsfacetten „Kohärenzherstellung“ und „Perspektivenübernahme“ als relevant. Die Rolle dieser Teilkomponenten der Schreibkompetenz wurde in einem dreistufigen Forschungsprogramm untersucht, über das zusammenfassend berichtet wird.

Phase 1: Diagnose. Bevor mögliche Fördereffekte untersucht werden, bedarf es einer diagnostischen Forschungsphase, in der die zumindest korrelativen Zusammenhänge zwischen den für relevant gehaltenen Teilfähigkeiten und den Zielkriterien (hier: Textqualität von drei verschiedenen Textarten als Indikator von Schreibkompetenz) untersucht und gegen andere erwartbare Einflussgrößen (z.B. intelligenznahe Fähigkeiten, Verarbeitungskapazität und -geschwindigkeit, Variablen der Lesefähigkeit) abgewogen werden. Eine entsprechende Erhebung an N=277 Schülerinnen und Schülern fünfter und neunter Klassen unterschiedlicher Schultypen ergab eine insgesamt hohe Vorhersagbarkeit der Textqualität (Regressionsanalysen; bis zu 57% Varianzaufklärung für Instruktionstexte bei Neuntklässlern) durch die erhobenen Variablen sowie isolierbare signifikante Vorhersageinkremente (jeweils > 10%) durch die Aufgabenbündel zu den postulierten Teilfähigkeiten der Kohärenzherstellung und der Perspektivenübernahme. Damit ist die statistische Relevanz der theoretisch abgeleiteten Fähigkeitsfacetten für die Schreibkompetenz bestätigt, ohne dass aber kausale Wirkungsaussagen getroffen werden können.

Phase 2: Förderung. In einer Prä-Post-Follow up-Interventionsstudie mit Kontrollgruppe (Graham & Harris, 2014) wurden jeweils 24 fünfte und neunte Schulklassen von Gymnasien und Gesamtschulen mit insgesamt ca. 1.200 Schülerinnen und Schülern untersucht. Die Hälfte der Klassen (Kontrollklassen) erhielt regulären Deutschunterricht; in den 24 Interventionsklassen wurde 11 jeweils einstündige, in ein Rahmenthema integrierte Unterrichtseinheiten durch die Deutschlehrkraft durchgeführt mit starker Bindung an Förder- und Aufgabenmaterialien, die von der Forschungsgruppe vorbereitet und didaktisiert wurden, um die Fähigkeitsaspekte der Kohärenzherstellung (5 Einheiten) und der Perspektivenübernahme (5 Einheiten; danach eine wiederholende Abschlusseinheit) zu vermitteln und einzuüben. Der didaktische Aufbau der Einheiten folgte vier Prinzipien: (1) vom Einfachen zum Komplexeren; (2) vom Rezeptiven zum Produktiven; (3) vom Bildhaften zum Textlichen; (4) von geschlossenen zu offenen Aufgabenformaten. Die Einheiten waren weitestgehend als Selbstlerneinheiten für die Schülerinnen und Schüler konzipiert, um Interventionsvarianz durch die jeweiligen Lehrkräfte zu vermeiden.

Die Leistungen der Kontroll- und Interventionsklassen am ersten Messzeitpunkt (Prätest) erwiesen sich als vergleichbar. Für die meisten Variablen ließen sich die erwartbaren Design-Effekte nachweisen (mehrfaktorielle Varianzanalysen): Neuntklässler sind besser als Fünftklässler, Gymnasiasten besser als Gesamtschüler. Ein familiärer sprachlicher Migrationshintergrund wirkt sich bei sprachnahen, aber nicht bei rein kognitiven Aufgaben nachteilig aus. Insofern kann die Datenqualität als hoch eingestuft werden. Eine überlegene Wirksamkeit der Intervention auf die Textqualität ließ sich statistisch (in Form von Wechselwirkungen des Bedingungsfaktors mit dem Messwiederholungsfaktor) allerdings nicht nachweisen. Gleichwohl lassen die regressionsanalytischen Vorhersageergebnisse der Textqualität erkennen, dass es sich bei den berücksichtigten Variablengruppen um Fähigkeitscluster handelt, die mit der Schreibkompetenz in einem empirisch relevanten Zusammenhang stehen. Die Anlage der Intervention erfüllt zehn von zwölf der bei Graham und Harris (2014) als Kennzeichen methodisch guter Interventionsstudien bei der empirischen Schreibforschung genannten Kriterien. Die Empfehlung „Apply a rigorous design“ wurde allerdings zu konsequenter verfolgt, indem eine hoch standardisierte, selbstlernzentrierte Operationalisierung (Sicherung der internen Validität) gegenüber der lehrkraftseitigen Berücksichtigung von Leistungs- und Voraussetzungsunterschieden, insbesondere in den Bereich der Lesefähigkeit und der selbstständigen Arbeit, präferiert wurde.

Phase 3: Praxistransfer. Um diesem Dilemma zwischen hoher Untersuchungskontrolle und unterrichtlichen Praxisanforderungen besser gerecht zu werden, wurde in einer dritten Phase ein Unterrichtskonzept (Lernmaterialien plus Lehrkraftmanual) entwickelt, das die Förderung der Teilfähigkeiten beibehält, aber die instruierende und verständnissichernde Rolle der Lehrkraft stärkt. Dieses Konzept wird an einer kleinen Stichprobe (4 Klassen an Gesamtschulen) erneut evaluiert; die Befunde liegen zum Jahresanfang 2018 vor.

Neben ausgewählten Befunden und Beispielen für das entwickelte Lernmaterial wird vor allem die gewählte interdisziplinär-bildungswissenschaftliche Forschungsstrategie vorgestellt und begründet.

## **Präsentationskompetenz fördern bei Schüler/innen: Effekte eines Präsentationstrainings beim Wettbewerb „Jugend präsentiert“**

**Fabian Ruth<sup>1,2,3</sup>, Evelin Herbein<sup>1,3</sup>, Benjamin Fauth<sup>1,3</sup>, Olaf Kramer<sup>1,2</sup>, Ulrich Trautwein<sup>3</sup>**

<sup>1</sup>Graduiertenschule & Forschungsnetzwerk LEAD, Universität Tübingen; <sup>2</sup>Forschungsstelle Präsentationskompetenz am Seminar für Allgemeine Rhetorik, Universität Tübingen; <sup>3</sup>Hector-Institut für Empirische Bildungsforschung, Universität Tübingen; [fabian.ruth@uni-tuebingen.de](mailto:fabian.ruth@uni-tuebingen.de)

Theoretischer Hintergrund

Die Präsentationskompetenz gehört zu den Kernkompetenzen des 21. Jahrhunderts (van Ginkel, Gulikers, Biemans, & Mulder, 2015). Sie beeinflusst den beruflichen (Campbell, Mothersbaugh, Brammer, & Taylor, 2001) aber auch den schulischen Erfolg.

Die Förderung von Präsentationskompetenz ist bedeutsam, da diese Kompetenz sich nicht automatisch und von selbst entwickelt. Die Präsentationskompetenz ist Bestandteil des Schulcurriculums (Kultusministerkonferenz, 2003) und beeinflusst mündliche Prüfungsleistungen.

Im Vergleich zum Hochschulkontext (Biggs & Tang, 2009; Dunbar et al., 2006; Joughin, 2007) gibt es im Sekundarschulbereich wenig Trainings und die wenigen Interventionen und Förderprogramme sind hinsichtlich ihrer Effektivität kaum überprüft.

Diese Studie untersucht die Effekte eines Präsentationstrainings beim Wettbewerb „Jugend präsentiert“ auf die Präsentationskompetenz von Schüler/innen der Sekundarstufe. Die Präsentationskompetenz wird als mündliches Vermitteln von Wissen verstanden (Alberts, Martin, & Nakayama, 2011), das sich aus einem Zusammenspiel von Präsentationswissen, -fähigkeiten und -motivation zusammensetzt. Es umfasst folgende Präsentationsfacetten: Adressatenorientierung, Körpersprache/Stimme, Visualisierung, Struktur, sprachlicher Ausdruck sowie Darstellung der Expertise.

Fragestellung

Hat das Präsentationstraining im Rahmen des Wettbewerbs „Jugend Präsentiert“ positive Effekte auf die Präsentationskompetenz der Teilnehmenden?

Methode

Die Effektivitätsstudie verwendet ein randomisiertes Warte-Kontroll-Gruppen-Design mit einem Pre- und einem Posttest. Das Präsentationstraining fand in der Präsentationsakademie von Jugend präsentiert statt und dauerte 2,5 Tage. Die Trainingsmethoden basierten auf Kernkomponenten überprüfter und effektiver Trainings des Hochschulkontexts: Übungsgelegenheiten, Rückmeldungen durch Trainer und Mitschüler, Theorieimpulse zu den Präsentationsfacetten, Transferleistungen sowie die Förderung von selbstreguliertem Lernen (z.B. Smith & Sodano, 2011; van Ginkel et al., 2015). Die Trainer waren Rhetorikexperten mit langjähriger Trainingserfahrung.

Die Stichprobe bestand aus 89 Schüler/innen (67 % weiblich, Alter: M = 15,58 Jahre, SD = 1.99, 94,4 % Gymnasiasten), die die Klassen 6 bis 13 besuchten.

Zu jedem Messzeitpunkt hielten die Schüler/innen dreiminütige, mediengestützte Präsentationen, die gefilmt wurden. Die Präsentationsaufgaben waren standardisiert, unterschieden sich jedoch hinsichtlich Themenwahl und Vorbereitungszeit. Während die Teilnehmer/innen bei T1 das Thema frei wählen konnten und unbegrenzte Vorbereitungszeit hatten, wurde dies bei T2 ebenso standardisiert.

Die Präsentationskompetenz wurde durch externe Fremdbewertung (Videoratings) und durch Selbsteinschätzungen erfasst.

Das Tübinger Instrument für Präsentationskompetenz (TIP) wurde bei beiden Messzeitpunkten eingesetzt und bestand aus 23 Items, die die sechs Präsentationsfacetten abdeckten. Auf einer vierstufigen Antwortskala (von „trifft gar nicht zu“ bis „trifft absolut zu“) bewerteten geschulte Rater/innen die gefilmten Schülerpräsentationen mit hoch-inferenten Beurteilungen. Die Interrater-Reliabilitäten erreichten bei 17 von 23 Items Intra-Klassen-Korrelationswerte von  $> .60$ . Die Skalen (Adressatenorientierung, Struktur, Körpersprache/Stimme, Medieneinsatz, Gesamtbewertung) wiesen Reliabilitäten bei T1 und T2 von  $\alpha = .67$  bis  $\alpha = .91$  auf.

Zu T2 wurden auch Selbsteinschätzungen der Schüler/innen hinsichtlich ihrer eigenen Präsentationskompetenz erfasst. Diese bestanden aus 29 Items, welche auf einer vierstufigen Antwortskala beantwortet wurden. Sie umfassten ebenso die sechs Präsentationsfacetten. Die Skalen der Präsentationsfacetten wiesen Reliabilitäten bei T2  $\alpha = .65$  bis  $\alpha = .92$ .

Zusätzlich wurde als Kovariaten mittels Fragebogen die Redeängstlichkeit (Spitznagel, Schlutt, & Schmidt-Atzert, 2003) sowie die Präsentations-Selbstwirksamkeit (basierend auf Stalder, 2013) erfasst. Die Reliabilitäten waren gut bis akzeptabel bei T1 und T2 sowohl bei der Sprechängstlichkeitsskala ( $\alpha = .91/.92$ ) als auch der Selbstwirksamkeit ( $\alpha = .59/.77$ ).

Ergebnisse

Zur Überprüfung der Effekte des Trainings auf die Präsentationsfacetten wurden multiple Regressionen verwendet. Bei der erfassten Präsentationskompetenz durch Selbstberichte und durch Videoratings wurden als Kovariaten die Sprechängstlichkeit, die Selbstwirksamkeit sowie das Alter und das Geschlecht aufgenommen. Die Analysen der selbsteingeschätzten Präsentationskompetenz ergaben keine signifikanten Trainingseffekte. Bei der Einschätzung der Präsentationskompetenz durch das Videorating wurden zusätzlich noch die Pretests als Kovariate aufgenommen. Dabei konnte ein positiver Interventionseffekt auf die Adressatenorientierung ermittelt werden ( $B = 0.41$ ,  $p = .038$ ).

Die Ergebnisse geben Hinweise darauf, dass das Intensivtraining unmittelbare positive Effekte auf eine der Facetten der Präsentationskompetenz hat. Zukünftige Studien sollten zusätzlich die Trainingseffekte in einem Follow-up erfassen, um nicht nur die unmittelbaren Effekte zu betrachten.

## **Setzung und Erreichung von Standards für Schreibkompetenzen im Fach Englisch der Sekundarstufe II**

**Johanna Fleckenstein<sup>1</sup>, Stefan Keller<sup>1</sup>, Olaf Köller<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule FHNW, Schweiz; <sup>2</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; [johanna.fleckenstein@fhnw.ch](mailto:johanna.fleckenstein@fhnw.ch)

Schreiben in der Fremdsprache Englisch ist eine Schlüsselkompetenz für Studierfähigkeit in einer zunehmend globalisierten Welt. Die Studie Measuring English Writing at Secondary Level (MEWS) untersucht argumentatives Schreiben in Englisch bei Schülerinnen und Schülern der Sekundarstufe II in Deutschland und der Schweiz. MEWS soll dazu beitragen, ein differenziertes Verständnis für die Messung und Entwicklung bzw. Förderung englischer Schreibfähigkeiten zu generieren. Um die Kompetenzen der Schülerinnen und Schüler einordnen zu können, wurden im Rahmen eines Standard Settings Kompetenzstufen basierend auf dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER; Europarat, 2001) definiert. Dieser Beitrag setzt sich sowohl mit dem Vorgehen bei der Kompetenzstufensetzung auseinander als auch mit Erreichung von Bildungsstandards beim Abitur bzw. bei der Matura.

Die Fähigkeiten im argumentativen Schreiben in Englisch von N = 2554 Schülerinnen und Schülern wurden anhand eines computer-basierten Tests zu zwei Zeitpunkten innerhalb des 11. Schuljahres erfasst. Die Schülerinnen und Schüler bekamen dabei zu jedem Messzeitpunkt zwei Schreibaufgaben (eine freie und eine materialgestützte Argumentation) basierend auf dem Test of English as a Foreign Language; TOEFL) vorgelegt, um die Produktion von Schreibtexten mit unterschiedlichen Anforderungen zu erzielen.

Zur empirischen Auswertung der Schülertexte wurde – als Ergänzung zu Beurteilungen durch menschliche Rater – eine spezielle Software (E-rater ©; Chodorow & Burstein (2004) eingesetzt, welche Schlüsselkomponenten von Schreibkompetenz

automatisiert beurteilen kann (z.B. lexikalische Komplexität, Syntax oder Textstrukturierung). Dadurch sollen einerseits die Erfassung einer grossen Stichprobe ermöglicht, andererseits Probleme menschlicher Textbeurteilungen (hohe Kosten, mangelnde Beurteilungskonsistenz, Müdigkeits- oder Reihenfolgeeffekte) abgemindert werden.

Die Teilnehmenden des Standard Settings waren 12 Expertinnen und Experten aus Deutschland und der Schweiz, insbesondere Lehrkräfte und in der Lehrkräfteausbildung tätige Personen. Diese Gruppe wurde vor dem Standard Setting Workshop anhand ausgewählter Skalen auf die Beurteilung von Schreibtexten auf Basis des GER vorbereitet. Darüber hinaus fand eine Familiarisierung mit den entsprechenden Schreibaufgaben statt. Während des Standard Setting Workshops wurden die Teilnehmenden dann gebeten anhand der authentischen Schülertexte globale Cut-Scores für die Untergrenzen der GER-Niveaus B1, B2, C1 (Runde 1) und C2 (Runde 2) festzulegen. In einem iterativen Verfahren einigten sich die Teilnehmenden auf die Festlegung der entsprechenden Kompetenzstufen.

Für die in MEWS getesteten Schülerinnen und Schüler ergab sich zu Beginn der 11. Klasse basierend auf den Resultaten des Standard Settings folgende Verteilung auf die GER-Niveaus: 5.0% auf A2, 19.4% auf B1, 65.9% auf B2 und 9.7% auf C1. Ausgehend von dieser Verteilung scheint die Erreichung beispielsweise der deutschen Bildungsstandards für die fortgeführte Fremdsprache Englisch (B2, in Teilen auch C1; KMK, 2014) zum Zeitpunkt des Abiturs realistisch. Die Schülerinnen und Schüler in der Schweiz schneiden dabei etwas besser ab ( $d = 0.3$ ), was unter anderem mit der im Vergleich zum deutschen Schulsystem stärkeren Selektivität der Sekundarstufe II in Zusammenhang gebracht werden kann.

Der Beitrag gibt Einblicke in die Messung und Beurteilung fremdsprachiger Schreibfähigkeiten sowie die Entwicklung von Kompetenzstufenmodellen. Die Befunde der Studie sind relevant für die Bildungssteuerung, da ein differenzierteres Bild der Leistungsfähigkeit junger Menschen in verschiedenen Bildungssystemen entsteht. Zudem haben die Erkenntnisse Implikationen für die Lehrerbildung und die Unterrichtspraxis, da untersucht wird, wo Defizite beim Schreiben in Englisch bestehen und welche Aspekte des Englischunterrichts fremdsprachliche Schreibfähigkeiten positiv beeinflussen.

## A02

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus S105

### **Zum Studienerfolgsverständnis von Professor\*innen des Studienbereichs Wirtschaft in dualen Studiengängen**

**Steffen Wild, Ernst Deuer**

DHBW Ravensburg, Deutschland; [wild@dhw-ravensburg.de](mailto:wild@dhw-ravensburg.de)

Zahlreiche Hochschulen buhlen um Studierende und müssen ‚Erfolgsquoten‘ u.a. gegenüber staatlichen Organen oder Hochschulrankings nachweisen, um sich zu legitimieren (Klein & Stocké 2016). Allerdings ist der Terminus des Studienerfolgs (Berthold et al. 2015) nicht eindeutig definiert (Ramseier 1980). Vielmehr kann der Begriff aus verschiedenen Perspektiven gesehen werden, wie etwa anhand der Fachkultur (Multrus 2004) oder anhand zahlreicher Theoriemodelle (Rindermann & Oubaid 1999; Kuh et al. 2007). Des Weiteren bestehen in der vielfältigen Hochschullandschaft in Deutschland unterschiedliche Schwerpunkte der Hochschulen. Ein Differenzierungsmerkmal ist die eher primäre Ausrichtung auf theoretische Inhalte bei Universitäten, die anwendungsorientierten Elemente bei Hochschulen für Angewandte Wissenschaften oder die Theorie-Praxis-Verknüpfung an dualen Hochschulen. Es stellt sich somit die Frage, welche Studienerfolgsverständnis herrscht bei unterschiedlichen Hochschulformen und Fachkulturen vor? Heublein et al. (2015) untersuchten beispielsweise Fakultäts- und Fachbereichsleitungen an Universitäten und Fachhochschulen ohne einen Fokus auf das duale Studium zu richten. Die vorliegende Studie untersucht in einem ersten Schritt das Studienerfolgsverständnis von Professor\*innen der Duale Hochschule Baden-Württemberg (DHBW) und fokussiert sich auf den Studienbereich Wirtschaft.

Anhand einer Teilstudie des Forschungsprojekts „Studienverlaufspanel – Weichenstellungen, Erfolgskriterien und Hürden im Verlauf des dualen Studiums an der DHBW“ (Deuer et al. 2017) konnte die Professor\*innenschaft der DHBW zum Verständnis von Studienerfolg in einer Vollerhebung befragt werden. Insgesamt wurden an den 9 Standorten der DHBW 359 Professor\*innen befragt. Auf der Basis von 156 Befragungsteilnehmer\*innen des Studienbereichs Wirtschaft konnte anhand von 6 Indikatoren zum Studienerfolg eine Datenanalyse vorgenommen werden. Aufgrund der verschiedenen Auffassungen des Begriffs Studienerfolgs waren die deskriptiven Analysen nur bedingt zielführend, da Antwortmuster, etwa aufgrund der hohen Standardabweichungen in den Items, hinter den Antworten vermutet wurden. Daher war es anschließend unausweichlich eine Latente Profilanalyse (LPA) für die Klassifizierung des Verständnisses von Studienerfolg vorzunehmen.

Es konnten 3 Typen von Studienfolg bei der Untersuchungspopulation analysiert werden: „Fachfokussierung“ (28%), „Studierendenfokussierung“ (58%) und „Arbeitsmarktfokussierung“ (14%). Lehrende des Typs Fachfokussierung hatten außer bei dem Studienerfolgskriterium „Absolventen mit sehr guten fachlichen Fähigkeiten“ immer die niedrigste Zustimmung bei den abgefragten Kriterien. Lehrende des Typs Studierendenfokussierung sind „niedrige Studienabbruchquoten“ und „gute Prüfungsnoten“ wichtig. Dagegen sind Lehrende des Typs Arbeitsmarktfokussierung ein „Auslandsaufenthalt“ und ins besonders eine „gute Vorbereitung auf den Arbeitsmarkt“ von zentraler Wichtigkeit. Das Ergebnis wurde anhand einer Diskriminanzanalyse evaluiert und zusätzlich anhand einer Kreuzvalidierung (Leave-one-out Methode) überprüft. Hier konnten insgesamt 70.8% der ursprünglichen und 66.7% der Kreuzvalidierten Fälle korrekt klassifiziert werden (Leave-one-out-Kreuzvalidierung).

Die drei Profile können als eine erste Charakterisierung für das Verständnis von Studienerfolg bei lehrenden Wirtschaftswissenschaftler\*innen in dualen Studiengängen aufgefasst werden und als Ausgangspunkt für Vergleiche verschiedener Stakeholderperspektiven sowie für hochschulinterne wie hochschulübergreifende Vergleichsanalysen zwischen Fachgruppen bzw. -kulturen dienen. Des Weiteren können derartiger Befunde Hinweise geben inwiefern die (bestehende oder sich wandelnde) Ressourcenverteilung an der Hochschule (für Studienberatung, Bibliothek, Auslandsamt) oder der Einsatz neuer Lehr- und Lernformen (E-Learning oder Lern-Apps) seitens der hauptamtlichen Kolleg\*innen in der Hochschullehre auf Akzeptanz und Unterstützung treffen können.

### **Geschlecht und wissenschaftliche Produktivität – Erklären Elternschaft und akademisches Sozialkapital Produktivitätsunterschiede in der Promotionsphase?**

**Steffen Jaksztat**

DZHW, Deutschland; [jaksztat@dzhw.eu](mailto:jaksztat@dzhw.eu)

Dieser Beitrag betrachtet den Zusammenhang zwischen dem Geschlecht und der wissenschaftlichen Produktivität während der Promotionsphase. Gerade für Forscherinnen und Forscher in frühen Karrierephasen, die in zunehmendem Maße um Stellen, Stipendien und Fördermittel konkurrieren (Rogge 2015), stellen wissenschaftliche Publikationen wichtige Produktivitäts- und Leistungssignale dar. Empirische Studien können entsprechend zeigen, dass die Berufungschancen in hohem Maße von der Zahl wissenschaftlicher Publikationen abhängen (Lang & Neyer 2004; Jungbauer-Gans & Gross 2013; Lutter & Schröder 2016). Insbesondere in der deutschsprachigen Wissenschaftssoziologie und Hochschulforschung fand das Thema Geschlecht und wissenschaftliche Produktivität bisher jedoch nur vergleichsweise wenig Beachtung.

Aufbauend auf theoretischen Überlegungen zu Geschlechterrollen und zur Entstehung von Unterstützungsnetzwerken wird untersucht, in welchem Umfang Produktivitätsunterschiede bereits während der Promotionsphase entstehen und inwiefern diese als Folge unterschiedlicher Effekte von Elternschaft oder als Folge einer geschlechterdifferenzen Ausstattung mit wissenschaftlichem Sozialkapital betrachtet werden können. Diese Forschungsfragen werden mit Daten des bundesweiten DZHW-Promoviertenpanels bearbeitet.

Mittels negativer Binomial-Regressionen wird der Effekt des Geschlechts auf die Anzahl veröffentlichter wissenschaftlicher Publikationen ermittelt (peer-review Artikel sowie alle Publikationen). In einem schrittweisen Vorgehen werden die oben beschriebenen Mediatormechanismen kontrolliert, um zu untersuchen, inwiefern diese mögliche Erklärungen für Geschlechterunterschiede liefern können.

Die Analysen zeigen, dass es Frauen während der Promotionsphase nicht in gleichem Umfang wie Männern gelingt, wissenschaftliche Publikationen zu veröffentlichen. Die in diesem Sinne geringere wissenschaftliche Produktivität von Wissenschaftlerinnen kann teilweise dadurch erklärt werden, dass Frauen von ihrem wissenschaftlichen Umfeld weniger gefördert und unterstützt werden. Entgegen den theoretischen Erwartungen können keine Effekte von Elternschaft auf die Publikationszahlen nachgewiesen werden. Produktivitätsunterschiede dürften einen erheblichen negativen Einfluss auf die Karrierechancen von Frauen innerhalb der Wissenschaft haben.

## **Forschungsleistung: Fachübergreifende Erfassung und Zusammenhänge mit Einstellungen, Verhalten und Erleben von Wissenschaftler(inne)n**

**Martin Daumiller, Markus Dresel**

Universität Augsburg, Deutschland; [Martin.Daumiller@phil.uni-augsburg.de](mailto:Martin.Daumiller@phil.uni-augsburg.de)

Im Zusammenhang mit einem generellen Interesse an der Quantifizierung der Hochschultätigkeit hat in letzter Zeit auch das Interesse an der Erfassung von Forschungsleistung zugenommen (Hennig-Thurau et al., 2004). Angesichts der hohen fachlichen Spezialisierung ist dies ein komplexes Unterfangen, das hohe Anforderungen an die Gewinnung adäquater und intersubjektiv nachvollziehbarer Kriterien zur Beurteilung der Forschungsleistung der einzelnen Forscher(innen) stellt. Beispielsweise umfasst Forschungsleistung sowohl quantitative als auch qualitative Aspekte, wobei gerade zur fachübergreifenden Erfassung ein Fokus auf nur einen dieser Aspekte unzureichend ist (Müller, 2012). Außerdem existieren noch keine inhaltlich allgemeingültigen Instrumente. Dementsprechend gibt es bislang nur wenige Erkenntnisse hinsichtlich personeller Bedingungsfaktoren der Forschungsleistung, die über verschiedene Fachrichtungen hinweg empirisch getestet wurden.

Ziel des vorliegenden Beitrags war es daher, ein ökonomisches, fachübergreifendes Messinstrument zur Erfassung der Forschungsleistung zu entwickeln und zu überprüfen sowie erste fachübergreifende Erkenntnisse zu personellen Bedingungsfaktoren seitens der einzelnen Wissenschaftler(innen) zu erlangen.

Basierend auf Überlegungen des Wissenschaftsrats (2013) sowie von Ochsner, Hug und Daniel (2012, 2013, 2014) wurden dazu vier inhaltliche Bereiche der Forschungsleistung unterschieden: (1) Forschungsqualität (z.B. Publikationen, Innovation, Wissenschaftlichkeit), (2) Reputation (z.B. Anerkennung, Offenheit, Professional Activities), (3) Forschungsmöglichkeit (z.B. Drittmittel, Wirkung auf akademische Gemeinschaft) sowie (4) Transfer an außerwissenschaftliche Adressaten (z.B. Wissensvermittlung). Zu deren Erfassung wurden dreizehn qualitative sowie quantitative Indikatoren entwickelt, anhand derer sich Wissenschaftler(innen) prozentual respektive ihrer relevanten fachlichen Vergleichsgruppe mit gleichem Status einschätzen sollten (z.B. „In den letzten 5 Jahren forschte ich innovativer und origineller als \_\_\_% der übrigen Forscher(innen) meines Status in meiner Community“).

Um entsprechende Einschätzungen wurden 1 303 Wissenschaftler(innen) aus den Fächern Geschichte, Physik sowie Psychologie (als einschlägige Vertreter der Geistes-, Natur- und Sozialwissenschaften) gebeten. Die Befragten (durchschnittlich 36.4 Jahre alt, SD=10.6; 41% Promovierte, 19% Professores) wurden außerdem um Angaben zu relevanten kognitiven, affektiven und behavioralen Aspekten ihrer Forschungstätigkeit gebeten, nämlich zu ihren Einstellungen zur Hilfesuche (basierend auf Dickhäuser et al., 2007), ihrer Aufschiebetendenz (Klingsieck & Fries, 2012) sowie ihrem Erleben positiven Affekts (adaptiert von Keller et al., 2014). Vierhundertachtzig von ihnen füllten darüber hinaus einen Zusatzfragebogen zu verschiedenen bereits abgesicherten, jedoch unökonomischen und selektiven Maßen zu quantitativen (z.B. Anzahl verschiedener Publikationsarten, Summe eingeworbener Drittmittel, Tätigkeiten als Herausgeber) sowie qualitativen (basierend auf Ochsner et al., 2012; Hug et al., 2013) Teilaspekten der Forschungsleistung aus. Diese wurden anschließend zu einem Validierungsmaß aggregiert.

Deskriptive Statistiken verwiesen neben einer geringen Anzahl fehlender Werte (<4.3%) darauf, dass nahezu die gesamte theoretische Spannweite der neu entwickelten Skala angenommen wurde (1–96.7), wobei die Verteilung leicht linksschief war ( $M=41.7$ ,  $v=-2.2$ ) und substantielle interindividuelle Unterschiede vorlagen ( $SD=24.2$ ). Konfirmatorische Faktoranalysen bestätigten, dass eine Einfaktorenlösung sehr gut zu den Daten passte ( $df=65$ ,  $\chi^2=142.30$ ,  $CFI=.98$ ,  $TLI=.97$ ,  $RMSEA=.02$ ,  $SRMR=.02$ ) und Alternativmodelle mit mehreren Faktoren die Daten nicht statistisch signifikant besser beschrieben. Weiterhin lagen keine problematischen Itemladungen vor und die Skala wies eine hohe interne Konsistenz ( $\alpha=.94$ ) auf. Mehrgruppenmodelle, in denen das Messmodell für Geschichte, Physik und Psychologie simultan angepasst und sukzessive restringiert wurde, verwiesen auf strikte Messinvarianz für die drei Fachrichtungen. Das neu entwickelte Messinstrument korrelierte mit dem aggregierten Validierungsmaß für Forschungsleistung zu  $r=.78$ . Latente Korrelationen verwiesen fachübergreifend auf positive Zusammenhänge mit günstigen Einstellungen zur Hilfesuche sowie dem Erleben positiven Affekts und auf negative Assoziationen mit der Aufschiebetendenz. Dabei lagen keine substantiellen Unterschiede zwischen den Zusammenhängen der neu entwickelten, ökonomischen und fachübergreifenden Skala mit jenen des Validierungsmaßes vor ( $\Delta\beta<.08$ ).

Zusammenfassend verweisen die Ergebnisse auf eine hohe Testgüte der neu entwickelten Skala. Sie kann daher als tragfähige und ökonomische Möglichkeit zur Erfassung von Forschungsleistung in unterschiedlichen Wissenschaftsdisziplinen erachtet werden und damit zur fachübergreifenden Erforschung von Bedingungen der Forschungsleistung von Wissenschaftler(inne)n beitragen.

## A03

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus S104

### Wie entwickelt sich die Selbstwirksamkeitserwartung und die Berufszufriedenheit von Studierenden und wie belastet fühlen sie sich während der Praxisphase? Ergebnisse einer Längsschnittstudie.

**Kira Elena Weber**

Leuphana Universität Lüneburg, Deutschland; [kweber@leuphana.de](mailto:kweber@leuphana.de)

#### Theoretischer Hintergrund

Zur professionellen Kompetenz von Lehrkräften gehören neben dem Wissen und Können einer Lehrkraft auch Werthaltungen und Überzeugungen (Baumert & Kunter, 2006). Zu den Überzeugungen zählen u.a. Selbstwirksamkeitserwartungen (vgl. Bandura 1997), die positive Auswirkungen auf motivationale Ressourcen von Lehrkräften haben können. So sind Lehrkräfte mit einer hohen Selbstwirksamkeit resilienter gegenüber beruflichen Belastungen, sowie negativen Gefühlen und Stress, zeigen eine größere Berufszufriedenheit, mehr Enthusiasmus im Unterricht und eine höhere Wertschätzung des Unterrichtens (Goddard, Hoy & Woolfolk Hoy, 2004; Baumert & Kunter, 2006; Klassen & Durksen, 2014; Oberle, Weschenfelder & Weißeno, 2015; Schüle, Besa, Schriek & Arnold, 2016). Ein zentrales Element der universitären Lehrerbildung in Deutschland stellen die Praxisphasen dar (Gröschner & Müller et al., 2015). Bisherige Studien kommen zu widersprüchlichen Ergebnissen im Hinblick auf die Entwicklung der Selbstwirksamkeit in Schulpraktika (Schüle et al., 2016). Da die Praxisphase ebenso wie die Berufseinstiegsphase mit einer Vielzahl von Anforderungen an angehende Lehrkräfte verbunden ist, kann diese zu einem erhöhten Belastungserleben führen (Brouwer & ten Brinke, 1995). Auch hier zeigen sich weitere Forschungsdesiderata im Hinblick auf Langzeitpraktika in der Lehrerbildung.

#### Fragestellung

Vor dem eingangs skizzierten theoretischen Hintergrund, stellen sich folgende Forschungsfragen: (a) wie entwickeln sich die Selbstwirksamkeitserwartung, die Berufszufriedenheit und das Belastungserleben im Rahmen der Praxisphase, (b) wie hängen die einzelnen Konstrukte miteinander zusammen, (c) welche Rolle spielen weitere Kontrollvariablen (u.a. bisherige Unterrichtserfahrungen, Erfahrungen mit heterogenen Klassen, Qualität der Praktikumsbetreuung).

#### Methode

Im Rahmen der in Niedersachsen eingeführten Neustrukturierung der Masterstudiengänge für die Lehramter an Grund-, Haupt- und Realschulen (GHR 300) absolvieren die Studierenden ein 18wöchiges Langzeitpraktikum. Die Lehramtsstudierenden wurden während der Praxisphase zu drei Messzeitpunkten (MZP) befragt. Der erste MZP (n = 185) fand vor Beginn des Praktikums statt, der zweite MZP (n = 114) in der Mitte des Praktikums und der dritte MZP (Oktober 2017) drei Monate nach dem Praktikum. Für die Operationalisierung der Selbstwirksamkeitserwartungen wurde die eindimensionale Skala der Lehrer-Selbstwirksamkeitserwartung (SWLE) von Schwarzer & Schmitz (1999), sowie zwei Subskalen aus den Kurzskalen zur inklusiven Einstellung und Selbstwirksamkeit von Lehrpersonen (KIESEL) von Bosse & Spörer (2014) verwendet. Das allgemeine Belastungserleben wurde in Anlehnung an Prümper, Hartmannsgruber und Frese (1995) sowie durch die deutsche Version des Maslach Burnout Inventars (MBI von Maslach, Jackson & Leiter, 1986) erfragt. Die Berufszufriedenheit wurde in Anlehnung an Ditton (BZUF, 2001) erfasst. Neben den verwendeten Messinstrumenten wurde nach Geschlecht, Alter, Studiengang, Studienfächern, Anzahl der Kinder, Unterrichtserfahrungen, Nebenjobs, theoretischen und praktischen Erfahrungen im Umgang mit Leistungsheterogenität und der Qualität der Praktikumsbetreuung gefragt.

#### Ergebnisse

Zwischen dem ersten und zweiten MZP konnten moderate Zuwächse bei der eindimensionalen Lehrer-Selbstwirksamkeitserwartung (MT1 = 4.52; MT2 = 4.73; sechsstufige Skala) und der Selbstwirksamkeit bezogen auf inklusiven Unterricht (MT1 = 3.87; MT2 = 4.01; sechsstufige Skala) festgestellt werden. Bei der Selbstwirksamkeit im Hinblick auf den Umgang mit Unterrichtsstörungen zeigte sich der höchste Zuwachs (MT1 = 3.78; MT2 = 4.24; sechsstufige Skala). Die Berufszufriedenheit war zu beiden MZP hoch (MT1 = 5.06; MT2 = 5.03; sechsstufige Skala). Eine Veränderung konnte nicht festgestellt werden, ebenso wenig wie beim Belastungserleben (MT1 = 3.49; MT2 = 3.56; fünfstufige Skala) sowie bei der emotionalen Erschöpfung (MT1 = 2.92; MT2 = 2.92; sechsstufige Skala). Es bestanden signifikante Korrelationen zwischen der Selbstwirksamkeitserwartung und der Berufszufriedenheit angehender Lehrkräfte ( $r = .50, p < .001$ ). Eine signifikante, negative Korrelation ( $r = -.23, p < .01$ ) ließ sich zwischen dem allgemeinen Belastungserleben und der Berufszufriedenheit feststellen. Auch die Qualität der Praktikumsbetreuung korrelierte erwartungsgemäß mit der Selbstwirksamkeit ( $r = .40, p < .001$ ), der Berufszufriedenheit ( $r = .38, p < .001$ ) und dem Belastungserleben ( $r = -.31, p < .01$ ). Detaillierte Ergebnisse zu den drei MZP werden auf der Konferenz präsentiert und praktische Implikationen für die Lehrerbildung diskutiert.

### Fachspezifische Selbstwirksamkeitsüberzeugungen angehender Naturwissenschaftslehrkräfte hinsichtlich inklusiven Unterrichts

**Melanie Basten<sup>1</sup>, Susanne Menzel<sup>2</sup>**

<sup>1</sup>Universität Bielefeld, Deutschland; <sup>2</sup>Universität Osnabrück, Deutschland; [melanie.basten@uni-bielefeld.de](mailto:melanie.basten@uni-bielefeld.de)

Die Einführung eines inklusiven Schulsystems (Art.24 BGBl 2008) stellt (angehende) Lehrkräfte vor zusätzliche Herausforderungen (Melzer et al., 2015). Für die Umsetzung inklusiven Unterrichts wird die Einstellung der Lehrpersonen als relevant erachtet (EA, 2011). Als stärkster Prädiktor der Einstellung gilt die Selbstwirksamkeit der Lehrkräfte hinsichtlich der Anforderungen inklusiven Unterrichtens (Sharma et al., 2012).

Im vorliegenden Projekt soll untersucht werden, ob Selbstwirksamkeitsüberzeugungen fachspezifisch erhoben werden sollten, da die Herausforderungen an inklusive Lehrkräfte fachspezifisch sind und in die didaktische Gestaltung des Unterrichts wirken (Nehring & Bohlmann, 2016). Obwohl naturwissenschaftsdidaktische Theorien noch nicht für inklusive Settings ausgearbeitet sind (ebd.), werden in der Literatur Anforderungen an inklusive Naturwissenschafts-Lehrkräfte formuliert. Diese beinhalten Anforderungen, die für alle Fächer (bspw. Weiß et al., 2013; Wilbert et al., 2016), alle naturwissenschaftlichen Fächer (bspw. Markic & Abels, 2013; Stroh, 2015) oder tatsächlich ein einzelnes Fach, hier Biologie (bspw. Dubs, 2009; Ortland, 2008; Zentel & Michaelys, 2015), Gültigkeit haben. Eine qualitative Studie ergab, dass Studierende hauptsächlich Anforderungen an inklusive Biologielehrkräfte erkennen, die in allen Fächern oder allen Naturwissenschaften bedeutsam sind. Sie sehen Anforderungen hinsichtlich der differenzierten Planung und Durchführung von Unterricht, Classroom-Management, Kooperationen (vgl. Wilbert et al., 2016) und der Planung und Durchführung von Experimenten (Basten & Menzel, 2017). In der

vorliegenden Studie wurde basierend auf diesen Befunden eine erste Fragebogenversion zur Selbstwirksamkeit hinsichtlich naturwissenschaftsspezifischer inklusiver Anforderungen erstellt und erprobt.

Die Fragestellungen lauten: F1: Wie selbstwirksam fühlen sich Lehramtsstudierende hinsichtlich inklusiven Unterrichts? F2: Für wie speziell relevant halten Lehramtsstudierende die abgefragten Anforderungen für den inklusiven naturwissenschaftlichen Unterricht? F3: Sagen die fachspezifischen Selbstwirksamkeitsüberzeugungen die Einstellung der Studierenden besser vorher als fächerübergreifende inklusive Selbstwirksamkeitsüberzeugungen?

Im Rahmen einer quantitativen Querschnittstudie füllten 170 Lehramtsstudierende (79% weiblich) der Biologie und des Sachunterrichts von fünf Universitäten (Alter:  $M=22,4$  ( $SD=2,8$ ); Fachsemester:  $M=5,5$  ( $SD=3,7$ )) einen Fragebogen aus. Dieser umfasste naturwissenschaftsspezifisch formulierte Anforderungen hinsichtlich Unterricht (UN\_INU,  $k=11$ , Cronbachs-Alpha=.87), Classroom-Management (CM\_INU,  $k=6$ , Cronbachs-Alpha=.82) und Kooperation (KO\_INU,  $k=9$ , Cronbachs-Alpha=.82), die zwar in allen Fächern Relevanz haben (vgl. Wilbert, et al., 2016), in den Naturwissenschaften jedoch spezifische Realisierungen aufweisen können. Je ca. ein Drittel der Studierenden beantwortete zu jedem Item eine Frage zur spezifischen Relevanz der Anforderung für naturwissenschaftlichen Unterricht ( $n=61$ ), für Schüler\_Innen mit geistigen ( $n=54$ ) oder körperlichen Besonderheiten ( $n=55$ ) (Antwortformat dichotom). Außerdem wurden die allgemeinen inklusiven Selbstwirksamkeitsskalen des KIESEL (Bosse & Spörer, 2014; Cronbachs-Alphas: Unterrichtsgestaltung KI\_UG=.73; Unterrichtsstörungen KI\_ST=.72; Elternarbeit KI\_EL=.79) und die Einstellung der Studierenden mit drei Items (basierend auf ORI deutsch (Benoit & Bless, 2014) und KIESEL; Cronbachs-Alpha=.75) erhoben.

Die Studierenden berichteten eine eher geringe Selbstwirksamkeit in den naturwissenschaftsspezifischen (UN\_INU:  $M=3,3$  ( $SD=0,59$ ); CM\_INU:  $M=3,3$  ( $SD=0,58$ ); KO\_INU:  $M=3,7$  ( $SD=0,62$ ); max.6) sowie den allgemeinen KIESEL-Skalen (KI\_UG:  $M=2,9$  ( $SD=0,81$ ); KI\_ST:  $M=3,1$  ( $SD=0,58$ ); KI\_EL:  $M=3,3$  ( $SD=0,58$ ); max.6). Die jeweilige fachspezifische Selbstwirksamkeit wurde höher eingeschätzt als die entsprechende allgemeine (UN\_INU>KI\_UG ( $Z=7,4$ ,  $p<.001$ ); CM\_INU>KI\_ST ( $Z=5,2$ ,  $p<.001$ ); KO\_INU>KI\_EL ( $Z=7,1$ ,  $p<.001$ )). Die fachspezifisch zugeschnittenen Anforderungen wurden von den Studierenden zudem als spezifisch relevanter für den naturwissenschaftlichen Unterricht eingeschätzt als die allgemein formulierten des KIESEL (UN\_INU:  $M=0,63$  ( $SD=0,30$ ); CM\_INU:  $M=0,45$  ( $SD=0,33$ ); KO\_INU:  $M=0,39$  ( $SD=0,31$ ); KI\_UG:  $M=0,37$  ( $SD=0,36$ ); KI\_ST:  $M=0,22$  ( $SD=0,34$ ); KI\_EL:  $M=0,12$  ( $SD=0,27$ ); max.1; UN\_INU>KI\_UG ( $Z=4,8$ ,  $p<.001$ ); CM\_INU>KI\_ST ( $Z=4,6$ ,  $p<.001$ ); KO\_INU>KI\_EL ( $Z=5,6$ ,  $p<.001$ )). Auch die Einstellung der Studierenden zu inklusivem Unterricht war sehr niedrig ( $M=1,3$  ( $SD=0,8$ ); max.6). Ihr stärkster Prädiktor war die Selbstwirksamkeit hinsichtlich naturwissenschaftsspezifischen Classroom-Managements (Beta CM\_INU=0,30). Der zweite signifikante Prädiktor war die allgemeine Selbstwirksamkeit hinsichtlich inklusiver Unterrichtsgestaltung (Beta KI\_UG=0,27). Die Varianzaufklärung war eher gering ( $R^2=.15$ ).

Die vorliegende Studie zeigt, dass Studierende naturwissenschaftsspezifisch ausformulierte Anforderungen als relevanter für den inklusiven Naturwissenschafts-Unterricht einschätzen als allgemein formulierte. Die Einstellung wurde jedoch von einer allgemein formulierten Selbstwirksamkeitsskala ähnlich gut vorhergesagt wie von einer fachspezifisch ausformulierten. Die geringe Varianzaufklärung (vgl. Sharma et al., 2012) ist vermutlich darauf zurückzuführen, dass die Einstellung in dieser Studie sehr eng gefasst und deutlich von (Selbstwirksamkeits-)Überzeugungen abgegrenzt erhoben wurde (vgl. Basten et al., 2016). Im nächsten Schritt sollen Selbstwirksamkeitsdimensionen auf ihre Aufklärungsleistung hin untersucht werden, die spezifischer für den naturwissenschaftlichen Unterricht sind.

## **„Ich würde ihm Mut machen, aber eine schlechte Note geben“: Inklusionsbezogene Vorstellungen angehender Lehrer\*innen**

**Veronika Schmid**

Heidelberg School of Education, Deutschland; [veronika.schmid@heiedu.ph-heidelberg.de](mailto:veronika.schmid@heiedu.ph-heidelberg.de)

In den vergangenen Jahrzehnten hat sich im deutschen Bildungssystem ein neuer Umgang mit Vielfalt etabliert (Prenzel, 1993; Sliwka 2010). Inklusion wird zunehmend als Querschnittsaufgabe professionellen pädagogischen Handelns verstanden, auf deren Bewältigung Lehramtsstudierende bereits während ihres Studiums vorbereitet werden sollen. Im Rahmen eines programmatisch „weiten“ Verständnisses von Inklusion geht es entsprechend darum, zukünftigen Lehrer\*innen ein Verständnis von Heterogenität/Diversität als Normalfall zu vermitteln (Steinbach 2016; Budde & Hummrich 2013). Daraus resultieren besondere Herausforderungen an die bisherige pädagogische Praxis. Als relevant wird in diesem Zusammenhang das Vermögen gesehen, die eigenen Vorstellungen und pädagogischen Praktiken in Hinblick auf mögliche diskriminierende Wirkungsweisen zu hinterfragen. Um unterschiedliche Argumentationsmuster von Lehramtsstudierenden zum Umgang mit Diversität und daraus resultierende Handlungskonzepte zu erfassen, wurde eine qualitative Vignettenstudie mit Prätest-Posttest-Messung entwickelt, die in drei Seminaren zum Themenfeld „Inklusion, Bildungsgerechtigkeit und Diversität“ durchgeführt wurde. Als Erhebungsinstrument dienten Dilemma-Vignetten, die im Umgang mit Diversität typische Ziel- oder Wertkonflikte im Praxisfeld Schule thematisieren.

Der Vignettenstudie liegen folgende Forschungsfragen zugrunde: 1. Art der Argumentation: Wie argumentieren die Studierenden angesichts der dargestellten Dilemma-Situationen? Ist es möglich, unterschiedliche inklusionsbezogene Argumentationsmuster zu identifizieren? 2. Stabilität der Argumentation: Verändern sich die Argumentationen, nachdem sich die Studierenden im Seminar intensiv mit den Themen Inklusion und Diversität beschäftigt haben? 3. Zusammenhänge zu übergeordneten Einstellungen und Überzeugungen: Sind die inklusionsbezogenen Argumentationsmuster Teil übergeordneter Einstellungs- und Überzeugungs-systeme?

Die Vignettenstudie wurde an der Pädagogischen Hochschule Heidelberg im Wintersemester 2016/17 und dem darauffolgenden Sommersemester 2017 durchgeführt ( $t1: n = 105$ ,  $t2: n = 63$ ). Die teilnehmenden Studierenden befanden sich am Beginn ihres Studiums (Primarstufe, Sekundarstufe I, Sonderpädagogik). Den Befragten wurden jeweils zu Veranstaltungsbeginn und zum Veranstaltungsende der jeweils analog aufgebauten Seminare Dilemma-Vignetten vorgelegt. Anschließend wurden die Studierenden aufgefordert, schriftlich zu begründen, wie sie sich als Lehrkraft in der dargestellten Situation verhalten würden. Darüber hinaus bearbeiteten die Studierenden u.a. Skalen zu Autoritarismus, Sozialer Dominanzorientierung und Vorurteilen.

Die Stellungnahmen der Befragten wurden inhaltsanalytisch und typenbildend ausgewertet. Dabei konnten drei unterschiedliche Argumentationsmuster bzw. Modi des Umgangs mit Diversität herausgearbeitet werden: Bedrohung der Ordnung/Störung der Routine, Formale Akzeptanz des Anderen, Wertschätzung des Anderen. Vergleicht man die Antworten der Studierenden des ersten und zweiten Erhebungszeitpunkts miteinander, zeigt sich, dass die Argumentationsmuster vergleichsweise stabil bleiben. Hierbei erwies sich insbesondere die Veränderungsresistenz des auch häufigsten Argumentationsmusters „Formale Akzeptanz des Anderen“ als signifikant. Das jeweils dominante Argumentationsmuster der



Studierenden ging mit unterschiedlichen Ausprägungen auf den Skalen zu übergeordneten Einstellungsmustern einher (z.B. Soziale Dominanzorientierung, Etabliertenvorrechte). Dabei zeigte sich u.a., dass Studierende, die zum Argumentationsmuster „Bedrohung der Ordnung/Störung der Routine“ tendierten, eine höhere soziale Dominanzorientierung aufwiesen als Studierende, die andere Argumentationsmuster präferierten.

Insgesamt sprechen die Befunde dafür, dass die inklusionsbezogenen Argumentationsmuster der Studierenden Teil von übergeordneten ideologischen Wert- und Überzeugungssystemen sind, die vergleichsweise stabil sind. Die theoretische Auseinandersetzung mit den Themen Inklusion und Diversität in Seminarform ist zwar wichtig, aber offenbar nicht ausreichend, um Studierende zu einem reflektierten Umgang mit Heterogenität anzuleiten. Für die Professionalisierung angehender Lehrer\*innen bedeutet dies, dass noch stärker darüber nachgedacht werden muss, mit welchen alternativen (hochschul-)didaktischen Lehr- und Lernarrangements dies möglicherweise besser gelingen kann.

## **Entwicklung von Überzeugungen bei Studierenden durch inklusionsorientierte Lehrveranstaltungen und Theorie-Praxis-Vernetzung im Lehramtsstudium**

**Saskia Opalinski, Katja Scharenberg**

Pädagogische Hochschule Freiburg, Deutschland; [saskia.opalinski@ph-freiburg.de](mailto:saskia.opalinski@ph-freiburg.de)

Im Kontext inklusiver Bildung werden gleichermaßen Fragen einer inklusionsorientierten Lehrer\*innenbildung und die Bedeutung inklusionsbezogener Überzeugungen bei (angehenden) Lehrkräften diskutiert (Heinrich, Urban & Werning, 2013). Überzeugungen werden konzeptuell als Element professioneller Kompetenz von Lehrkräften verstanden, beeinflussen deren Handeln und wirken auf die Leistungen von Schüler\*innen (Baumert & Kunter, 2006; Moser, 2017). Bezogen auf inklusive Bildung liegen mittlerweile zahlreiche, wenn auch widersprüchliche, Befunde zu den Überzeugungen von (angehenden) Lehrkräften vor, die jedoch überwiegend aus querschnittlichen Untersuchungen stammen und damit keine Aussagen zur Entstehung bzw. Veränderung der Überzeugungen erlauben (de Boer, Pijl & Minnaert, 2011). Es gibt aber empirische Hinweise darauf, dass inklusionsorientierte Lehrveranstaltungen zu positiveren Überzeugungen bezogen auf Inklusion führen können (z.B. Kopp, 2009; Lancaster & Bain, 2007; Laubner & Lindmeier, 2017). Gleichzeitig sind inklusionspädagogische Anteile im Lehramtsstudium, ungeachtet der großen Bedeutung, die den Themenkomplexen Heterogenität und Inklusion in der Praxis in Schule und Unterricht zukommt, noch immer deutlich unterrepräsentiert und verbleiben üblicherweise ohne Theorie-Praxis-Vernetzung. Die Implementation von Praxisanteilen erwies sich jedoch als relevant für die Herausbildung einer kritisch-reflexiven, und damit im Kontext inklusiver Bildung bedeutsamen, Haltung (Van Laarhoven, Munk, Lynch, Bosma & Rouse, 2007).

Im Beitrag soll mittels Prä-Post-Design der Frage nachgegangen werden, inwiefern durch inklusionspädagogische Lehrveranstaltungen und konzeptuell verankerte Theorie-Praxis-Vernetzung im bildungswissenschaftlichen Studienanteil der Lehrämter Primar- und Sekundarstufe Überzeugungen von Studierenden entwickelt werden können und inwiefern dabei der Lehramtsstudiengang (Primar- oder Sekundarstufe) von Bedeutung ist. Erwartet wird eine positive Veränderung hinsichtlich der Inklusionsorientierung und der Orientierung auf die individuellen Lernausgangslagen der Schüler\*innen. Es werden außerdem positivere Überzeugungen der Studierenden des Lehramts Primarstufe angenommen, da die Grundschule der Lernort mit einer sehr heterogenen Schülerschaft ist.

Die Datenerhebungen (n=47) erfolgten jeweils zu Beginn und Ende des Semesters in zwei einsemestrigen, zweistündigen inklusionspädagogischen Seminaren, die sich durch eine konzeptuell in das Seminkonzept eingebettete Exkursion auszeichneten. Im Rahmen der Exkursion hospitierten die Studierenden an Schulen, die als good practice-Beispiele im Umgang mit der Heterogenität ihrer Schüler\*innen und der Realisierung inklusiver Bildung betrachtet werden können. Erfasst wurden die Überzeugungen der Studierenden mittels des ‚Beliefsinventar Lehrkräfte im Bereich schulischer Förderung (BILF)‘ (Moser, Kuhl, Redlich & Schäfer, 2014). Die Überzeugungen der Studierenden wurden anhand einer 4-stufigen Likert-Skala mit insgesamt 35 Items erfasst, verteilt auf die drei Subskalen ‚Inklusionsorientierung (IO)‘ (12 Items) ( $\alpha=.79$ ), ‚individuell förderbezogen, Lebenslagen und diagnostisch orientiert (IFLD)‘ (14 Items) ( $\alpha=.61$ ) sowie ‚psychiatrisch-therapeutisch orientiert (PT)‘ (6 Items) ( $\alpha=.69$ ). Die Auswertung erfolgte mittels multivariater Varianzanalysen.

Erste Befunde zeigen einen signifikanten Haupteffekt für den Besuch der Lehrveranstaltung auf die Inklusionsorientierung ( $p<.05$ ;  $\eta^2=.16$ ). Dagegen erwies sich der Seminarbesuch hinsichtlich der Orientierung der Studierenden auf die individuellen Lernausgangslagen der Schüler\*innen sowie die psychiatrisch-therapeutische Orientierung als nicht bedeutsam. Wechselwirkungen mit dem Geschlecht der Studierenden oder dem Lehramtsstudiengang konnten nicht gezeigt werden. Insgesamt zeichnen sich die Studierenden nach Besuch des Seminars durch eine deutliche Inklusionsorientierung (Skala IO,  $M=3.12$ ,  $SD=0.42$ ) bei gleichzeitig unveränderter, deutlicher Orientierung auf die individuellen Lernausgangslagen der Schüler\*innen (Skala IFLD,  $M=3.23$ ,  $SD=0.30$ ) und einer moderaten psychiatrisch-therapeutischen Orientierung (Skala PT,  $M=2.86$ ,  $SD=0.50$ ).

Anhand der Befunde und bezogen auf eine Theorie-Praxis-Vernetzung werden in unserem Beitrag abschließend professionstheoretische und hochschuldidaktische Fragen sowie Aspekte der organisationalen Realisierung inklusionsorientierter Lehrer\*innenbildung diskutiert.

## A04

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Rosshofgasse, S02

### Wahrnehmung von Quantitäten in bildlichen statistischen Diagrammen: Linear, quadratisch oder kubisch?

Tobias Rolfes

Universität Koblenz-Landau, Deutschland; [rolfes@uni-landau.de](mailto:rolfes@uni-landau.de)

Theoretischer Hintergrund. Die Visualisierung von statischen Daten mit Hilfe von graphisch gestalteten Diagrammen ist in den Medien weit verbreitet. Beispielsweise wurde in der ADAC-Zeitschrift Motorwelt („Der erste Schritt“, 2017) der Stickoxid-Ausstoß verschiedener Dieselmotoren anhand unterschiedlich großer Abgaswolken dargestellt. Unklar ist dabei, in welcher Weise die Relationen der Quantitäten in den Diagrammen von Menschen wahrgenommen werden. Tuft (2008) weist darauf hin, dass gerade bei Flächendiagrammen häufig das in einem Diagramm dargestellte Quantitätsverhältnis nicht mit dem tatsächlichen Quantitätsverhältnis übereinstimmt. De Bock et al. (2002, 2007) konnten in zahlreichen Untersuchungen nachweisen, dass Schülerinnen und Schüler einer „Linearitätsillusion“ unterliegen. So berechneten die Testpersonen, dass man für das Ausmalen eines dreimal so großen Weihnachtsmannes an einem Fenster auch das Dreifache an Farbe benötigt, obwohl in Wirklichkeit das Neunfache an Farbe benötigt wird. Dagegen liegt bei Verwendung von Abbildungen von dreidimensionalen Objekten aus theoretischer Sicht ein kubischer Zusammenhang nahe. So stellt eine doppelt so große Mülltonne die achtfache Müllmenge dar, da sich das Volumen verachtfacht, wenn sich Länge, Breite und Höhe jeweils verdoppeln (vgl. Krämer, 2008). Daher stellt sich die zentrale Frage, wie Quantitäten in Diagrammen von Menschen wahrgenommen werden.

Hypothesen. H1/2/3: Skalierungen in Diagrammen werden linear/quadratisch/kubisch wahrgenommen, d. h. eine Skalierung um 1.3 resultiert in einer wahrgenommenen Veränderung der Quantität mit dem Faktor 1.3/1.69/2.197.

Methode. N = 63 Studierende nahmen an dem Experiment teil. Das mittlere Alter betrug 21.4 Jahre (SD = 2.2). Drei Testlets wurden entwickelt. In dem ersten Testlet repräsentierte eine Strichzeichnung einer Wolke die Menge an Kohlendioxid-Emission einer Fabrik. Im zweiten Testlet stellte ein Foto einer Mülltonne die Müllmenge dar, die ein Haushalt produziert. Im dritten Testlet repräsentierte eine Strichzeichnung eines Stücks Würfelzucker die Menge an Zucker, die eine Person in einem bestimmten Land konsumiert. Jedes Testlet bestand aus vier Items. Bei jedem Item wurden auf der linken Seite 100 Einheiten durch eine Wolke, eine Mülltonne oder ein Würfelzuckerstück repräsentiert. Auf der rechten Seite jedes Items wurde die Wolke, die Mülltonne oder das Würfelzuckerstück mit einem Faktor zwischen 0.3 und 1.9 skaliert. Die Aufgabe der Versuchspersonen bestand bei jedem Item darin, ohne Berechnungen intuitiv abzuschätzen, welche Quantität die Darstellung auf der rechten Seite des Items darstellt.

Ergebnisse. Im ersten Testlet führte eine Skalierung der 100 Tonnen Kohlendioxid repräsentierenden Wolke um den Faktor 0.3 zu einer mittleren wahrgenommenen Quantität von 10.9 Tonnen (95% KI [9.8, 11.9]). Entsprechend resultierten die Faktoren 0.9, 1.3 bzw. 1.6 in einer mittleren wahrgenommenen Quantitäten von 80.9 Tonnen (95% KI [79.0, 82.7]), 175 Tonnen (95% KI [159.3, 190.7]) bzw. 290.2 Tonnen (95% KI [265.4, 314.9]). Im zweiten Testlet bewirkte eine Skalierung des 100 Liter darstellenden Fotos einer Mülltonne um den Faktor 0.5, 0.7, 1.5 bzw. 1.9 zu einer mittleren wahrgenommenen Quantität von 28.4 Liter (95% KI [25.3, 31.4]), 54.5 Liter (95% KI [50.4, 58.6]), 189.3 Liter (95% KI [177.3, 201.2]) bzw. 344.4 Liter (95% KI [307.3, 381.5]). Im dritten Testlet führte eine Skalierung der 100 Gramm darstellende Strichzeichnung eines Würfelzuckerstücks um den Faktor 0.4, 0.8, 1.4 bzw. 1.8 zu einer mittleren wahrgenommenen Quantität von 19.2 Gramm (95% KI [16.9, 21.5]), 67.9 Gramm (95% KI [64.6, 71.2]), 195.0 Gramm (95% KI [181.9, 208.1]) bzw. 365.6 Gramm (95% KI [320.1, 411.2]). Somit musste bei allen Items die Hypothesen H1 und H3 verworfen werden, da die Quantitäten, die aus einer linearen bzw. kubischen Quantitätsveränderung resultieren würden, außerhalb der jeweiligen 95%igen Konfidenzintervalle lagen. Bei fünf der zwölf Items konnte die Hypothese H2 nicht abgelehnt werden. Wurde bei jedem Item der Zusammenhang zwischen dem Skalierungsfaktor  $x$  und der mittleren wahrgenommenen Quantität  $Q$  mit Hilfe einer Potenzfunktion  $Q = x^k$  bestimmt, so nahmen die Exponenten  $k$  Werte zwischen 1.6 und 2.3 an.

### Überschätzungen verringern, Transfer fördern: Refutationtexte können den negativen Einfluss von statistischen Fehlkonzepten auf die Genauigkeit der Selbsteinschätzung des Textverständnisses kompensieren und den Wissenstransfer begünstigen

Anja Prinz, Stefanie Golke, Jörg Wittwer

Albert-Ludwigs-Universität Freiburg, Deutschland; [anja.prinz@ezw.uni-freiburg.de](mailto:anja.prinz@ezw.uni-freiburg.de)

Im Bereich der Statistik haben Lernende häufig Fehlkonzepte (vgl. Huck, 2016). Statistische Fehlkonzepte führen nicht nur zu einem schlechteren Verständnis eines Statistiktextes, sondern auch dazu, dass Lernende ihr Textverständnis überschätzen (Autoren, 2017). In der Folge investieren die Lernenden vermutlich nicht genügend Anstrengung und Zeit, um ihr Verständnis zu verbessern (z. B. Thiede, Anderson & Theriault, 2003). Refutationtexte, die Fehlkonzepte explizit adressieren und deren Unangemessenheit durch die normativ korrekte Erklärung herausstellen, sind gegenüber typischen Lehrbuchtexten effektiver, um das Textverständnis zu fördern (vgl. Guzzetti, Snyder, Glass & Gamas, 1993). Ob Refutationtexte auch dazu geeignet sind, die Genauigkeit der Selbsteinschätzung des Textverständnisses zu begünstigen, ist bislang jedoch unklar. Offen ist ebenfalls, inwiefern Refutationtexte den Transfer von Wissen begünstigen. Im Bereich der Statistik erscheint dies besonders relevant. Lernende sollten beispielsweise in der Lage sein, ihr Statistikwissen anzuwenden, um empirische Forschungsergebnisse zu verstehen.

Um diese offenen Fragen zu beantworten, haben wir eine experimentelle Studie mit between-subjects Design durchgeführt. Zunächst bearbeiteten N = 80 Lehramtsstudierende einen Fehlkonzepttest bezüglich Kovarianz. Anschließend lasen die Studierenden einen Statistiktext über Kovarianz, wobei sie zufällig der Standardtext-Bedingung ( $n = 39$ ) oder der Refutationtext-Bedingung ( $n = 41$ ) zugeordnet wurden. Im Refutationtext wurde ein Fehlkonzept zunächst beschrieben, dessen Unangemessenheit geäußert und anschließend die korrekte Erklärung geliefert. Im Standardtext wurde dementsprechend ausschließlich die korrekte Erklärung gegeben, ohne das entsprechende Fehlkonzept explizit zu nennen. Nach dem Lesen des Statistiktextes schätzten die Studierenden ihr Verständnis ein und beantworteten Inferenzfragen. Darauf folgte ein Anwendungstext, in dem Forschungsergebnisse in Bezug auf Kovarianz zwischen verschiedenen Variablen (z. B. Informiertheit über Umweltprobleme und Leistung in Naturwissenschaften) aus der Bildungsforschung präsentiert wurden. Schließlich schätzten die Studierenden ihr Verständnis dieses Textes ein und beantworteten Inferenzfragen.

Im Hinblick auf den Statistiktext zeigten die Ergebnisse, dass in der Standardtext-Gruppe Fehlkonzepte zu einem niedrigeren Textverständnis ( $\beta = -.57, p < .001$ ) sowie zu einer größeren Überschätzung des Textverständnisses ( $\beta = .44, p = .005$ ) führten. Im Gegensatz dazu hatten in der Refutationstext-Gruppe Fehlkonzepte keinen signifikanten Einfluss auf das Textverständnis ( $\beta = .10, p = .528$ ) und auf die Genauigkeit der Selbsteinschätzung des Textverständnisses ( $\beta = -.04, p = .798$ ). Diese Befunde machen deutlich, dass ein Refutationstext den negativen Einfluss von Fehlkonzepten nicht nur auf das Textverständnis, sondern auch auf die Genauigkeit der Selbsteinschätzung des Textverständnisses kompensieren kann. Im Hinblick auf den Anwendungstext zeigten die Ergebnisse, dass nach Lesen des Standardtextes Fehlkonzepte auch hier zu einem niedrigeren Textverständnis ( $\beta = -.40, p = .010$ ) sowie zu einer größeren Überschätzung des Textverständnisses ( $\beta = .33, p = .031$ ) führten. Dementgegen zeigte sich, dass nach Lesen des Refutationstextes Fehlkonzepte sogar tendenziell mit einem besseren Textverständnis ( $\beta = .31, p = .052$ ) und einer genaueren Selbsteinschätzung des Textverständnisses ( $\beta = -.47, p = .003$ ) einhergingen. Dies verdeutlicht, dass die kompensatorischen Effekte des Refutationstextes auch noch zum Tragen kommen, wenn Lernende ihr neu erworbenes Wissen nutzen, um einen weiteren anwendungsbezogenen Text zu verstehen und ihr Verständnis dieses Textes einzuschätzen. Zusammengefasst sprechen die Befunde dafür, dass Refutationstexte eine effektive instruktionale Methode darstellen, um das selbstregulierte Lernen aus Texten und den Wissenstransfer zu begünstigen. In Domänen, in denen Lernende häufig Fehlkonzepte aufweisen, sollten Refutationstexte deshalb vermehrt eingesetzt werden.

### Bedingungsfaktoren für den Transfer metakognitiver Strategien vom Training in den Fachunterricht

Corinna Schuster<sup>1</sup>, Ferdinand Stebner<sup>1</sup>, Joachim Wirth<sup>1</sup>, Detlev Leutner<sup>2</sup>

<sup>1</sup>Ruhr-Universität Bochum, Deutschland; <sup>2</sup>Universität Duisburg-Essen; [corinna.schuster@rub.de](mailto:corinna.schuster@rub.de)

Vielfach wurde durch Forschung die Bedeutsamkeit von Trainings zur Förderung der Kompetenz des selbstregulierten Lernens (SRL) betont (u.a. Artelt et al., 2003). Um das Strategiewissen, welches Schülerinnen und Schüler (SuS) in Trainings zum SRL erlangen, auch außerhalb von Trainings nutzen zu können, müssen sie erlernte Strategien in den Fachunterricht transferieren. Transfer wird dabei als Prozess beschrieben, in dem vorheriges Lernen verallgemeinert sowie kognitive Muster und Formen auf neue Probleme, Kontexte und Inhalte ausgerichtet werden (Saxe, 1989; White & Frederiksen, 1998). Dabei beinhaltet naher Transfer Situationen, in denen eingeübte Routinen auf im Vergleich zum ursprünglichen Lernkontext ähnliche Situationen angewendet werden, während beim weiten Transfer die Übertragung nicht von oberflächlichen Ähnlichkeiten abhängt (Perkins & Salomon, 1988).

Bisherige Forschung zum Transfer beim SRL zeigt, dass SuS vor allem der nahe Transfer gelingt (u.a. Rump, Leopold & Leutner, 2005). Weiter Transfer, wie z.B. von einem Training in den Fachunterricht, kann entweder selten nachgewiesen werden (Mokhlesgerami, Souvignier, Rühl & Gold, 2007) oder findet auf einem sehr geringen Level statt (Schuster et al., 2007). Folglich stellt sich die Frage, wovon es abhängt, dass dieser weite Transfer gelingt. Im Sinne von Klauer (2011) ließe sich annehmen, dass dies neben der Ähnlichkeit zwischen Lern- und Transferaufgabe vor allem auf die individuellen Voraussetzungen der Lernenden, wie Vorwissen, Intelligenz und Leistungsniveau, zurückzuführen ist. Weiter Transfer kann demnach eher von leistungsstarken als von leistungsschwachen SuS geleistet werden.

Um dies zu untersuchen, erhielten insgesamt 96 SuS ( $M=10.5$ ;  $SD=0.53$ ; 50% weiblich) über ein Schulhalbjahr hinweg ein erfolgreich evaluiertes Training zum selbstregulierten Lernen (Stebner et al., 2015), in dem sie eine kognitive Experimentierstrategie (Variablen-Kontroll-Strategie; Klahr & Dunbar, 1988) kombiniert mit metakognitiven Strategien (Ziel setzen, überprüfen, reagieren; Bandura, 1986) erlernten. Darüber hinaus erhielten diese SuS zeitgleich einen auf die Trainingsinhalte abgestimmten Deutschunterricht. Als mögliche Einflussfaktoren auf den Strategietransfer wurden das kognitive und metakognitive Strategiewissen vor und nach dem Training mithilfe des Multistrategietests (MST; Stebner et al., 2015; Cronbachs Alpha = .70), die kognitiven Fähigkeiten (KFT; Heller et al., 1985) und die Schulnoten erhoben. Der metakognitive Strategietransfer wurde mithilfe einer auf den Deutschunterricht adaptierten Version des MST gemessen und stellt somit ein weite Transferaufgabe für die SuS dar. Es wurden Varianzanalysen mit wiederholtem Messzeitpunkt und dem Strategiewissen im Fach Deutsch (weiter Transfer) als abhängige und dem Strategiewissen vor und nach dem Training, den kognitiven Fähigkeiten und den Schulnoten als unabhängige Variable berechnet.

Die Ergebnisse der Analyse mit dem Strategiewissen vor dem Training als unabhängige Variable zeigen einen Haupteffekt des Messzeitpunktes ( $F(1,94)=42.39, p=.00, \eta^2=.31$ ), aber keinen Haupteffekt des Strategiewissens vor dem Training ( $F(1,95)=2.33, p=.13, \eta^2=.02$ ) und keine Interaktion zwischen Messzeitpunkt und Strategiewissen vor dem Training ( $F(1,95)=3.23, p=.08, \eta^2=.03$ ). Für die unabhängige Variable Strategiewissen nach dem Training zeigt sich ein Haupteffekt für den Messzeitpunkt ( $F(1,95)=37.60, p=.00, \eta^2=.28$ ), ein Haupteffekt für das Strategiewissen nach dem Training ( $F(1,95)=6.23, p=.01, \eta^2=.06$ ) und eine Interaktion zwischen Messzeitpunkt und Strategiewissen nach dem Training ( $F(1,95)=5.42, p=.02, \eta^2=.05$ ). Demnach gelingt es den leistungsstarken SuS des Trainings besser, metakognitive Strategien im Fach Deutsch zu transferieren als den leistungsschwachen. Die Variablen kognitive Fähigkeiten und Schulnoten beeinflussen die Transferleistung der SuS nicht signifikant.

Hinsichtlich der Frage nach Bedingungsfaktoren für erfolgreichen weiten Strategietransfer beim SRL lässt sich zusammenfassend festhalten, dass das kognitive und metakognitive Strategiewissen nach dem Training einen signifikanten Einflussfaktor für den weiten Strategietransfer in den Deutschunterricht darstellt. Diejenigen SuS, die also im Training erfolgreich waren, konnten die metakognitiven Strategien auch transferieren. Allerdings konnte weder ein Einfluss des Strategiewissens vor dem Training noch von den kognitiven Fähigkeiten und den Schulnoten auf den metakognitiven Strategietransfer festgestellt werden. Diese Ergebnisse betonen die Wichtigkeit von Trainings zur Förderung des SRL und machen die Komplexität von Transferprozessen beim schulischen Lernen deutlich.

## A05

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Rosshofgasse, S01

### Evidence on the instructional validity of a newly developed video-based tool for measuring teaching competence

**Hannes Saas, Christiane Kuhn, Olga Zlatkin-Troitschanskaia, Christina Parr, Sebastian Brückner**

Johannes Gutenberg Universität Mainz, Deutschland; [hannes.saas@uni-mainz.de](mailto:hannes.saas@uni-mainz.de), [troitschanskaia@uni-mainz.de](mailto:troitschanskaia@uni-mainz.de)

Researchers agree that teaching competence is situation-specific and enables teachers to act in accordance with the objectives of the situation (Seidel & Shavelson, 2007). This competence can be assessed in a valid manner only if the assessment involves an appropriate representation of actual teaching practices (Darling-Hammond et al., 2013). Video recordings allow for proximal assessment of teachers' actions (Alonzo & Kim, 2015).

Although video-based tools to assess teaching competence exist for the fields of mathematics and science, in business and economics (B&E) such approaches are scarce. The few existing instruments have no comprehensive validity evidence as recommended by AERA et al. (2014). Instructional validity is "the extent to which an assessment is aligned with curriculum and instruction [...]" (Pellegrino et al., 2016, p. 63). Based on previous research, we developed a video-based tool to assess the teaching competence of B&E teachers and comprehensively analyzed its validity. In this paper, we focus on the instructional validity and on the following questions:

1. Does the new assessment align with instructional demands in teachers' instructional practice?
2. Do the cognitive processes elicited while solving video-based tasks align with the expected cognitive processes related to the construct definition of teaching competence?

Two fields of teaching competence can be distinguished that represent different specific instructional demands: (1) during class (e.g., immediate actions under time pressure) and (2) before and after class (e.g., planning and evaluation) (Lindmeier, 2011). Thus, our model differentiates between action-related competence (AC), necessary to spontaneously handle specific situations during instruction, and reflective competence (RC), necessary to prepare and evaluate pre- and post-instructional phases.

The model also differentiates between three levels of teacher expertise (student teachers, teacher trainees, in-service teachers), which is assumed to increase over the different stages of teacher education (university education, practical training, and professional practice). To assess teachers' AC and RC, we embedded 18 video clips (1-2 minutes in length) of genuine teacher-student classroom interactions into computer-based tasks, considering the situation-specific instructional context (i.e., type of school class, topic). In AC situations, the test takers were required to react immediately via microphone to a classroom situation; in RC situations, the test takers were required to prepare a follow-up lesson to improve students' learning. These 18 tasks were validated to ensure they corresponded with learning opportunities in the teacher education curriculum (cf., Authors, 2017).

Results: 48 experienced teacher educators from university and practical training rated the tasks on a 4-point scale according to various analysis criteria such as comprehensibility, genuineness, and both the relevance and representativeness of instructional methods in the respective content area. Results confirmed that 17 of the 18 video-based items are highly relevant and genuine representations of teachers' instructional skills.

The analysis of 42 cognitive interviews with pre- and in-service teachers using the think aloud method (Leighton et al., 2011) confirmed the theoretically expected cognitive processes in the sense of the construct definition when responding to the video-based tasks. We transcribed, coded, re-coded twice, and quantified the data (Chi, 1997). Additionally, we administered a socio-biographical background questionnaire (e.g., attended courses at university) and determined that teaching practical experience is the strongest indicator of successful task solving.

Moreover, the results from the expert survey align with the findings from the cognitive interviews. This can be considered as first evidence of instructional validity, as the expert ratings, the cognitive labs, and their reciprocal relations highly correspond in showing the relevance and suitability of the new video-based tasks for teaching practice. Implications for instruction, learning and valid assessment in teacher education and consequences for use of the video-based tool in teacher practice will be discussed.

### Einfluss von Lerngelegenheiten auf das Professionswissen von Lehramtsstudierenden im Fach Mathematik

**Florina Stefanica**

BWT Universität Stuttgart, Deutschland; [stefanica@bwt.uni-stuttgart.de](mailto:stefanica@bwt.uni-stuttgart.de)

Das Professionswissen der Lehrkräfte, das üblicherweise in Fachwissen, fachdidaktisches und pädagogisches Wissen unterteilt wird, gestützt durch deren Überzeugungen und motivationale Orientierungen, ist prädiktiv für situationsbezogene und verhaltensnahe Maße der Lehrkräfte, die konkrete Unterrichtsgestaltung der Lehrkräfte sowie die Lernfortschritte und die Leistungsängstlichkeit der unterrichteten Schülerinnen und Schüler und reduziert die emotionale Erschöpfung und die Burnout-Wahrscheinlichkeit der Lehrkräfte. Dies wurde im Rahmen von Studien wie COACTIV, COACTIV-R, TEDS-FU, LEK, BilWiss, MLMT empirisch belegt.

Die Relevanz des Professionswissens von Lehrkräften wird somit an diversen Stellen deutlich. Doch wovon hängt die Entwicklung dieses Professionswissens ab?

Der vorliegende Beitrag fokussiert Lerngelegenheiten als institutionelle Determinanten für Unterschiede im Professionswissen und dessen Entwicklung. Dabei werden (1) die Lehrerbildung allgemein, (2) die unterschiedlichen Studienstandorte, (3) die von den Studierenden gewählten Zweifächer und (4) ausgewählte Qualitätsmerkmale von Lehrveranstaltungen als Lerngelegenheiten betrachtet. Ein detaillierter Überblick zu Lerngelegenheiten im Rahmen der Professionalisierung von Lehrkräften findet sich bei Ștefănică (2018).

Bezogen auf (1) die generelle Wirksamkeit der Lehrerbildung konnten Wissenszuwächse in den fachbezogenen Dimensionen des Professionswissens verzeichnet werden: Dabei wurden Querschnittuntersuchungen durchgeführt, unter Einbezug der Studieneingangsphase und dem Ende der ersten und zweiten Phase (Krauss et al. 2011; Voss/ Kunter 2017; Blömeke et al. 2008b). Untersuchungen des Studienverlaufs fanden bisher nicht statt. Hingegen wurde für das fachübergreifende Wissen die ganze Spanne der Lehrerbildung in Quer- und Längsschnittuntersuchungen beforscht (Seifert/ Schaper 2012; König 2012; König/ Seifert 2012; König/ Klemenz 2015; Blömeke et al. 2008c; Voss/ Kunter 2017).

(2) Der Studienstandort weist Einflüsse auf die Wissensstände und –entwicklung im pädagogischen Wissen auf (König 2012; Seifert/ Schaper 2012; König et al. 2012; König/ Seifert 2012; Kunina-Habenicht et al. 2013; Kunter et al. 2017). Analoge Studien für das fachbezogene Wissen existieren bisher nicht.

Bezogen auf (3) das Zweifach zeigen Blömeke et al. (2013), dass Studierende mit einer Naturwissenschaft als Zweifach besser im Gesamtkonstrukt Professionswissen abschneiden als Studierende mit einer Sprache oder einer Geisteswissenschaft als Zweifach. Studien zu den fachbezogenen oder zur fachübergreifenden Dimension des Professionswissens stehen noch aus.

(4) Qualitätsmerkmale des Studiums als Lerngelegenheiten zeigen Korrelationen mit dem Fachwissen und dem fachdidaktischen Wissen auf (Stancel-Piatak et al. 2013; Blömeke et al. 2011; Felbrich et al. 2008), jedoch nicht mit dem pädagogischen Wissen (König et al. 2012; König/ Seifert 2012; Stancel-Piatak et al. 2013). Dabei wurden Qualitätsmerkmale für das bisherige Studium insgesamt erfasst.

Im Rahmen der vorliegenden Studie wird untersucht, ob im Studienverlauf Zuwächse in den einzelnen Dimensionen des Professionswissens verzeichnet werden können. Dabei werden Unterschiede nach Standort bzw. Zweifach analysiert. Zusätzlich wird geprüft, ob die wahrgenommenen Qualitätsmerkmale von Fachdidaktik-Lehrveranstaltungen mit dem fachdidaktischen Wissen zum Ende dieser Lehrveranstaltungen zusammen hängen.

Bei Studierenden der Universitäten Stuttgart und Tübingen konnte ein Quasi-Längsschnitt mit drei Messzeitpunkten (N zwischen 125 und 175) sowie ein echter Längsschnitt mit zwei Messzeitpunkten realisiert werden (N = 102). Die TUM School of Education konnte als Vergleichsstandort gewonnen werden. Auf Basis von IRT-Skalierungen der Leistungsdaten werden Zusammenhangs- und Regressionsanalysen mit den Lerngelegenheiten gerechnet.

Erste Analysen zeigen dass, in den fachbezogenen Wissensdimensionen stärkere Zuwächse als im fachübergreifenden Wissen verzeichnet werden. Weiterhin verfügen Studierende an den unterschiedlichen Standorten zu diversen Zeitpunkten im Lehramtsstudium über unterschiedliche Wissensstände im Fachwissen und fachdidaktischen Wissen (Hedges g bis 0,6); im pädagogischen Wissen sind die Unterschiede zwischen den Standorten geringer. Zusätzlich finden sich Unterschiede zwischen den Wissenskomponenten jener Studierenden, die MINT- bzw. Nicht-MINT-Zweifächer gewählt haben: Hierbei liegen die Vorteile in den fachbezogenen Wissenskomponenten bei der MINT-Gruppe, wohingegen die Nicht-MINT-Gruppe im fachübergreifenden Wissen besser abschneidet.

Wahrgenommene Qualitätsaspekte der Lehrveranstaltungen weisen erwartungskonforme Zusammenhänge zum fachdidaktischen Wissen auf. Jedoch findet man bei getrennter Betrachtung nach Standort bzw. Zweifach differenzierte Ergebnisse.

## **Der Einfluss von kontrafaktischen Gedanken auf die affektiven Reaktionen und Handlungsintentionen von angehenden Lehrkräften: Eine Interventionsstudie**

**Daria Katharina Benden, Fani Lauer mann**

Universität Bonn, Deutschland; [dariabenden@uni-bonn.de](mailto:dariabenden@uni-bonn.de)

### THEORIE

Kontrafaktisches Denken bezieht sich auf Gedanken über alternative Ergebnisse und Ereignisse, die anders hätten stattfinden können und/oder sollen (Epstude & Roese, 2008; Markman, Gavanski, Sherman, & McMullen, 1993). Eine zentrale Funktion kontrafaktischen Denkens liegt in der Steuerung zukünftigen Verhaltens. So werden kontrafaktische Gedanken oft durch das Erkennen eines Problems ausgelöst. Der spezifische Gedanke liefert dann eine Information darüber, welches Verhalten notwendig gewesen wäre, um das Problem zu vermeiden (z.B. „Ich hätte mich mehr anstrengen sollen!“). Aus dieser Einsicht können sich wiederum Handlungsintentionen und darauffolgende Handlungen ableiten. Dementsprechend weisen kontrafaktische Gedanken durch ihren Fokus auf zukünftige Handlungen Potential zur Verbesserung von Leistungen auf (Markman, McMullen, & Elizaga, 2008; Roese, 1994). Beispielsweise können experimentell induzierte kontrafaktische Gedanken von Lernenden ihre zukünftigen akademischen Leistungen verbessern (Markman et al., 2008).

Bisher mangelt es jedoch an Studien, in denen die Theorie des kontrafaktischen Denkens auf den Lehrberuf übertragen wurde. So stellt sich die Frage, welchen Einfluss kontrafaktische Gedanken von Lehrkräften auf deren Handlungsintentionen und Reaktionen haben können (z.B. Unterstützung der Schüler/innen). Dabei ist nicht nur die Leistungsbereitschaft der Lehrkraft, sondern auch die der Lernenden relevant, was die Komplexität des Entscheidungsprozesses erhöhen kann.

### FRAGESTELLUNG

Aufbauend auf diese Evidenz untersucht die vorliegende Studie die experimentell induzierten Effekte von kontrafaktischen Gedanken auf die affektiven Reaktionen (Mitgefühl, Ärger, Schuldgefühl) und Handlungsintentionen (z.B. Hilfsbereitschaft gegenüber einem Schüler) von angehenden Lehrkräften. Effekte auf die Selbstwahrnehmungen der angehenden Lehrkräfte (z.B. Eigenverantwortung und Kontrollüberzeugungen) werden ebenfalls untersucht, um mögliche Wirkmechanismen aufzudecken.

### METHODE

Drei Gruppen von angehenden Lehrkräften (n=179) wurden angeleitet, kontrafaktische Gedanken in Bezug auf die schlechte akademische Leistung einer Schülerin/eines Schülers („A.B.“) zu formulieren. Eine Vignette mit der Situation A.B.s wurde aus einer früheren Studie im Kontext der Attributionstheorie adaptiert (Reyna & Weiner, 2001).

Die Teilnehmer/innen wurden randomisiert einer Kontrollgruppe und zwei Experimentalgruppen zugeordnet. In der „AB-Gruppe“ (EG1) beschreiben die Teilnehmenden, was A.B. hätte tun sollen, um eine bessere Leistung zu erzielen. In der „LuL-Gruppe“ (EG2) listeten die Teilnehmenden kontrafaktische Gedanken über sich selbst in der Rolle der Lehrkraft auf. Eine Kontrollgruppe (KG) formulierte nur allgemeine Einstellungen zum Thema der Studie.

Daraufhin wurden Fragen zu den affektiven Reaktionen der Lehrkraft (Ärger über A.B., Mitgefühl mit A.B., Schuldgefühl), zur Bewertung der Verantwortung und Kontrolle bezüglich A.B. und sich selbst in dieser Situation sowie zur Hilfsbereitschaft der Lehrkraft gegenüber A.B. gestellt. Weiterhin wurde in einem offenen Format erfasst, welche selbstentworfenen Strategien die angehenden Lehrkräfte einsetzen würden, um A.B. zu helfen.

### ERGEBNISSE

In Bezug auf die erfassten standardisierten Skalen, zeigten sich signifikante Gruppenunterschiede nur hinsichtlich des Mitgefühls mit A.B. ( $F(2,173)=5.15, p<.01, \omega=.21$ ). Post-Hoc Vergleiche mit der Sidak-Korrektur zeigten, dass die

Teilnehmenden der „LuL-Gruppe“ signifikant mehr Mitgefühl mit A.B. hatten ( $M=3.09$ ,  $SD=1.33$ ) im Vergleich zur „AB-Gruppe“ ( $M=2.46$ ,  $SD=1.25$ ) sowie zur Kontrollgruppe ( $M=2.38$ ,  $SD=1.09$ ). Beobachtete Gruppenunterschiede bezüglich Ärger, Schuldgefühlen, Verantwortungszuschreibungen, Kontrollüberzeugungen und allgemeiner Hilfsbereitschaft der angehenden Lehrkräfte waren in der Tendenz ebenfalls theoriekonform aber nicht signifikant.

Die selbstentworfenen Strategien zur Unterstützung von A.B. wurden mittels qualitativer Inhaltsanalyse ausgewertet (Interrater-Übereinstimmung: 82%). Poisson-Regressionen zeigten, dass Teilnehmende der „LuL-Gruppe“ im Vergleich zur „AB-Gruppe“ signifikant häufiger Strategien entwerfen, die Anpassungen des eigenen Unterrichts erfordern, während Teilnehmende der „AB-Gruppe“ die Problemlösung signifikant häufiger auf andere Personen übertragen (z.B. Eltern oder A.B. selbst).

Diese Ergebnisse zeigen, dass kontrafaktische Gedanken die affektiven Reaktionen angehender Lehrkräfte nur bezüglich des Mitgefühls mit Lernenden signifikant beeinflussen und darüber hinaus – abhängig vom Bezug der Gedanken auf die eigene Person oder die Schülerin/den Schüler – die darauffolgenden Handlungsintentionen angehender Lehrkräfte verändern können. Die divergenten Ergebnisse hinsichtlich der selbstentworfenen Strategien in beiden Experimentalgruppen verdeutlichen dabei die Relevanz kontrafaktischer Gedanken für Entscheidungsprozesse im Kontext des Lehrberufs. Weitere Forschung bezüglich der angenommenen Wirkmechanismen ist nötig.

## **Die evidenzbasierte Entwicklung eines Masterseminars zur Förderung der professionellen Wahrnehmung**

**Sabrina Dämmer, Verena Zucker, Timo Reuter, Miriam Leuchter**

Universität Koblenz Landau, Deutschland; [daemmer@uni-landau.de](mailto:daemmer@uni-landau.de), [reutertimo@uni-landau.de](mailto:reutertimo@uni-landau.de)

Schülerinnen und Schüler besitzen vielfältige Konzepte zu naturwissenschaftlichen Inhalten, von denen die meisten jedoch wissenschaftlichen Ansprüchen nicht standhalten (Harlen, 2001; Vosniadou, Vamvakoussi, & Skopeliti, 2008). Um die Schülerinnen und Schüler bei einer Umstrukturierung zu wissenschaftlich annähernden Konzepten zu unterstützen, müssen Lehrkräfte im Rahmen des Formative Assessment a) bestehende Konzepte der Schülerinnen und Schüler diagnostizieren (Greenfield, 2015; Kretschmann, 2000) und b) Rückmeldungen geben (Cowie & Bell, 1999; van den Boom, Paas, & van Merriënboer, 2007).

Für den naturwissenschaftlichen Unterricht konnte gezeigt werden, dass Formative Assessment im Sachunterricht lernwirksam ist (Decristan et al., 2015), dies den Lehrkräften jedoch oft schwerfällt (Gotwals & Birmingham, 2015). Dies könnte damit zusammenhängen, dass Lehrkräfte Schwierigkeiten haben, für das Formative Assessment geeignete Lehr-Lernsituationen wahrzunehmen.

Professionelle Wahrnehmung wird als Voraussetzung für das professionelle Handeln in Lehr-Lern-Situationen angesehen (Seidel & Stürmer, 2014). Sie umfasst das Erkennen lernbedeutsamer Situationen (Noticing) und das Interpretieren dieser Situationen (Reasoning) als zwei dynamische, sich einander ergänzende, wissensbasierte Prozesse (Schäfer & Seidel, 2015; Sherin & van Es, 2009).

Erste empirische Befunde aus dem naturwissenschaftlichen Grundschulunterricht verdeutlichen, dass die professionelle Wahrnehmung keine allgemeine, sondern eine domänen- und themenspezifische Fähigkeit ist (Steffensky, Gold, Holodynski, & Möller, 2015; Sunder, Todorova, Steffensky, & Möller, 2014). Studien zeigen auch, dass die professionelle Wahrnehmung in universitären Seminaren gefördert werden kann (Meschede, Steffensky, Wolters, & Möller, 2015; Sunder, Todorova, & Möller, 2017).

Im Rahmen des Projekts Modulare Schulpraxiseinbindung als Ausgangspunkt zur individuellen Kompetenzentwicklung (MoSAiK) der Qualitätsoffensive Lehrerbildung an der Universität Koblenz-Landau am Campus Landau wurde ein Seminar entwickelt, in dem a) Ausschnitte komplexer Lehr-Lern-Situationen fremder Lehrkräfte (Video- und Textvignetten) von den Studierenden analysiert werden (Gartmeier, 2014; Kleinknecht & Syring, 2014; Krammer et al., 2016) und b) komplexitätsreduzierte Lehr-Lernsituationen (Biaggi, Krammer, & Hugener, 2013) von den Studierenden durchgeführt, videographiert und analysiert werden.

Forschungsfrage

Kann das Erkennen von diagnostischen Maßnahmen und Rückmeldung bei Grundschullehrantsstudierenden durch das Seminar konzept gefördert werden?

Methode

Das qualifizierte Erkennen von Diagnostik und Rückmeldung (insgesamt 18 Items) wurde mithilfe von je zwei text- und videobasierten Ausschnitten realer Unterrichtsszenen (Vignetten) erfasst (Zucker & Leuchter, eingereicht). Unmittelbar nach der mehrfachen Sichtung der jeweiligen Vignette schätzten die Teilnehmenden Items auf einer 5-stufigen Likertskala („1=trifft zu“ bis „4=trifft nicht zu“ sowie „5=weiß nicht“) ein, die bestimmte Maßnahmen der Diagnostik und der Rückmeldung vorgaben. Diese Maßnahmen waren entweder in der Vignette zu erkennen (Zustimmungs-Items) oder nicht enthalten (Ablehnungs-Items).

Die Studie erfolgte im Ein-Gruppen-Prä-Post-Design mit  $N=82$  Grundschullehrantsstudierenden im Master der Universität Koblenz-Landau. Für die Analyse wurde die Skala dichotomisiert und Mittelwerte gebildet, anschließend eine messwiederholte Varianzanalyse mit den Faktoren MZP (Prä-, Posttest), Vignettentyp (Text, Video), Art der Maßnahme (Diagnose, Rückmeldung) sowie Itemtyp (Zustimmung, Ablehnung) mit der Lösungsrate als abhängiger Variable gerechnet.

Ergebnisse

Eine messwiederholte Varianzanalyse ergab einen signifikanten Haupteffekt des Messzeitpunkts. Über den ganzen Test hinweg war, die Lösungsrate im Posttest jedoch wider Erwarten geringer als im Prätest,  $M_{prä} = 0.63$ ,  $SD_{prä} = 0.07$ ,  $M_{post} = 0.59$ ,  $SD_{post} = 0.09$ ,  $F(1,81) = 6.28$ ,  $p = .014$ ,  $\eta^2 = .072$ . Das traf jedoch in erster Linie auf Items der Diagnostik zu, die abzulehnen waren (vgl. Tabelle 1).

Tabelle 1

Mittelwerte, Standardabweichungen und Ergebnisse der -Tests von Prä- und Posttest

Die Lösungsrate der Teilnehmenden im Prä-Posttest-Vergleich

(1) blieb stabil, wenn nach Diagnosetätigkeiten gefragt wurde, die in den Text- oder Videovignetten vorkommen,

(2) verschlechterte sich jedoch, wenn nach Diagnosetätigkeiten gefragt wurde, die nicht in den Text- oder Videovignetten zu sehen sind

(3) verschlechterte sich ebenfalls, wenn nach Rückmeldetätigkeiten gefragt wurde, die nicht in den Videovignetten zu sehen sind. In den Textvignetten blieb sie stabil.

(4) verbesserte sich, wenn nach Rückmeldetätigkeiten gefragt wurde, die in den Text- oder Videovignetten vorkommen.

Diese und weitere Ergebnisse werden in dem Beitrag berichtet und diskutiert.

## A06

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus S212

### Studienunabhängige Selbstprofessionalisierung in der Schule

**Kerstin Bäuerlein, Christian Reintjes**

Pädagogische Hochschule FHNW, Schweiz; [kerstin.baeuerlein@fhnw.ch](mailto:kerstin.baeuerlein@fhnw.ch)

#### Theoretischer Hintergrund

Im Studienbereich der Berufspraktischen Studien an der PH Nordwestschweiz ist konzeptionell eine dreiphasige Einführung der Studierenden in die Lehrpersonenrolle vorgesehen (PH FHNW, 2015): In einer Basisphase werden grundlegende Konzepte und Kompetenzen vermittelt, worauf eine immersive Praxisphase folgt. Eine Phase individueller Fokussierung zielt abschließend auf die im Rahmen einer einphasigen Lehrpersonenbildung intendierte elementare Berufsfähigkeit.

Die zweite Phase der partizipativen Immersion bildet das Herzstück der Konzeption. Basierend auf theoretischen Überlegungen und Erfahrungen ähnlich ausgerichteter Konzepte (z.B. Professional Development Schools in den USA; Darling-Hammond, 2005) wurde für diese Phase ein Partnerschulkonzept entwickelt. Dabei erleben Studierende durch kontinuierliche Anwesenheit und Einbindung in eine Partnerschule über ein Jahr die Vielfalt professionellen Handelns (Fraefel et al., 2016). Sie arbeiten mit qualifizierten Praxislehrpersonen in professionellen Arbeits- und Lerngemeinschaften (vgl. Wenger, 1998) und werden von Hochschuldozierenden in reflexionsfördernden Begleitformaten bei der Verarbeitung der Praxiserfahrungen unterstützt (PH FHNW, 2015).

Diesem theoretisch fundierten Ausbildungskonzept steht die Beobachtung gegenüber, dass ein großer Teil der Studierenden im Bildungsraum Nordwestschweiz neben dem Studium bereits eine studienunabhängige Lehrtätigkeit im Schulfeld ausübt. Daher stellt sich die Frage, ob die sukzessive Einführung der Studierenden in die Lehrpersonenrolle der Realität entspricht oder ob dieses Konzept durch eine Selbstprofessionalisierung in der Schule konterkariert wird. Die Studie hatte zum Ziel, eine Bestandsaufnahme hinsichtlich der studienunabhängigen Lehrtätigkeit der Studierenden vorzunehmen, um daraus im Herbstsemester 2017/2018 Implikationen für die Konzeption der Berufspraktischen Studien abzuleiten.

#### Fragestellungen

- Wie viele Studierende arbeiten während ihres Studiums im Schulfeld? In welchem Umfang? Unter welchen Bedingungen? In welchen Fächern?
- Unterscheiden sich Studierende mit studienunabhängiger Lehrtätigkeit von Studierenden ohne studienunabhängige Lehrtätigkeit?

#### Methode

Alle 858 Studierenden des Instituts für Sekundarstufe der PH Nordwestschweiz wurden im Frühlingsemester 2017 in einem freiwilligen Online-Survey u.a. zu Folgendem befragt:

- Studium: u.a. Studiensemester, Fächerkombination, Studiengang/Zielstufe
- allfällige schulische Lehrtätigkeit: u.a. Umfang, Unterrichtsfächer, Schulart, Betreuung (Erleben des Mentoring, Richter et al., 2011)
- Personenmerkmale: u.a. Alter, Geschlecht, FIT-Choice (Richardson & Watt, 2006), LPA (Brandstätter & Mayr, 1994), LIS (Mayr, 1998)

Auf eine deskriptive Datenanalyse folgten t-Tests für unabhängige Stichproben zum Vergleich der Studierenden mit und ohne studienunabhängige Lehrtätigkeit (explorativ ohne Alpha-Adjustierung). Werte auf unterschiedlichen Skalen wurden mit t-Tests für abhängige Stichproben bzw. Varianzanalysen mit Messwiederholung (ggf. anschließende paarweise Vergleiche mit Bonferroni-Korrektur) verglichen.

#### Ausgewählte Ergebnisse

- Von den 203 antwortenden Studierenden (Rücklauf: 24%) gaben 53% an, neben dem Studium (un-)befristet als Lehrperson zu arbeiten (49% ausschließlich in Studienfächern, 66% ausschließlich auf studiengangskonformen Schulstufen); diese Studierenden befinden sich in höheren Studiensemestern als Studierende ohne Lehrtätigkeit,  $t(204)=2.61$ ,  $p=.010$ .
- 52% der als Lehrperson Tätigen gab an, vor Stellenantritt das Partnerschuljahr abgeschlossen zu haben.
- Die PH-unabhängige Lehrtätigkeit hat für die Studierenden am stärksten entwicklungsfördernde Wirkung, gefolgt von informierender, erkundungsanregender, reflexionsanregender und steuernder Wirkung (alle Unterschiede signifikant,  $p\leq.013$ ).
- Knapp 18% gaben an, bei ihrer PH-unabhängigen Lehrtätigkeit betreut zu werden. Aussagen zur emotionalen wie auch zur instruktionalen Unterstützung werden von 94% bzw. 89% als (eher) zutreffend bewertet. Die Zusammenarbeit ist den Angaben der Studierenden zufolge tendenziell stärker von konstruktivistischen als von transmissiven Interaktionen geprägt,  $t(17)=1.88$ ,  $p=.077$ , wobei 83% bzw. 67% entsprechende Aussagen als (eher) zutreffend bewerten.
- Für Studierende mit Lehrtätigkeit spielen die selbstwahrgenommene Lehrbefähigung, der Wunsch, einen sozialen Beitrag für die Gesellschaft zu leisten, sowie der Wunsch, mit Kindern und Jugendlichen zu arbeiten, eine größere Rolle für die Berufswahl als für Studierende ohne Lehrtätigkeit,  $t(201)=4.12$ ,  $p<.001$ ;  $t(201)=2.31$ ,  $p=.022$ ;  $t(201)=2.24$ ,  $p=.027$ .
- Studierende mit Lehrtätigkeit haben zudem ein größeres Interesse an der Förderung sozialer Beziehungen,  $t(201)=2.46$ ,  $p=.015$ .
- Ansonsten unterscheiden sich die Gruppen nicht hinsichtlich ihrer Berufswahlmotive oder Lehrer-Interessen, und auch nicht hinsichtlich der erhobenen Persönlichkeitsmerkmale.

Die Ergebnisse werden in Hinblick auf das Professionalisierungskonzept der Hochschule diskutiert. Zudem werden Überlegungen zu Implikationen für die Aus-, Fort- und Weiterbildung von Lehrpersonen getätigt.

### Förderung von Klassenführungscompetenz in der ersten Phase der Lehrkräfteausbildung: Ein Vergleich zweier Ansätze didaktischen Arbeitens mit Videofällen

**Tom Schep, Udo Rauin**

Goethe Universität Frankfurt am Main, Deutschland; [schep@em.uni-frankfurt.de](mailto:schep@em.uni-frankfurt.de)



## Theoretischer Hintergrund:

Für den Erwerb von Klassenführungs Kompetenzen in der ersten Phase der Lehrkräfteausbildung wird der Teilnahme an universitär begleiteten Praxisphasen sowie der Analyse von videobasierten Praxissimulationen eine bedeutsame Rolle zugeschrieben (u.a. Larrivee 2006, Jones 2006). Empirische Studien haben den Effekt der Teilnahme an Schulpraktika auf die professionelle Wahrnehmung (Stürmer et. al. 2013) und das Wissen zu Klassenführungsthemen (Blüthmann et. al. 2010) nachgewiesen. Auch zeigen verschiedene Studien die Effektivität von videogestützten Seminaren bzgl. professioneller Wahrnehmung von Klassenführung (u.a. Gold et. al. 2013; Barth 2016; Kramer et. al. 2017). Allerdings fehlt es bislang an Studien, die systematisch untersuchen, wie mit der Gestaltung von videobasierten Lerngelegenheiten effektiv auf die praktischen Erfahrungen von Lehramtsstudierenden Bezug genommen werden kann. Verschiedene empirische Studien (u.a. Syring et. al. 2015) zeigen eine hohe kognitive Belastung von Lehramtsstudierenden bei der Beobachtungen von Videovignetten. Diese erschwere es, relevante Unterrichtsmerkmale zu identifizieren und zu analysieren (Wolff et. al. 2016). Diese Erkenntnisse lassen vermuten, dass Studierende ohne Praxiserfahrung eher von einem theoriegeleiteten, stark strukturierten Ansatz profitieren, der darauf abzielt, kognitive Überlastung zu vermeiden. Dahingegen lassen Studien zum situierten Lernen (Korthagen 2010; Gruber 1996) vermuten, dass Studierende mit Erfahrungen im Praxissemester eher von einem explorativen Ansatz profitieren, der darauf abzielt, eigene Praxiserfahrungen in Videos zu reflektieren.

## Fragestellung:

In diesem auf einer Dissertationsstudie basierenden Beitrag wird der Frage nachgegangen, inwieweit Praxiserfahrungen von Studierenden einerseits und die didaktische Arbeitsweise in videobasierten Seminaren andererseits differentielle Effekte auf die Fähigkeit haben, klassenführungsrelevante Ereignisse in Videos zu erkennen und zu interpretieren.

## Methode:

Um diese Frage zu beantworten, wurden in einem quasi-experimentellen Pre-Post-Forschungsdesign eine Gruppe von Lehramtsstudierenden, die am Praxissemester teilgenommen haben (n=42), und eine Gruppe von Studierenden ohne Praxiserfahrung (n=43) jeweils randomisiert zwei videobasierten Blockseminaren zum Thema Klassenführung zugewiesen. In Anlehnung an eine Studie von Seidel et. al. (2013) wurde in einem der Seminare ein exploratives Vorgehen realisiert, bei dem die Studierenden ohne vorherigen Theorieinput aufgefordert wurden, unter Bezugnahme ihrer eigenen Praxiserfahrungen aus Videovignetten Strategien für effektive Klassenführung abzuleiten (example-rule). Dagegen wurden Studierende im anderen Seminar mit theoriegeleitetem Ansatz aufgefordert, zuvor vermittelte theoretische Konzepte im Rahmen stark strukturierter Aufgabenstellungen auf Videovignetten anzuwenden (rule-example). In beiden Seminaren wurden jeweils die gleichen Videos mit Hilfe von Tablets über eine Online-Lernplattform vor Ort zur Verfügung gestellt. Für die Erfassung der Fähigkeit, klassenführungsrelevante Ereignisse zu erkennen und zu interpretieren, wurden zwei Instrumente eingesetzt. Zum einen wurde der Videotest „Classroom Management Expertise“ (CME, König & Lebens 2012) eingesetzt. Zum anderen wurde ein neuer Videotest verwendet, der in Anlehnung an CME entwickelt und mit neuen Videovignetten und Fragen um theoretische Modelle zu Klassenführung nach Borich (2011) erweitert wurde. Darüber hinaus wurden ein Aufmerksamkeitstest (FAIR-2, Moosbrugger & Oehlschlägel, 2011) sowie Skalen zur Selbstwirksamkeit und Seminarbeurteilung in die Analyse mit einbezogen.

## Ergebnisse:

Erste Auswertungen anhand von Mittelwertvergleichen des selbstentwickelten Videotests zeigen einen großen signifikanten Haupteffekt auf Testebene für alle Seminargruppen zwischen Pretest und Posttest ( $F(1, 81) = 49.818$ ;  $p = .000$ ;  $\eta^2 = .381$ ). Allerdings konnten wider Erwarten keine signifikanten Interaktionseffekte erster und zweiter Ordnung zwischen Testscore, Praxiserfahrung und Seminartyp festgestellt werden. Im Vortrag werden diese Ergebnisse sowie zusätzlich die Ergebnisse der Varianzanalysen auf der Ebene der Subskalen präsentiert. Anhand der vorgestellten Ergebnisse werden Empfehlungen für die Gestaltung von Lerngelegenheiten mit Videovignetten sowie Perspektiven für zukünftige Forschung diskutiert.

## **Reflexion im Praxissemester Physik**

**Anna Nowak, Andreas Borowski**

Universität Potsdam, Deutschland; [anna.nowak@uni-potsdam.de](mailto:anna.nowak@uni-potsdam.de)

## Theoretischer Hintergrund

Erweiterte Praxisphasen sind ein wichtiger Bestandteil der deutschen LehrerInnenbildung. Die Erwartung an den Lernzuwachs innerhalb dieser Phasen ist sehr hoch (Rothland & Boecker, 2015), wie hoch er allerdings tatsächlich ist, ist eine langwährende Frage (Gröschner, Schmitt & Seidel, 2013). Helmke (2015) hat Reflexionskompetenz als Schlüsselbedingung für die Verbesserung des eigenen Unterrichts bezeichnet. Ergebnisse vorheriger qualitativer (Interventions-)studien haben gezeigt, dass Studierende kaum strukturiert und kritisch reflektieren, sondern eher beschreibend (z.B. Hatton & Smith, 1995). Basierend auf den Modellen von Korthagen & Kessels (1999) und Windt & Lenke (2016) wurde ein eigenes Reflexionsmodell entwickelt. Aus beiden Modellen wurde die Achse „Elemente einer Reflexion“ leicht verändert übernommen, hinzugefügt wurden die Achsen „Wissensbasis“ und „Begründung“. Eine Reflexion hat in diesem Modell fünf Elemente. Das erste Element, die Rahmenbedingungen, umfasst den theoretischen Plan der Stunde, also u.a. eine Darstellung der Lernziele oder des Schülervorwissens. Das nächste Element ist die objektive Beschreibung der tatsächlichen Unterrichtssituation gefolgt von der Bewertung der selbigen. Sollte diese negativ sein, so muss eine alternative Handlung entwickelt werden, andernfalls kann dieselbe Handlung erneut stattfinden (Element Alternativen). In Kontrast zum Element Alternativen sind die Konsequenzen langfristige Ziele für die Lehrperson, welche auch losgelöst von der reflektieren Unterrichtsstunde gefasst werden können, wie bspw. das Ziel, das grundätzlich mit Elementarisierung zu beschäftigen. Jedes der Elemente kann ohne oder mit einer Begründung dargestellt werden. Die Themen innerhalb der einzelnen Elemente können entweder aus dem Bereich des Fachwissens (FW), Fachdidaktischen Wissens (FDW), Pädagogischen Wissens (PW) oder Sonstigem (Sonst.) stammen.

## Fragestellung und Methode

Untersucht wird die Validität des Modells, indem die Strukturiertheit der Reflexionen (sinnlogische Reihenfolge der einzelnen Elemente entsprechend dem Modell) und die Vollständigkeit (alle Elemente müssen enthalten sein) über Qualitative Inhaltsanalyse nach Mayring (2010) mittels eines deduktiven Kategoriensystems analysiert wird. Weiterhin wird mittels einer Qualitativen Inhaltsanalyse (induktiv-deduktives Kategoriensystem) untersucht, welche Inhalte aus welcher Wissensbasis die Studierenden thematisieren und wie sich der Fokus im Laufe des Praxissemesters verändert.

Die hier vorgestellte Studie findet im Rahmen des Projekts PSI Potsdam (Professionalisierung – Schulpraktische Studien – Inklusion) statt. Datengrundlage sind schriftliche Reflexionen von Masterstudierenden, die das 4-monatige Praxissemester

absolvieren. Ein Arbeitsauftrag der Studierenden ist es, sechs schriftliche Reflexionen, jeweils zwei zu den Themen Experiment, Aufgabe und Unterrichtseinstieg, zu schreiben. Die Daten stammen von sieben Praxissemesterstudierenden aus dem Sommersemester 2017.

#### Ergebnisse

Von den 42 Reflexionen sind 69% sehr gut strukturiert, d.h. dass ein Element sinnlogisch auf das andere folgt, entsprechend des Modells. Bei einer unstrukturierten Reflexion sind alle Elemente durcheinander gemischt. Vier Studierende schrieben durchgängig strukturierte Reflexionen. Zwei zeigten gemischte Ergebnisse und ein Studierender schrieb stets unstrukturiert. Die Hälfte der Reflexionen sind vollständig, umfassen also alle fünf Elemente. Die andere Hälfte ist fast vollständig, hier fehlt jeweils die Konsequenz. Drei Studierende schrieben vermehrt vollständige Reflexionen, der Rest vermehrt fast vollständige. Die Inter-Rater-Reliabilität für die einzelnen Elemente des Modells ist jeweils gut (Cohen's Kappa je Element: Rahmenbedingungen 0.90, Beschreibung 0.865, Bewertung 0.73, Alternativen 0.863, Konsequenzen 0.96). Die Ergebnisse sind ein erster Hinweis darauf, dass es für Studierende möglich ist, strukturiertes und kritisches Reflektieren zu erlernen und dass das Praxissemester eine gute Lerngelegenheit ist.

Drei der Studierenden haben über das gesamte Praxissemester hinweg fast ausschließlich pädagogische Themen, wie bspw. Unterrichtsstörungen oder Mitarbeit der Schülerinnen und Schüler, reflektiert. Zwei Studierende haben anfangs hauptsächlich pädagogische Themen reflektiert, im Laufe des Praxissemesters verschob sich der Fokus dann jedoch zu fachdidaktischen Themen, wie bspw. Gelingensbedingungen von Experimenten oder Schülervorstellungen. Zwei weitere Studierende reflektierten von Anfang an hauptsächlich fachdidaktische Themen. Dies gibt uns erste Hinweise, dass der Fokus im Praxissemester möglicherweise verschoben werden sollte: zu Beginn des Praxissemesters könnten mehr erziehungswissenschaftliche Begleitseminare hilfreich sein, gefolgt von mehr fachdidaktischen Seminaren zum Ende hin.

### **Podcasting zur Förderung reflexiver Kompetenzentwicklung zum Thema Klassenführung – eine quasi-experimentelle Studie in der Lehrerbildung**

**Annette Busse, Dorit Bosse**

Universität Kassel, Deutschland; [abusse@uni-kassel.de](mailto:abusse@uni-kassel.de), [bosse@uni-kassel.de](mailto:bosse@uni-kassel.de)

Die Bedeutung professioneller Klassenführung für die erfolgreiche Tätigkeit von Lehrkräften ist seit langem nachgewiesen (Stough, Montague, Landmark & Williams-Diehm, 2015; Freeman, Simonsen, Briere & MacSuga-Gage, 2014; Evertson & Weinstein, 2006; Kounin, 1970). Zugleich wird für die Ausbildung angehender Lehrkräfte die unzureichende Passung zwischen den tatsächlich benötigten Handlungskompetenzen im Schulalltag und der Art universitärer Lerngelegenheiten beklagt, in denen sich bei Studierenden das erforderliche Wissen zum Thema Klassenführung reflexiv entwickeln kann (Freeman, Simonsen, Briere & MacSuga-Gage, 2014;). Dieses zu erwerbende Wissen basiert auf konzeptuellem pädagogischen, fachdidaktischem und psychologischem Wissen (Evertson & Weinstein, 2006) sowie erfahrungsbasierten Überzeugungen. Um den Erwerb von konzeptuellem Wissen zum Thema Klassenführung reflexiv aufbauen zu können, ist ein hochschuldidaktisches Setting erforderlich, das die entsprechende Breite der Wissens- und Erfahrungsbasis berücksichtigt und den Prozesscharakter des Lernvorgangs festhält.

Bestandteil eines solchen Lernsettings kann das Erstellen von Podcasts sein, die die Möglichkeit bieten, den eigenständigen Erarbeitungsprozess eines neuen Themas zu dokumentieren (Taylor & Clark, 2010; Armstrong, 2009; McGarr, 2009; Lazzari, 2008; Siemens, 2004). Durch Podcasting zu Beginn und am Ende eines solchen Prozesses können die Lernenden ihre unterschiedlichen Wissensstände festhalten und reflexiv einordnen. Im vorliegenden Projekt produzieren die Lehramtsstudierenden partnerweise zwei Podcasts, um die Gelegenheit zu erhalten, sich ko-konstruktiv mit dem Thema Klassenführung und ihrem Erarbeitungsprozess auseinandersetzen zu können. Die ko-konstruktive Kommunikation im Medium des Mündlichen erleichtert es Studierenden zudem, sich spontan – ohne die kontrollierte Zurückhaltung, die sich im Schriftsprachlichen unweigerlich ergibt – über ein unbekanntes Thema auszutauschen und gegenseitig gedanklich anzuregen.

Durch Podcasting soll die reflexive Kompetenzentwicklung (Gillen & Dehnbostel, 2007) im Bereich Klassenführung im Lehramtsstudium unterstützt und in einer quasi-experimentellen Studie im Prä-Post-Kontrollgruppendesign untersucht werden. Studierende des ersten Studienjahres (N = 180) produzieren während eines Semesters im Rahmen einer Einführungsvorlesung der Schulpädagogik mit Tutorien in Zweiertteams zwei Podcasts zur Frage, was sie unter "Klassenführung" verstehen: Der erste Podcast wird als eine unvorbereitete Spontanaufnahme zu Studienbeginn vor dem Hintergrund erfahrungsbasierter Überzeugungen erstellt, die sich auf die eigene Schulzeit und ersten Schulkontakte während des Studienbeginns beziehen. Der zweite Podcast wird gegen Ende des Semesters geplant und vorbereitet auf der Basis des erworbenen konzeptuellen Wissens zu Klassenführung produziert. Zwischen beiden Podcasts werden in den Veranstaltungen zentrale Bereiche der Schulpädagogik thematisiert und erarbeitet (Experimentalgruppe). Die Studierenden der Kontrollgruppe besuchen ebenfalls eine Einführungsvorlesung der Schulpädagogik mit Tutorien, erstellen aber keine Podcasts (N = 120).

In der laufenden Studie soll untersucht werden, inwieweit Podcasting als hochschuldidaktisches Element geeignet ist, um die professionelle Wahrnehmung von Klassenführung auf Grundlage von erfahrungsbasierten Überzeugungen und konzeptuellem Wissen von Studienanfängern reflexiv aufzubauen. Die inhaltliche Dimensionalität der Podcasts wird mittels evaluativer Inhaltsanalyse (Kuckartz, 2014) von drei Ratern unabhängig voneinander geratet, die Interraterreliabilität ( $\alpha$ ) wird derzeit an Daten aus der abgeschlossenen Pilotierung (n = 68) etabliert. Darüber hinaus wird die Qualität der professionellen Wahrnehmung von Klassenführung über einen standardisierten videobasierten Online-Test erfasst (Gold & Holodinsky, 2017). Die Frage, welchen Einfluss erfahrungs- und wissensbasiert erstellte Podcasts auf die professionelle Wahrnehmung von Klassenführung im Videotest und auf die Selbsteinschätzung der Klassenführungskompetenz (Tschannen-Moran, Woolfolk Hoy & Hoy, 1998) haben, wird mittels einer Varianzanalyse mit Messwiederholung untersucht.

In der Auswertung der Podcasts aus der Pilotierung im Prä-Post-Design (n = 68) zeigt sich, dass Studierende zu t1 weniger explizit auf Erfahrungen aus ihrer eigenen Schulzeit zurückgreifen als angenommen und diese zu t2 erwartungskonform kaum noch einbeziehen, eine stärker prospektive Perspektive in Bezug auf erfahrungsbezogene Überzeugungen annehmen und das konzeptuelle Wissen steigt.

Das laufende Projekt wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert.

## A07

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Bernoullianum 32, H120

### Förderung diagnostischer Methodenkompetenz – Interventionseffekte eines Lehrveranstaltungsmoduls

**Tobias Alexander Joos, Ulrike Spörhase**

Pädagogische Hochschule Freiburg, Deutschland; [tobias.joos@ph-freiburg.de](mailto:tobias.joos@ph-freiburg.de)

#### Theoretischer Hintergrund

Diagnosekompetenz gilt als wichtige Voraussetzung für die erfolgreiche Planung, Durchführung und Reflexion schulischen Lernens (z.B. HELMKE, 2015) und erfordert fachliches wie fachdidaktisches Wissen und Können. Einige Autoren fokussieren in diesem Zusammenhang verstärkt auf Wissen über diagnostische Methoden und die Fähigkeit ihrer sachgerechten Anwendung (z.B. SCHRADER, 2014; VON AUFSCHNAITER et al., 2015). Diese als methodisches Wissen zusammengefasste Facette der diagnostischen Kompetenz ist eng mit fachdidaktischen Aspekten verbunden, wenn es um die Auswahl, Bewertung und Konzeption geeigneter Aufgaben geht, die im Unterricht zur Diagnose fachlicher Lernprozesse eingesetzt werden können (HELMKE et al., 2004). Eine so verstandene diagnostische Methodenkompetenz fehlt nahezu in der universitären Lehrerbildung (TRITTEL et al., 2014).

#### Fragestellung

Im Kontext universitärer Lehrerbildung werden Instruktionsstrategien empfohlen, die zusätzlich zu einem fundierten, theoretischen Wissen über Diagnostik auch das Können schulen (TRITTEL et al., 2014). Eine von uns entwickelte, theorie- und evidenzbasierte Lehrveranstaltung zur Förderung diagnostischer Methodenkompetenz wurde als Modul für Gymnasiallehramtsstudierende des Faches Biologie in die Ausbildung integriert. Unsere Fragestellung „Kann diagnostische Methodenkompetenz mithilfe des Moduls gefördert werden?“ bezieht sich auf die Wirksamkeit dieses Veranstaltungsteils.

#### Methode

Interventionskonzeption. Mithilfe einer systematischen Übersicht wurden aus-bildungsrelevante Methoden und Inhalte identifiziert, auf deren Basis eine Intervention zur Förderung der diagnostischen Kompetenz entwickelt wurde (Tabelle 1). Als Modul wurde die Intervention im Wintersemester 2016/17 an der Universität Freiburg durchgeführt.

##### 1. Einheit:

- (1) Einführung: Ziele und Notwendigkeit pädagogischer Diagnostik (90 min)
- (2) Qualitätskriterien: Systematischer Ablauf und Beurteilungsfehler (90 min)

##### 2. Einheit:

- (3) Aufgabenbeurteilung (90 min)
- (4) Aufgabenkonzeption (90 min)

##### 3. Einheit:

- (5) Analyse: Aufgabebearbeitungen für Prozessdiagnosen nutzen (90 min)
- (6) Interpretation: Fördermaßnahmen ableiten (90 min)

Tabelle 1: „Diagnose und individuelle Förderung“ im Fach Biologie (insg. 540 min)

Nach der Vermittlung grundlegender Aspekte von Diagnose und Förderung wird diagnostische Methodenkompetenz über die fachdidaktische Auseinandersetzung mit Lernaufgaben geschult. Die Teilnehmer nutzen hierbei die Inhalte einer begleitenden fachwissenschaftlichen Veranstaltung, um Lernaufgaben aus fach-didaktischer Perspektive zu analysieren sowie eigene Diagnoseaufgaben als Erhebungsinstrumente zu erstellen. Ein Schwerpunkt der Veranstaltung liegt somit auf diagnostischem Können.

#### Abbildung 1: Studiendesign

Datenaufnahme. Das Modul wurde durch problemzentrierte Interviews begleitet und mithilfe eines Fragebogens im Pre- / Posttestdesign evaluiert (Abbildung 1).

Eingesetzte Instrumente. Summativ wurden die KLiP-Skalen (GRÖSCHNER, 2008) „Unterrichten“ und „Diagnostizieren“ zur Selbsteinschätzung sowie ein Wissenstest zur pädagogischen Diagnostik eingesetzt (Eigenentwicklung: 14 Items, geschlossene (MC-Aufgaben) sowie halboffene (Zuordnungs-/Ergänzungs-) und offene Aufgaben). Sechs Items erfassen die Einstellungen zu pädagogischer Diagnostik. Im Posttest wurden zusätzlich Fragebögen zur Evaluation des Moduls eingesetzt (COHEP: FARIN et al., 2013) und heiQ: OSBORNE et al., 2007). In den problemzentrierten Gruppeninterviews wurde die Veranstaltung mit Fokusgruppen formativ evaluiert sowie der Umgang mit Aufgaben als Diagnoseinstrumente untersucht.

#### Ergebnisse

Ein Teil der Wirksamkeitstestung bezog sich auf Selbsteinschätzungen der Teilnehmenden im Bereich „Unterrichten“ und „Diagnostizieren“. Ein Vergleich der Mittelwerte aus Pre- und Posttestmessung zeigt bei den zusammengefassten Skalen eine Zunahme am Ende der Intervention (Tabelle 2). Beim „Diagnostizieren“ ist ein signifikanter Zugewinn zu verzeichnen, während die Zunahme beim „Unterrichten“ statistisch nicht signifikant ist.

Auf Itemebene ist dies auf die Einschätzungen zu Kompetenzen der Unterrichtsplanung und Begründung schulischer Inhalte zurückzuführen, wohingegen die Items, welche sich auf Aufgabenkonstruktion und kritische Bewertung des eigenen Vorgehens im Unterricht beziehen, ebenfalls hohe Zugewinne aufweisen.

#### Interventionsgruppe

n M SD

Diagnostizieren\*

Pretest 33 2.78 1.13

Posttest 33 3.43 .81

Unterrichten

Pretest 33 3.71 .90

Posttest 33 3.86 .71

\*p < .05

Tabelle 2: Zunahme des Kompetenzerwerbs nach Selbsteinschätzung der Teilnehmer (KLIP-Skalen)

Die Ergebnisse aus den Fragebögen (Wissenstest, Einstellungen) sowie aus den qualitativen Befragungen weisen darauf hin, dass die Diagnosekompetenz durch die Intervention nachhaltig gefördert wird. Abschließend werden die Ergebnisse diskutiert.

## **Lernzielintentionen, Lernzielvorstellungen und deren Einfluss auf den Erfolg im Chemiepraktikum**

**Thomas Elert, Maik Walpuski**

Universität Duisburg-Essen, Deutschland; [thomas.elert@uni-due.de](mailto:thomas.elert@uni-due.de)

Immer häufiger wird das Laborpraktikum als fester Bestandteil eines Chemiestudiums (Hofstein & Lunetta, 2004; Reid & Shah, 2007) kritisch hinterfragt. Dabei wird in der Literatur besonders oft bemängelt, dass es keine Einigkeit über die Intentionen und Ziele von Laborpraktika gibt (Dillon, 2008; Reid & Shah, 2007) sodass Forderungen für eine Neuerfindung des chemischen Laborpraktikums in Hochschulen laut werden (Eilks, Bäumer & Byers, 2010). Die vorliegende Studie möchte dazu eine erste Bestandaufnahme machen, nicht zuletzt, weil es kaum Erkenntnisse darüber gibt, wann Studierende erfolgreich im Laborpraktikum sind und weil vieles über die Prozesse in und Nützlichkeit von Laborpraktika unbekannt ist. Am Beispiel eines Chemiepraktikums für Erstsemesterstudierende mit Lehramtsoption werden dabei sowohl die Perspektive der Lehrenden als auch die der Lernenden in den Blick genommen.

Der Untersuchung liegen zwei Modelle zur Effektivität von Laborpraktika zu Grunde (Abrahams & Millar, 2008; Bussey, Orgill & Crippen, 2013). Abrahams und Millar (2008) nehmen an, dass die Lehrenden ein Laborpraktikum stets mit bestimmten Intentionen gestalten, welche in Form von Lernzielen ausgedrückt werden. Auffällig hierbei allerdings ist, dass die Lehrenden zwar in der Lage sind diese Lernziele zu benennen (Bretz, Fay, Bruck & Towns, 2013), die Lernziele selbst allerdings selten in den Praktikumskripten zu finden sind (Meester & Maskill, 1995). Gleichzeitig vermuten sie eine starke Beziehung zwischen dem, was Lernende glauben tun zu müssen und tatsächlich im Labor tun. Nach Bussey, Orgill und Crippen (2013) kann der Mangel an klar vermittelten Lernzielen dazu führen, dass Lernende auf ihr Vorwissen zurückgreifen und eigene Vorstellungen von den Zielen des Praktikums entwickeln. Damit laufen Studierende die Gefahr, dass sie andere Lernziele wahrnehmen als intendiert sind und somit geringere Erfolgsaussichten im Laborpraktikum haben, da sie nur diejenigen Lernziele direkt erreichen können, die sie auch erkennen. Neuere Studien haben zwar bereits die Lernzielintentionen von Lehrenden für Laborpraktika (u. a. Bretz, Fay, Bruck & Towns, 2013) sowie die Lernzielwahrnehmungen der Studierenden (Galloway & Bretz, 2015a, 2015b, 2015c) untersucht, allerdings gibt es kaum Untersuchungen, die beide Perspektiven direkt kontrastieren (Wilkinson & Ward, 1997) und diese in Bezug zum Lernen und Erfolg im Laborpraktikum setzen.

Für die Erfassung der Lernzielintentionen und -wahrnehmungen wurde ein Online-Survey zur Beurteilung von Lernzielen des Laborpraktikums entwickelt und den Praktikumsassistenten (N = 5) wie auch den Studierenden eines Lehramtsjahrgangs (N = 71, 56 % weiblich, Alter = 20.15 Jahre, SD = 2.35) dargeboten. Die Ratings der jeweiligen Assistenten und der entsprechenden zu betreuenden Studierenden wurden anschließend einander zugeordnet und die Übereinstimmung über gewichtete Cohen's k Werte berechnet, welches als Maß für die Lernzielpassung genutzt wurde. Darüber hinaus wurden mit einem zuvor pilotierten Wissenstest (Elert & Walpuski, 2017) Daten zum Vorwissen der Studierenden erhoben, um dessen Einfluss auf die Lernzielvorstellungen zu überprüfen. Ebenfalls lagen Daten zu den Endnoten im Laborpraktikum als Maß für Praktikumserfolg vor.

Die Ergebnisse der Lernzielratings zeigen, dass die Studierenden im Verlauf des Laborpraktikums einen Paradigmen-Wechsel vornehmen: Von den insgesamt 32 allgemeinen Lernzielen wurden zwischen Anfang und Ende 12 signifikant wichtiger eingeschätzt, von welchen sich die meisten auf das klassische Laborpraktikum beziehen. Zuvor standen für sie insbesondere Lernziele eines inquiry labs im Vordergrund. Darüber hinaus zeigt sich eine große Spannweite in der Passung der Lernziele zwischen Assistenten und Studierenden ( $k_w = .000-.490$ ), welche aber theoriekonform sehr schlecht ist. Lineare Regressionsanalysen zeigen außerdem, dass das Vorwissen der Studierenden 9.5 % der Varianz in der Lernzielpassung aufklären ( $F(1,69) = 7.14, p = .009$ ), sodass von einem Zusammenhang zwischen Expertise im Fach und Lernzielwahrnehmung ausgegangen werden kann und für die angenommenen Modelle spricht. Trotzdem kann die Lernzielpassung keine Varianz in den Endnoten des Laborpraktikums aufklären ( $F(1,69) = .452, p = .504$ ). Diese und weitere Ergebnisse werden auf der Tagung selbst diskutiert.

## **Innovative Lehrveranstaltungen in der Lehrerbildung. Ansätze und Evaluationsergebnisse aus dem Projekt Level**

**Sebastian Stehle, Johannes Appel, Holger Horz**

Goethe-Universität Frankfurt, Deutschland; [stehle@psych.uni-frankfurt.de](mailto:stehle@psych.uni-frankfurt.de)

Die universitäre Lehrerbildung in Deutschland steht seit geraumer Zeit in der Kritik. Zu den am häufigsten genannten Kritikpunkten gehört die Fragmentierung der Studieninhalte durch die zumeist wenig aufeinander abgestimmte Vermittlung von Fachwissen, fachdidaktischem Wissen und allgemeinem pädagogischen Wissen in spezifische Curricula und durch verschiedene Institutionen. Daraus resultieren sowohl fehlende Verbindungen zwischen den verschiedenen Studieninhalten als auch Redundanzen, was es den Studierenden insgesamt erschwert, im Laufe des Studiums ein kohärentes Wissen aufzubauen (Terhart, 2004). Ein weiterer zentraler Kritikpunkt zielt auf das mangelhafte Theorie-Praxis-Verhältnis in der universitären Lehrerbildung sowie fehlende Konzepte für einen angemessenen Übergang von der universitären Ausbildung in die zweite Lehrerbildungsphase (Kotthoff & Terhart, 2013; Rumpf, 2002). Das heißt, auch an dieser Stelle gelingt es nicht in ausreichendem Maße, die verschiedenen Wissens- und Erfahrungsbereiche zu einem kohärenten Gesamtbild zu verbinden. Und schließlich verlangt die wachsende Heterogenität der Studierenden nach einer Flexibilisierung und Individualisierung des Lehramtsstudiums (Dräger et al., 2014).

Das im Rahmen der Bund-Länder Initiative Qualitätsoffensive Lehrerbildung (QL) an der Goethe-Universität Frankfurt implementierte Projekt "Level - Lehrerbildung vernetzt entwickeln" adressiert diese Defizite durch verschiedene Maßnahmen. (1) Zur Reduktion inhaltlicher Brüche zwischen Fächern und Ausbildungsphasen werden interdisziplinäre Lehr- und Forschungsk Kooperationen etabliert (Kunina-Habenicht, et al., 2012), in denen inhaltlich aufeinander abgestimmte Lehrangebote konzipiert und durchgeführt werden. (2) Zur Förderung anwendungsbezogener Aspekte professioneller Lehrkompetenz im Sinne der Theorie-Praxis-Verknüpfung werden Lernangebote entwickelt, in denen Videoaufzeichnungen authentischer Unterrichtssituationen analysiert werden (Seidel & Stürmer, 2014). Der Vorteil dieses methodischen Zugangs besteht darin,

dass alltägliche Handlungsprobleme, die in der Berufspraxis von Lehrpersonen virulent werden, anhand von Videomitschnitten des Unterrichtsgeschehens einer wiederholten Beobachtung und Analyse mit unterschiedlichen Fragestellungen und aus verschiedenen allgemein- und fachdidaktischen Perspektiven zugänglich gemacht werden können (vgl. Krammer, Lipowsky, Pauli, Schnetzler & Reusser, 2012). (3) Der durch die wachsende Heterogenität der Studierenden notwendigen Flexibilisierung und Individualisierung des Lehramtsstudiums wird mit dem Aufbau einer videobasierte Online-Lernplattform begegnet, mit deren Hilfe das Angebot an Blended Learning-Formaten vergrößert wird (Horz, 2015).

Übergeordnetes Ziel dieser Maßnahmen zur Weiterentwicklung der universitären Lehrerbildung ist die Verbesserung der professioneller Unterrichtswahrnehmung der Studierenden, d.h. der Fähigkeit, bedeutsame Unterrichtereignisse professionstypisch vor dem Hintergrund theoretischen Wissens wahrzunehmen und zu interpretieren (van Es & Sherin, 2008). Professionelle Unterrichtswahrnehmung gilt als wichtiger Indikator erfolgreicher Kompetenzentwicklung in der Lehrerbildung (Seidel & Stürmer, 2014) und bildet die theoretische Grundlage für die videobasierten Lernangebote im Projekt. Damit weisen Lehrveranstaltungen im Rahmen des Level-Projekts neben ihrem jeweiligen inhaltlichen Kern durchgängig eine einheitliche methodische sowie auch inhaltliche Akzentuierung auf und bieten damit die Basis für interdisziplinäre Kooperationen.

Zur Evaluation der Wirksamkeit der Lernangebote wurden im Sommersemester 2016 und im Wintersemester 2016/17 im Rahmen der Veranstaltungsevaluation Studierenden- und Dozierendenbefragungen in 39 Projekt-Lehrveranstaltungen durchgeführt (N=1.265 Studierende und 34 Lehrende). Die Verbindung beider Perspektiven gibt Aufschluss darüber, wie videobasierte und interdisziplinär konzipierte Blended Learning-Veranstaltungen gestaltet werden müssen, um einen erfolgreichen Kompetenzerwerb zu unterstützen.

Erhoben wurden Daten zur Evaluation allgemeiner und projektspezifischer Veranstaltungsmerkmale sowie personenbezogene Daten zur Kompetenzentwicklung der Studierenden (Prä-Post) und zur Lehrkompetenz der Dozierenden (Post). Die Reliabilität der eingesetzten Skalen zur Kompetenzeinschätzung hinsichtlich professioneller Unterrichtswahrnehmung sowie zu den anderen projektspezifischen Veranstaltungsmerkmalen ist zufriedenstellend bis gut ( $.71 < \alpha > .94$ ). Während sich über alle Projektveranstaltungen hinweg ein signifikanter Zuwachs selbsteingeschätzter professioneller Unterrichtswahrnehmung nachweisen lässt, zeigen multivariate Analysen erwartungskonforme Zusammenhänge zwischen bestimmten Veranstaltungsmerkmalen (UVs) und unterschiedlichen Kriterien für Veranstaltungsqualität (AVs). Im Beitrag werden die Ergebnisse der ersten beiden Projektsemester vorgestellt und Konsequenzen für die Weiterentwicklung der Lehr-Lern-Formate diskutiert.

## **Zum Umgang mit Vorwissensunterschieden zu Studienbeginn - Ein Beispiel eines differenzierenden Chemievorkurses**

**Carolin Eitemüller, Sebastian Habig**

Universität Duisburg-Essen, Deutschland; [carolin.eitemueller@uni-due.de](mailto:carolin.eitemueller@uni-due.de)

### Theoretischer Hintergrund

Nicht zuletzt vor dem Hintergrund der nach wie vor hohen Studienabbruchquoten (Heublein, 2017) in naturwissenschaftlichen Studienfächern in Deutschland beschäftigen sich Forschungsprojekte zunehmend mit den Fragen nach Studieneingangsvoraussetzungen und welche kognitiven und affektiven Faktoren zu einem Verbleib im Studium und letztendlich zum Studienerfolg beitragen (Albrecht, 2011; Buschhüter, Spoden, & Borowski, 2016; Busker, Parchmann, & Wickleder, 2010; Sumfleth & Leutner, 2016). Für das Studienfach Chemie ist insbesondere das fachspezifische Vorwissen zu Studienbeginn als ein wichtiger Prädiktor für Studienerfolg, gemessen an Modulabschlussnoten, identifiziert worden (Averbeck, Fleischer, Sumfleth, Leutner, & Brand, 2017).

Vor diesem Hintergrund stellt sich für viele Hochschulen die Frage nach dem richtigen Umgang mit Unterschieden im fachspezifischen Vorwissen von angehenden Studierenden. Als Instrument zur Abfederung der Vorwissensunterschiede existieren an vielen Universitäten Vor- und Brückenkurse, die vor Studienbeginn angeboten werden und angehenden Studierenden die Möglichkeit bieten sollen, ihr Vorwissen in unterschiedlichen Disziplinen aufzuarbeiten. Innerhalb weniger Wochen wird oberstufenrelevanter Unterrichtsstoff durchgearbeitet, was nicht selten dazu führt, dass Studierende mit geringen Vorkenntnissen überfordert und diejenigen mit hohem Vorwissen unterfordert werden. Obwohl es einzelne empirische Befunde zur Wirksamkeit von chemiebezogenen Vorkursen gibt (z. B. Busker, Klostermann, Herzog, Huber, & Parchmann, 2011), ist bisher wenig dazu bekannt, wie sich Vorkurse optimieren lassen, sodass sowohl angehende Studierende mit geringem Vorwissen, als auch angehende Studierende mit eher hohem Vorwissen davon profitieren. Ein möglicher Ansatzpunkt wäre ein differenzierendes Vorkursangebot, das am entsprechenden Vorwissensstand der Studierenden ansetzt und darauf aufbaut.

### Fragestellung und Zielsetzung

Die hier vorgestellte Untersuchung setzt an dieser Ausgangslage an und geht der Frage nach, ob mit einem chemiebezogenen Vorkursangebot, das vorwissenshomogene Übungsgruppen mit angepassten Übungsmaterialien vorsieht, ein höherer Lernzuwachs erzielt werden kann als durch ein reguläres Vorkursangebot ohne Differenzierungsmaßnahmen. Die zugrundeliegende Forschungsfrage lautet demnach:

Lässt sich mit einer äußeren Differenzierung des Vorkursangebots nach fachspezifischen Vorkenntnissen der Teilnehmenden eine Erhöhung des Lernzuwachses erzielen?

### Konzeption des Vorkurses und Untersuchungsdesign

Für alle chemieaffinen Studiengänge wird mit Ausnahme der Ingenieurwissenschaften an der Universität Duisburg-Essen ein gemeinsamer zweiwöchiger Chemievorkurs unmittelbar vor dem Beginn des ersten Semesters angeboten, der sich aus einer täglich stattfindenden Vorlesung und Übung zusammensetzt. Zur Ermittlung des Lernzuwachses wird das Fachwissen der Vorkursteilnehmenden in einem Prä-post-Design am ersten und letzten Tag des Vorkurses mit einem erprobten Chemie-Fachwissenstest (Freyer, 2013) erhoben. Neben demografischen Daten und der Abiturgesamtnote werden außerdem die schulische Vorbildung in den naturwissenschaftlichen Fächern und Mathematik erfasst. Am ersten Tag des Vorkurses werden die Teilnehmenden anhand ihrer Chemiekurswahl in der Oberstufe randomisiert einer von zwei Übungsgruppen zugeordnet. Da die Chemiekurswahl in der Oberstufe einer der stärksten Prädiktoren für das Vorwissen in der Allgemeinen Chemie darstellt (Averbeck et al., 2017), ist davon auszugehen, dass sich die Übungsgruppen bei dieser Zuordnung hinsichtlich ihrer fachspezifischen Kenntnisse signifikant voneinander unterscheiden. In einem Interventions-Kontrollgruppen-Design wird die Wirksamkeit der beschriebenen Differenzierungsmaßnahmen auf den Lernerfolg durch den Vergleich mit regulären Lehrbedingungen in einer Kontrollgruppe überprüft. Diese setzt sich aus Vorkursteilnehmenden mit verschiedenen Vorkenntnissen zusammen, welche mit Übungsaufgaben unterschiedlicher Schwierigkeitsniveaus lernen.

## Ergebnisse

Am Ende des Vorkurses liegen 133 vollständige Datensätze vor, die für Analysen genutzt werden können. Der Fachwissenstest weist eine mit  $\alpha = .75$  zufriedenstellende interne Konsistenz auf. Erste Ergebnisse zeigen, dass die Vorkursteilnehmenden im Mittel deutlich dazulernen ( $t(132) = 15.253$ ,  $p < .001$ ,  $d = 1.12$ ). Ein Vergleich der Lernzuwächse zwischen den Übungsgruppen mittels univariater Varianzanalyse mit dem Lernzuwachs als abhängiger Variable zeigt darüber hinaus, dass beide vorwissenshomogenen Übungsgruppen höhere Lernzuwächse erzielen als die Kontrollgruppe ( $F(2, 114) = 6.918$ ,  $p < .001$ ,  $d = 0.70$ ).

Demnach scheint ein differenzierendes Vorkursangebot, im Vergleich zu einer nicht differenzierenden Vorkurskonzeption, förderlich zu sein. Weitere Detailanalysen zu differenziellen Wirkungen der Differenzierungsmaßnahme sollen im Rahmen der GEBF-Jahrestagung präsentiert werden.

## A08

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Bernoullianum 32, H103

### **Professionelles Lehrerhandeln für die Implementation digitaler Medien – Eine Typologie von Lehrkräften auf Basis von ICILS 2013**

**Birgit Eickelmann, Kerstin Drossel**

Universität Paderborn, Deutschland; [birgit.eickelmann@uni-paderborn.de](mailto:birgit.eickelmann@uni-paderborn.de)

Theoretischen Ansätzen wie beispielsweise dem TPACK-Modell (Mishra & Koehler, 2006) zufolge benötigen Lehrkräfte, um digitale Medien erfolgreich in Lehr-Lern-Prozesse zu integrieren, sowohl allgemeindidaktische, fachdidaktische als auch technologische Kompetenzen. Im internationalen Vergleich zeigt sich auf empirischer Basis jedoch, dass insbesondere Lehrpersonen in Deutschland vergleichsweise selten in den vorgenannten Bereichen und mit Bezug auf den unterrichtlichen Einsatz digitaler Medien an Lehrerprofessionalisierungsmaßnahmen teilnehmen (Gerick, Schaumburg, Kahnert & Eickelmann, 2014). In bereits vorliegenden Studien werden diesbezüglich allerdings bisher vielfach nur Mittelwerte ausgewiesen, ohne dass das Professionalisierungsverhalten von Lehrpersonen genauer untersucht und intersubjektive Unterschiede sowie deren Zusammenhang mit der Nutzung neuer Technologien im Unterricht betrachtet werden. Diesem Desiderat geht der vorliegende Beitrag nach und untersucht, ob sich in Deutschland und im Vergleich dazu in anderen europäischen Ländern hinsichtlich der vorbeschriebenen Professionalisierungsbereiche verschiedene Lehrertypen identifizieren lassen (Forschungsfrage 1). Darüber hinaus wird analysiert, inwieweit sich Unterschiede zwischen den ermittelten Lehrertypen im Hinblick auf die unterrichtliche Nutzung digitaler Medien sowie auf weitere zentrale Prädiktoren (vgl. u.a. Fraillon et al., 2014) ergeben (Forschungsfrage 2).

Zur Bearbeitung der Forschungsfragen werden Sekundäranalysen der international vergleichenden Schulleistungsstudie IEA-ICILS 2013 (International Computer and Information Literacy Study, vgl. Bos, Eickelmann, Gerick et al., 2014) durchgeführt. Ergänzend zu den Lehrerdaten aus Deutschland werden Daten (N=8.709) aus denjenigen fünf europäischen Ländern (Dänemark, Niederlande, Norwegen, Polen und Tschechien) ausgewählt, deren Schülerinnen und Schüler über signifikant höhere computer- und informationsbezogene Kompetenzen verfügen als Jugendliche in Deutschland. Mittels Latent-Class-Analyse (Hagenaars & McCutcheon, 2002) werden unter Einbezug von jeweils drei Items zu technisch bzw. pädagogisch ausgerichteten Professionalisierungsmaßnahmen zunächst Lehrertypen identifiziert (Mplus; Version 7.7). In einem zweiten Schritt werden die identifizierten Typen im Zusammenhang mit der unterrichtlichen Computernutzung sowie dem IT-bezogenen Selbstkonzept und der Förderung von IT-bezogenen Schülerkompetenzen betrachtet (IDB-Analyzer; Version 3.1).

Im Ergebnis zeigt sich anhand der Informationskriterien BIC und AIC (Kuha, 2004) über alle Länder hinweg, dass eine Vier-Gruppen-Lösung die Daten am besten beschreibt. Lehrertyp 1, der fast die Hälfte aller Lehrpersonen ausmacht, ist dadurch charakterisiert, dass sowohl pädagogische als auch technische Professionalisierungsmaßnahmen mit geringer Wahrscheinlichkeit wahrgenommen werden. Zwei weitere Typen (2 und 3) repräsentieren jeweils etwa ein Fünftel der Lehrkräfte, wobei diese Lehrkräfte durchweg häufiger Professionalisierungsmaßnahmen durchführen und vor allem Lehrkräfte, die dem zweiten Typ zuzuordnen sind, häufiger an pädagogischen Professionalisierungsmaßnahmen teilnehmen. Nur etwa 10 Prozent der Lehrkräfte sind dem vierten Typ zuzuordnen und nehmen, wie in der Theorie als erforderlich abgebildet, sowohl an Maßnahmen mit pädagogischem als auch technisch ausgerichtetem Schwerpunkt teil. Betrachtet man vertiefend die Verteilung der Lehrpersonen in Deutschland auf die ermittelten Lehrertypen, so lassen sich zwei Drittel (67%) dem ersten Typ und lediglich 6 Prozent dem vierten Typ zuordnen.

Bezüglich der zweiten Forschungsfrage wird ersichtlich, dass Lehrkräfte, die sich umfassend fortbilden und Typ 4 zuzuordnen sind, in allen betrachteten Bildungssystemen zu höheren Anteilen als die anderen Typen angeben, regelmäßig digitale Medien im Unterricht einzusetzen. Dabei fällt vor allem der Unterschied im Vergleich zum ersten Typ deutlich aus. Auch das IT-bezogene Selbstkonzept ist bei Lehrkräften des vierten Typs ausgeprägter, ebenso wie die Förderung IT-bezogener Fähigkeiten der Schülerinnen und Schüler im Unterricht dieser Lehrkräfte. In Deutschland fällt dieser Unterschied besonders deutlich aus.

Folgt man dem vorliegenden theoretischen Ansatz und zieht hinzu, dass Lehrpersonen die ‚key stone species‘ für die erfolgreiche Implementierung digitaler Bildung in Schulen sind (vgl. Davis, Eickelmann & Zaka, 2014), unterstreichen die Ergebnisse auf empirischer Grundlage die Relevanz und Notwendigkeit des Angebots angemessener pädagogischer und technischer Professionalisierungsmaßnahmen. Für zukünftige Entwicklungen erscheint es zudem wichtig, Schulleitungshandeln zu fördern, das eine entsprechende Qualifizierung von Lehrkräften durch Maßnahmen unmittelbar auf der Schulebene unterstützt und passende innerschulische Angebote und anspruchsvolle Formen der Kooperation fördert (vgl. Drossel, Schulz-Zander & Eickelmann, 2016).

### **Professionswissen von Lehrpersonen der Sekundarstufe I in Deutschland zum Unterrichten mit digitalen Medien – Typenanalysen zur TPACK-Selbsteinschätzung und Zusammenhang mit der unterrichtlichen Computernutzung**

**Manuela Endberg, Ramona Lorenz**

TU Dortmund, Deutschland; [manuela.endberg@tu-dortmund.de](mailto:manuela.endberg@tu-dortmund.de)

Um den veränderten Anforderungen, die sich im Rahmen der Digitalisierung und Mediatisierung für die Schule ergeben, gerecht werden zu können, benötigen Lehrpersonen zunehmend Wissen und Können über die lernförderliche Einbindung digitaler Medien in schulische Lehr- und Lernprozesse (Aktionsrat Bildung, 2017; BMBF, 2016; KMK, 2016). Die dafür erforderlichen Kompetenzen werden hierzulande zumeist mittels des Ansatzes der medienpädagogischen Kompetenz (Blömeke, 2000) beschrieben. In Theorie- und Analysemodellen zur Erfassung des Lehrerprofessionswissens in Deutschland wird die medienpädagogische Kompetenz allerdings nicht berücksichtigt.

International ist mit dem Ansatz des Technological Pedagogical Content Knowledge (TPACK, Mishra & Koehler, 2006; Koehler & Mishra, 2009) ein Modell etabliert, mit dem Shulmans (1986) Ansatz des Pedagogical Content Knowledge (PCK) – als Kern des Lehrerprofessionswissens – um die Komponente des technischen Wissens erweitert wurde. Das TPACK-Modell setzt sich aus sieben Wissensbereichen zusammen, die in drei grundlegende Wissensbereiche (CK, PK, TK), drei kombinierte Wissensbereiche (PCK, TCK, PCK) und den Überschneidungsbereich aller Wissensbereiche (TPACK) gegliedert werden können. TPACK beschreibt das Wissen von Lehrpersonen darüber, wie effektives, kompetentes Lehren und Unterrichten mit digitalen Medien gelingen kann. Durch die direkte Anknüpfung an Shulmans Ausführungen zum Lehrerprofessionswissen, die auch in Deutschland etabliert sind (Baumert & Kunter, 2006), stellt TPACK einen Ansatz dar, mit dem das professionelle

Wissen und Handeln von Lehrpersonen im Kontext der unterrichtlichen Integration digitaler Medien beschrieben wird (Voogt, Fisser, Roblin, Tondeur & van Braak, 2013).

Obwohl der Diskurs um TPACK international breit geführt wird und wissenschaftliche wie auch bildungspraktische Implikationen liefert, ist der TPACK-Ansatz in Deutschland bisher kaum rezipiert. Vor dem Hintergrund, dass in Deutschland derzeit umfangreiche Untersuchungen, in denen der Zusammenhang zwischen dem Professionswissen von Lehrkräften und dem Einsatz digitaler Medien im Unterricht in den Blick genommen wird, und Instrumente zur Erfassung entsprechender Kompetenzen fehlen, wurde mit dem im Vortrag vorgestellten Forschungsvorhaben erstmalig das selbsteingeschätzte TPACK-Wissen von Lehrkräften der Sekundarstufe I in Deutschland erfasst. Dazu wurde ein valides TPACK-Fragebogeninstrument (Schmidt et al., 2009) eingesetzt, das übersetzt und für den Untersuchungskontext adaptiert wurde. Datengrundlage der Auswertungen bildet die repräsentative Lehrerbefragung (N = 1210) der Studie „Schule digital – der Länderindikator 2016“ (Bos et al., 2016). Im Vordergrund des Forschungsinteresses steht die Generierung von Erkenntnissen darüber, (1) wie Lehrkräfte in Deutschland ihr Professionswissen im Bereich des Einsatzes digitaler Medien entlang der TPACK-Wissensbereiche einschätzen, (2) ob sich unterschiedliche Lehrertypen hinsichtlich des eingeschätzten TPACK-Wissens unterscheiden lassen und (3) inwiefern das Wissen in diesem Bereich einen Zusammenhang mit der Nutzungshäufigkeit von Computern im Unterricht aufweist.

Die Ergebnisse zeigen, dass auf deskriptiver Ebene durchaus Unterschiede in der TPACK-Selbsteinschätzung der Lehrpersonen bestehen. Das pädagogische Inhaltswissen (PCK), als Kern des Professionswissens, wird im Durchschnitt am höchsten eingeschätzt, die weiteren Wissensbereiche, die einen direkten Bezug zum technischen Wissen aufweisen, werden im Mittel etwas geringer eingeschätzt. Hinsichtlich der zweiten Forschungsfrage lassen sich mittels Latent-Class-Analyse vier unterschiedliche Lehrertypen bezüglich der TPACK-Selbsteinschätzung klassifizieren, wobei insbesondere zwei Extremgruppen (sehr hohe TPACK-Selbsteinschätzung vs. sehr geringe TPACK-Selbsteinschätzung) kontrastiert werden können. Des Weiteren zeigt sich anhand von Regressionsmodellen in Bezug auf die dritte Forschungsfrage, dass die Ausprägung der TPACK-Selbsteinschätzung einen Zusammenhang mit der Nutzungshäufigkeit von Computern im Unterricht aufweist: Je höher die Kompetenzeinschätzung ausfällt, desto häufiger werden Computer im Unterricht genutzt. Der Effekt bleibt auch unter Kontrolle der Hintergrundvariablen Geschlecht und Alter bestehen.

Die Befunde stellen einen wichtigen Beitrag zum bestehenden Diskurs um Medienbildung und medienpädagogische Kompetenzen von Lehrpersonen in Deutschland dar. Zukünftig sollte die Entwicklung objektiver Kompetenztests für (angehende) Lehrpersonen verstärkt verfolgt werden, um zusätzlich zu selbsteingeschätzten Fähigkeiten auch messbare Kompetenzen erfassen zu können. Damit wäre weiterhin auch die Betrachtung von Effekten der Lehrerkompetenzen auf die Unterrichtsqualität und die Lernerfolge der Schülerinnen und Schüler denkbar.

### **Lehrkräfte als Digitalmuffel? Einfluss von fachdidaktischen Wissen und technologischen Überzeugungen auf die kognitive Aktivierung medienbasierten Unterrichts**

**Iris Backfisch<sup>1,2</sup>, Andreas Lachner<sup>1,2</sup>, Katharina Scheiter<sup>1</sup>, Frank Loose<sup>2</sup>, Christoff Hische<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für Wissensmedien, Tübingen, Deutschland; <sup>2</sup>Eberhard-Karls-Universität Tübingen; [i.backfisch@iwm-tuebingen.de](mailto:i.backfisch@iwm-tuebingen.de)

Digitale Medien im Unterricht bieten vielerlei Möglichkeiten die kognitive Aktivierung von Schüler\*innen zu fördern. In Mathematik können dynamische Visualisierungen Schüler\*innen unterstützen Konzepte besser nachzuvollziehen (Renkl & Scheiter, 2017); oder die Integration adaptiver Lernprogramme kann helfen, Lernaufgaben ad-hoc bereitzustellen (Alevin & Koedinger, 2002). Jedoch deuten aktuelle Befunde darauf hin, dass Lehrkräfte digitale Medien oft nur unzureichend und ohne deren Potential auszuschöpfen in ihrem Unterricht einsetzen (Harris et al., 2009, BertelsmannStiftung, 2017). Neben Faktoren auf der Schul- und der Infrastrukturebene, kann der mangelnde Einsatz digitaler Medien auch auf kognitive und motivationale Voraussetzungen der beteiligten Lehrkräfte zurückgeführt werden (Petko, 2012). Auf kognitiver Ebene, sollten dem TPACK-Modell folgend, Lehrkräfte über integriertes fachdidaktisches, fachwissenschaftliches, und technologisches Wissen verfügen, um digitale Medien didaktisch sinnvoll im Unterricht einsetzen zu können (Mishra & Koehler, 2006). Diese Wissensfacetten tragen nach Large-Scale-Studien auf Grundlage von Selbsteinschätzung der Lehrkräfte dazu bei, dass Technologien häufiger eingesetzt werden (Chai, Koh & Tsai, 2016). Zusätzlich wurde vor allem in qualitativen Studien ein Zusammenhang von Überzeugungen und Einstellungen von Lehrkräften und deren Technologie-Einsatz gezeigt (Tondeur et al., 2016).

Bisherige Analysen basieren eher auf der Untersuchung von Sichtstrukturen des Unterrichts, wie die Häufigkeit des Technologie-Einsatzes und nicht auf der Analyse der zugrundeliegenden Tiefenstrukturen, wie der kognitiven Aktivierung der Schüler\*innen durch den Einsatz digitaler Medien im Unterricht. Daher untersuchten wir, in einem Quasiexperiment mit Lehrkräften unterschiedlicher Statusgruppen (Lehramtsstudierende, Referendar\*innen, Lehrkräfte), ob und wie die kognitive Aktivierung medienbasierten Unterrichts durch die Wissensfacetten und motivationalen Voraussetzungen der Lehrkräfte bedingt ist. An der Untersuchung nahmen 29 Mathematiklehramtsstudierende, 13 Mathematikreferendar\*innen und 12 Mathematiklehrkräfte teil. Diese wurden gebeten ein medien-basiertes Unterrichtskonzept zum Satz des Pythagoras für eine 8. Gymnasialklasse zu erstellen. Vor der Unterrichtskonzeption wurden die kognitiven und motivationalen Voraussetzungen (Braak et al., 2004) der Lehrkräfte erfasst. Die Unterrichtsplanungen wurden hinsichtlich der kognitiven Aktivierung nach Hugener et al. (2009) inhaltsanalytisch ausgewertet ( $\kappa = .83$ ). Zudem wurde nach Puentedura (2006) die Ausschöpfung des Medienpotentials innerhalb des Unterrichtskonzepts nach einem hierarchischen 4-stufigen-Rating ausgewertet.

Wir erwarten, dass Lehrkräfte über mehr Wissen in den einzelnen Facetten (fachlich, fachdidaktisch und technologisch) verfügen, als Referendare und Studierende. Dennoch erwarten wir, dass die Studierenden am Offensten gegenüber dem Einsatz digitaler Medien im Unterricht sind. Bezüglich der Unterrichtskonzepte wird angenommen, dass Lehrkräfte die höchste kognitive Aktivierung mit digitalen Medien erreichen, und dabei deren Potential am höchsten ausschöpfen.

Hinsichtlich der kognitiven Voraussetzungen verfügten erwartungskonform Lehrkräfte und Referendar\*innen über mehr Fachwissen ( $F(2,50)=5.09$ ;  $p < .05$ ,  $\eta^2 = .18$ ), und mehr fachdidaktisches Wissen, ( $F(2,50)=4.27$ ;  $p < .05$ ,  $\eta^2 = .15$ ) als Studierende. Jedoch unterschieden sich die Statusgruppen nicht hinsichtlich ihres technologischen Wissens ( $F(2,50)=0.12$ ;  $p > .05$ ). Auf motivationaler Ebene teilten die Lehrkräfte dem Technologie-Einsatz im Unterricht überraschenderweise eine signifikant höhere Relevanz zu als Referendare und Studierende ( $F(2,50)=3.66$ ;  $p < .05$ ,  $\eta^2 = 0.13$ ).

Hinsichtlich der kognitiven Aktivierung, zeigte sich erwartungskonform, dass Lehrkräfte kognitiv aktivierende Unterrichtskonzepte gestalteten als Referendar\*innen und Studierende, ( $F(2,49)=5.4$ ,  $p < .01$ ,  $\eta^2=0,18$ ). Eine multiple hierarchische Regressionsanalyse mit kognitiver Aktivierung als Kriterium und den Wissensfacetten und Relevanzbewusstsein als Prädiktoren zeigte, dass lediglich das fachdidaktische Wissen ( $b=.29$ ,  $t(46)=2.15$ ,  $p<.05$ ) und das Relevanzbewusstsein ( $b=.51$ ,  $t(46)=4.03$ ,  $p<.001$ ) einen signifikanten Anteil der kognitiven Aktivierung der Unterrichtskonzepte erklärten,  $R^2 = .32$ ,



$F(4,64)=5.30$ ,  $p < .001$ . Das Fachwissen und technologische Wissen der Teilnehmer\*innen hingegen trugen nicht signifikant zur kognitiven Aktivierung bei. Hinsichtlich des Potentials digitaler Medien, zeigte sich, dass dies alle Statusgruppen in ihren Unterrichtskonzepten gleichermaßen nicht voll ( $M=2.72$ ,  $SD=.78$ ) ausschöpfen ( $F(2,47)=1.03$ ,  $p < .05$ ).

Aus theoretischer Perspektive deuten die Befunde darauf hin, dass fachdidaktisches Wissen und Relevanzbewusstsein notwendige Grundvoraussetzungen für die Planung medienbasierten Unterrichts sind. Aus praktischer Perspektive zeigen die Befunde, dass weitere Forschung notwendig ist, wie Lehrkräfte dabei unterstützt werden können, das Potential digitaler Medien gänzlich auszuschöpfen.

## **Professionelles Handeln von Lehrpersonen der Sekundarstufe I in Deutschland zur Förderung der computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern**

**Ramona Lorenz, Manuela Endberg**

TU Dortmund, Deutschland; [ramona.lorenz@tu-dortmund.de](mailto:ramona.lorenz@tu-dortmund.de)

Ein kompetenter Umgang mit digitalen Medien ist zu einer unverzichtbaren Kompetenz in der Wissens- und Informationsgesellschaft geworden. Mit dem Beschluss ‚Bildung in der digitalen Welt‘ der Kultusministerkonferenz (KMK, 2016) ist die Vermittlung medienbezogener Kompetenzen an Schülerinnen und Schüler daher eine verpflichtende Aufgabe der Lehrkräfte aller Unterrichtsfächer geworden.

In theoretischen Modellen der Schuleffektivität wird zwischen den Ebenen des Inputs, des Prozesses und des Outputs unterschieden. Bezogen auf den Bereich der schulischen Medienbildung mit besonderem Fokus auf Lehrpersonen lassen sich als Voraussetzungen auf der Ebene des Inputs die Einstellungen von Lehrkräften zum Einsatz digitaler Medien im Unterricht sowie ihre medienbezogenen Kompetenzen verorten. Auf der Prozessebene kann im Hinblick auf schulische Prozesse die medienbezogene Kooperation von Lehrkräften hervorgehoben werden. Einen zentralen unterrichtlichen Prozess stellt die Nutzung digitaler Medien dar, die für die Förderung der Kompetenzen von Schülerinnen und Schülern notwendig erscheint.

Der bisherige Forschungsstand umfasst Hinweise auf unzureichende computer- und informationsbezogene Kompetenzen der Schülerinnen und Schüler in Deutschland (Bos et al., 2014) sowie Analysen zu Zusammenhängen einzelner lehrkraftbezogener Merkmale im Medienkontext mit Schülerkompetenzen bzw. der Förderung dieser (u.a. Lorenz & Endberg, 2016). Ein Desiderat besteht allerdings noch in der gemeinsamen Betrachtung von lehrkraftbezogenen Merkmalen auf den verschiedenen Ebenen, die sich in diesem Kontext als relevant erwiesen haben (u.a. Drossel et al., 2016; Petko, 2012; Voogt et al., 2013), und deren Effekte auf die Förderung der Schülerkompetenzen. Daraus ergibt sich die Forschungsfrage: Welche lehrkraftbezogenen Faktoren auf der Input- und Prozessebene weisen einen Zusammenhang zur Förderung computer- und informationsbezogener Kompetenzen der Schülerinnen und Schüler auf?

Dieser Frage wird mit der Studie ‚Schule digital – der Länderindikator‘ nachgegangen, die im Erhebungszyklus 2016 eine repräsentative Datengrundlage von 1210 Lehrpersonen der Sekundarstufe I in Deutschland umfasst. Mittels Strukturgleichungsmodell (Mplus 7) werden lehrkraftbezogene Merkmale im Kontext schulischer Medienbildung berücksichtigt und ihr Zusammenhang zur Förderung der computer- und informationsbezogenen Kompetenzen der Schülerinnen und Schüler analysiert. Berücksichtigt werden dabei auf der Inputebene die Einstellungen gegenüber dem schulischen Medieneinsatz sowie die Selbsteinschätzung ihrer Kompetenzen zur Integration digitaler Medien in den Unterricht. Auf der Prozessebene wird die medienbezogene Kooperation von Lehrkräften einbezogen. Diese drei Merkmale wurden jeweils als latente Konstrukte modelliert. Zudem wurde die Nutzungshäufigkeit digitaler Medien im Unterricht manifest berücksichtigt. Als abhängige Variable dient ein latent modelliertes Konstrukt der Förderung computer- und informationsbezogener Kompetenzen, das fünf Variablen zu Arbeitsweisen im Unterricht umfasst. Die Fitwerte deuten insgesamt auf eine gute Modellpassung hin.

Im Ergebnis zeigt sich, dass die selbsteingeschätzten Lehrerkompetenzen sowie die Einstellungen der Lehrpersonen signifikant positiv miteinander korrelieren und zudem einen positiven Zusammenhang zu den Merkmalen der Kooperation und der Nutzungshäufigkeit auf der Prozessebene aufweisen. Die beiden Prozessmerkmale hängen wiederum signifikant und positiv mit der Förderung computer- und informationsbezogener Kompetenzen zusammen. Unter zusätzlicher Berücksichtigung indirekter Effekte der beiden Merkmale auf der Inputebene weist nur die medienbezogene Kooperation unter Lehrkräften einen signifikanten Zusammenhang zur Förderung der Schülerkompetenzen auf, jedoch nicht mehr die Nutzungshäufigkeit digitaler Medien. Die Effektstärken weisen darauf hin, dass die Selbsteinschätzung der medienbezogenen Kompetenzen von Lehrkräften den relevantesten Prädiktor unter den herangezogenen Konstrukten darstellt.

Implikationen lassen sich insbesondere für die Lehreraus- und -fortbildung ableiten, für die es zu klären gilt, wie die medienbezogenen Kompetenzen von Lehrkräften in der Fachdidaktik, der Schulpädagogik oder der Medienpädagogik aber auch im Referendariat sinnvoll eingebunden werden kann, sodass ein phasenübergreifender, sukzessiver Kompetenzaufbau ermöglicht wird. Zudem muss der Bereich der Lehrerfortbildung so gestaltet werden, dass Lehrkräfte flächendeckend zum lernförderlichen (in Bezug auf das fachliche Lernen) und adaptiven (in Bezug auf die Medienkompetenz) Einsatz digitaler Medien befähigt werden. Die Ergebnisse heben zudem deutlich hervor, dass nicht die Nutzungshäufigkeit digitaler Medien, sondern insbesondere lehrkraftbezogene Merkmale auf der Input- und Prozessebene eine wesentliche Bedeutung im Kontext schulischer Medienbildung zukommt.

## A09

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Bernoullianum 32, H223

### Kooperation zwischen Lehrkräfte an deutschen Schulen

**Julio Gregorio Mora-Ruano, Eveline Wittmann**

Technische Universität München, Deutschland; [juliogregorio.mora-ruano@tum.de](mailto:juliogregorio.mora-ruano@tum.de), [eveline.wittmann@tum.de](mailto:eveline.wittmann@tum.de)

#### 1. Theoretischer Hintergrund

Lehrerkooperation wird häufig als Dreh- und Angelpunkt der Schul- und Personalentwicklung, sowie für die Schulleistung und die Schulwirksamkeit (Steinert et al. (2006); KMK (2003)). Ditton (2000) und Scheerens (1990) platzieren sie auf der Prozessebene von Modellen der Schulqualität sowie. Zudem zeigt sich ein hoher Zusammenhang mit der Selbstwirksamkeit und der Arbeitszufriedenheit der Lehrkräfte (OECD, 2016). Bisherige Studien weisen darauf hin, dass das Ausmaß der Zusammenarbeit von Lehrpersonen an deutschen Schulen insgesamt relativ gering ist (Gräsel, 2010), dass Gymnasial-Lehrkräfte am wenigsten kooperieren (Kullmann (2008); Projekt-MARKUS (2002)) und dass es keinen signifikanten Unterschied bezüglich des Geschlechts gibt (Richter & Pant, 2016).

#### 2. Fragestellung

Ziel der Arbeit ist es, offen zu legen, wie stark bzw. schwach die Kooperation zwischen Lehrkräften an deutschen Schulen unter Berücksichtigung verschiedener Schulmerkmale (z.B. Schulart) und soziodemographische Variablen (z.B. Geschlecht, Alter) ist.

Wir möchten die folgenden Forschungsfragen beantworten:

1. Inwiefern beeinflusst die Schulart die Kooperationshandlungen zwischen den Lehrkräften?
2. Inwiefern beeinflusst das Geschlecht die Kooperationshandlungen zwischen den Lehrkräften?
3. Inwiefern beeinflusst das Alter die Kooperationshandlungen zwischen den Lehrkräften?
4. Gibt es einen signifikanten Interaktionseffekt zwischen Schulart und Geschlecht?

#### 3. Methode

Diese Studie führt eine Sekundäranalyse durch und es werden Daten von PISA 2012 aus dem deutschen nationalen Fragebogen für Lehrerinnen und Lehrer verwendet (Fragebogen von Bosker & Hendrick, 1997). Durch das IPN (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik) wurden drei verschiedene Arten von Lehrerkooperation aus den Items gebildet (Unterrichts- Projekt- und Organisationbezogene Kooperation).

##### 3.1. Design

Das Design handelt sich um eine repräsentative Querschnittserhebung.

##### 3.2. Stichprobe

2084 Lehrer in Deutschland, davon 1040 weiblich, 917 männlich ,127 nicht spezifiziert, aus verschiedene Schularten (Hauptschule: 231, MBG: 305, Realschule: 397, Gesamtschulen: 251, Gymnasium: 753, Berufliche Schulen: 47, Sonder- und Förderschule: 98) im Alter von 25 bis 71 Jahren (M = 46 Jahre, SD = 11,4) haben teilgenommen.

##### 3.2. Statistische Auswertung

Eine zweifaktorielle MANCOVA (Schulart \* Geschlecht mit Alter als Kovariate) wurde durchgeführt um Interaktions- und Haupteffekte zwischen den schon genannten Variablen herauszufinden. Die verwendeten Skalen hatten eine Zuverlässigkeit zwischen  $\alpha = 0,70$  und 90.

#### 4. Ergebnisse

Ein bemerkenswertes Ergebnis unserer Analyse ist, dass nur die Schulart eine große und signifikante Auswirkung auf die gesamte Kooperation und ihre getrennten Formen hat. Das Geschlecht und das Alter ergaben gemischte Ergebnisse bezüglich ihrer Signifikanz, jedoch sind die Auswirkungen auf die Häufigkeit der Kooperation klein. Das Geschlecht hatte nur einen signifikanten Effekt auf die organisationsbezogene Kooperation, das Alter hingegen hatte einen positiven signifikanten Effekt auf die unterrichts- und projektbezogene Kooperation.

Darüber hinaus, unabhängig von der Schulart, hatte die unterrichtsbezogene Kooperation den höchsten Mittelwert, während die projektbezogene Kooperation den geringsten Mittelwert hat.

Des Weiteren haben die Gymnasial-Lehrkräfte den kleinsten Mittelwert bei der projektbezogenen Kooperation sowie bei der organisationsbezogenen Kooperation. Nur auf der unterrichtsbezogenen Kooperationsebene sind die Realschul-Lehrkräfte die Gruppe, die am wenigsten kooperieren.

Männer kooperieren weniger als Frauen bei allen Kooperationsebenen, jedoch sind diese Unterschiede nur bezogen auf die organisationsbezogene Kooperation signifikant.

### Ganztags- versus Halbtagschule: Effekte auf Leistungsniveau und Bildungsungleichheit

**Isa Steinmann, Rolf Strietholt**

Technische Universität Dortmund, Deutschland; [isa.steinmann@tu-dortmund.de](mailto:isa.steinmann@tu-dortmund.de)

Schüler/-innen verbringen international auch außerhalb des regulären Unterrichts Zeit in der Schule (OECD, 2015; Plantenga & Remery, 2013). Häufig haben solche Ganztagschulmodelle eine Leistungsniveauförderung und einen Bildungsungleichheitsabbau zum Ziel (vgl. Ecarius et al., 2013; Mahoney et al., 2005). Theoretische Modelle legen gleichzeitig nahe, dass Lernmehrwerte nur unter bestimmten Bedingungen erwartet werden können. Beispielsweise sollten die zusätzlichen Zeiträume lernunterstützend gestaltet sein, mit dem regulären Unterricht verbunden und von qualifiziertem Personal beaufsichtigt (vgl. Miller & Truong, 2009; Vandell et al., 2007). Diese Annahmen werden durch internationale Befunde gestützt. Sommerschul- oder Nachmittagsprogramme zeigen sich (nur) als förderlich, wenn sie leistungssteigernd ausgerichtet sind, bestimmte Zielgruppen fokussieren, durch professionelles Personal betreut werden, etc. (Durlak et al., 2010; Lauer et al., 2006; Patall et al., 2010; Roth et al., 2010). Solche Prozessmerkmale und Definitionen ganztägiger Beschulung variieren international jedoch stark (z.B. OECD, 2015; Plantenga & Remery, 2013).

In Deutschland werden Ganztagschulen beispielsweise eher anhand struktureller (zeitlicher Rahmen, formelle Zuständigkeit, etc.) als pädagogischer (Inhalte, Curricula, etc.) Kriterien definiert, was zu einer großen Ausgestaltungsvariation beiträgt (z.B. Fischer & Klieme, 2013; Fischer et al., 2014). Vor dem Hintergrund theoretischer Annahmen und internationaler Befunde überrascht demnach nicht unbedingt, dass die einzige vorliegende Studie, die Leistungs- und Bildungsungleichheitsunterschiede zwischen Halb- und Ganztagschulen robust untersucht, keine signifikanten Effekte findet (Strietholt et al., 2015). Andere kausal interpretierbare Studien untersuchen Teilnahmeeffekte innerhalb von Ganztagschulen und finden keine oder kleine positive Leistungseffekte (Bellin & Tamke, 2010; Fischer et al., 2016; Lossen et al., 2016).

#### Fragestellung

In der vorliegenden Studie ergänzen wir den limitierten Forschungsstand in Deutschland und untersuchen, ob Ganztagschulen einen Mehrwert gegenüber Halbtagschulen in Hinblick auf eine Steigerung des Leistungsniveaus und des Abbaus von Bildungsungleichheiten bilden.

#### Methode

Wir nutzen Daten der repräsentativen National Educational Panel Study (NEPS; Blossfeld et al., 2011). 4343 Fünftklässler/-innen an 191 Schulen wurden ab 2010/11 zwei Jahre verfolgt. Standardisierte Leistungstests mit EAP/PV-Reliabilitäten von mindestens 0,8 in Mathematik, Lesen und Rechtschreibung wurden in Klasse 5 und 7 erhoben. Zu Beginn wurden außerdem Hintergrundinformationen wie der sozio-ökonomische Status und der Migrationshintergrund erfragt. In Klasse 6 wurden Daten zur Ganztagsstruktur (Halbtagschule, offene Ganztagschule, teilgebundene Ganztagschule, vollgebundene Ganztagschule) und zu angebotenen Programmarten (Hausaufgabenhilfe, Deutsch/Literatur, Mathematik) bei den Schulleitungen erfragt. Die Schüler/-innen gaben an, wie häufig sie an Nachmittagsangeboten teilnahmen. Die Schüler/-innenangaben wurden aggregiert, um die Teilnahmeintensität auf Schulebene zu quantifizieren. Neben der Ganztagsstruktur betrachten wir demnach zwei Prozesscharakteristika: die angebotenen Programmarten und die mittlere Teilnahmeintensität.

Wir schätzen separate Mehrebenenregressionsmodelle für die drei abhängige Variablen und beide Variablen sozialer Herkunft. In den Hauptanalysen regressieren wir auf Schüler/-innenebene die Leistung in Klasse 7 auf die soziale Herkunft (Random Slope) und kontrollieren die Vorleistung in Klasse 5 sowie die Zeit zwischen den Tests. Auf Schulebene regressieren wir sowohl Intercept als auch Random Slope auf die Ganztagsstruktur-Dummies und kontrollieren den Schultyp. Für die Gruppe der Ganztagschulen führen wir vertiefende Analysen zu den Prozesscharakteristika durch. Auf Schulebene regressieren wir hierbei den Intercept auf jeweils zwei inhaltlich verbundene Programmarten, die durchschnittliche Teilnahmeintensität und den Schultyp.

#### Ergebnisse

Die Ergebnisse bieten keine Evidenz für Effekte der Ganztagsstruktur auf die Schulleistung in Mathematik, Lesen und Rechtschreibung. Dasselbe gilt für den Zusammenhang zwischen Leistungsmaßen und dem sozio-ökonomischen Status bzw. Migrationshintergrund. Demnach unterscheiden sich Halbtagschulen, offene, teilgebundene und vollgebundene Ganztagschulen weder in Bezug auf das Leistungsniveau, noch die Bildungsungleichheit.

Wir finden in den weiteren Analysen darüber hinaus keinen Effekt der Prozesscharakteristika auf die Leistungs- oder Ungleichheitsmaße. Ganztagschulen, die Angebote in den untersuchten Bereichen bereitstellen oder eine hohe Teilnahmeintensität erzielen, weisen demnach keine höheren Leistungsniveaus oder niedrigeren Bildungsungleichheiten auf, als andere Ganztagschulen.

## **Zwischen pädagogischem Anspruch und De-Professionalisierungsfurcht. Die Mitwirkung von Lehrpersonen im schulischen Ganztagsbetrieb**

**Christine Steiner**

Deutsches Jugendinstitut e.V. (DJI), Deutschland; [steiner@dji.de](mailto:steiner@dji.de)

Die Zusammenarbeit von Lehrpersonen mit pädagogischen Fachkräften und Laien wurde zuletzt durch den Ausbau von Ganztagschulen nachhaltig gefördert und mit hohen Erwartungen an die Professionsentwicklung verbunden. So wird von einer multiprofessionellen Zusammenarbeit eine generell höhere pädagogische Qualität etwa durch die Veränderung klassischer Lehr-Lern- Arrangements erwartet. Zugleich wurden aber De-Professionalisierungsbefürchtungen laut.

Bisher vorliegende Befunde zur Arbeit in multiprofessionellen Ganztagssteams zeigen jedoch, dass beide Erwartungen nicht zutreffen. Vielmehr wird die Zusammenarbeit bei weitreichender Kritik an den bestehenden Arbeitsbedingungen als wechselseitige Ergänzung wahrgenommen (dazu auch Böttcher u.a. 2011, S. 108ff.) und insbesondere von Lehrkräften als gewinnbringend und entlastend beschrieben (u.a. Dinzinger u.a. 2011, S. 108ff.). Breuer (2011) zeigte, dass diese positive Einschätzung auch darauf zurückzuführen ist, dass sich Teams auf eine Arbeitsteilung entlang der jeweiligen Profession als typisch erachteten Tätigkeiten verständigen (S. 95f.). Dies führe zwar zu einer ungleichen Arbeitsteilung, zugleich bliebe dadurch die jeweilige professionelle Autonomie gewahrt und ein eigenständiges pädagogisches Handeln möglich (S. 97).

Prozesse der Arbeitsteilung und der Wahrung von bzw. dem Streben nach Autonomie sind mit dem herkömmlichen Professionsverständnis allerdings kaum thematisierbar. Im Rückgriff auf Luhmanns Zurechnung von Entscheidungen auf die Organisation und von Interaktionen auf die Profession wird Teamarbeit folgerichtig als weitergehende Ausdifferenzierung der Organisation betrachtet. Um sie in den Blick zu bekommen, haben Kolbe und Reh (2008) den Begriff der Kollegialität vorgeschlagen. Kollegialität unterlaufe nicht nur weitreichende Ansprüche an Teamarbeit, sie schütze den Einzelnen auch vor inkompetenter Einflussnahme und wahre die individuelle Verantwortung für Schüler/-innen.

Unter der Hand wird damit jedoch ein anderes Professionsverständnis aufgerufen, das vor allem auf Prozesse der sozialen Schließung abhebt (vgl. Stock 2006). Im Mittelpunkt des Beitrages steht am Beispiel des Engagement von Lehrkräften im Ganztags die Frage, inwieweit dieses Engagement durch Aspekte sozialer Schließung geprägt ist. Dabei sind nicht nur Grenzziehungen im Sinne einer klaren Arbeitsteilung bzw. Beschränkung auf spezifische Tätigkeiten von Interesse, sondern auch die gezielte Gestaltung von zulässigen Grenzüberschreitungen.

Empirische Basis des Beitrags sind Daten der „Studie zur Entwicklung von Ganztagschulen“ (StEG), in deren Rahmen zwischen 2005 und 2009 an über 300 Schulen des Primarbereiches und der Sekundarstufe I aus 14 Bundesländern insgesamt drei umfangreiche Erhebungen aller am Ganztagsbetrieb beteiligten Personengruppen durchgeführt wurden. Für die Analysen wurden die Daten der im Jahr 2009 durchgeführten Lehrkraftbefragung genutzt. Es handelt sich dabei um 6.211 Lehrende an 313 Schulen, darunter 85 Primarschulen. Analysiert werden auf Basis von Cluster- und Mehrebenen-Modellen insbesondere zwei Aspekte der Arbeit in multiprofessionellen Ganztagssteams analysiert, nämlich zum einen die Intensität der aktiven

Mitwirkung der Lehrkräfte im Ganztagsbetrieb und ihrer individuellen Arbeitsschwerpunkte. Zum andern wird untersucht, wer unter den Lehrkräften über ein eher „professionsuntypisches“ Tätigkeitsprofil verfügt.

Die Analysen zeigen zunächst sowohl für den Primärbereich als auch für Lehrkräfte in der Sekundarstufe I, dass der Umfang ihrer Mitwirkung im schulischen Ganztagsbereich vor allem von ihrem individuellen pädagogischen Selbstverständnis, aber auch vom absoluten Umfang mitarbeitender Kolleg/-innen und der Organisationsform des Ganztages abhängt. Zugleich zeigen sich auch alters-, geschlechts- und regionalspezifische Unterschiede. Gleichwohl wird unter den vier für Lehrkräfte ermittelten distinkten Tätigkeitsprofilen dasjenige am häufigsten von Lehrpersonen vertreten, in dem fast ausschließlich unterrichtsbezogenen Angebotsformate versammelt sind. Lehrpersonen ohne diese Schwerpunktsetzung unterscheiden sich von ihren Kolleg/innen durch Merkmale, die ihr pädagogisches „Grenzgängertum“ begründbar und damit auch legitimierbar macht.

## **Haben Kooperation und Kollegialität in Lehrerkollegien an Ganztagsgrundschulen einen Einfluss auf ihre Innovationsbereitschaft? – Effekte auf Individual- und Schulebene**

**Frederik Osadnik<sup>1</sup>, Karin Lossen<sup>1</sup>, Wolfram Rollett<sup>2</sup>**

<sup>1</sup>TU Dortmund, Deutschland; <sup>2</sup>PH Freiburg; [frederik.osadnik@tu-dortmund.de](mailto:frederik.osadnik@tu-dortmund.de)

Theoretischer Rahmen und Forschungsstand

Der konsequente und bundesweite Ausbau von Ganztagsgrundschulen in Deutschland (KMK 2015,2004) stellt Schulen als Organisation vor neue Herausforderungen und setzt bei allen beteiligten Akteuren ein hohes Maß an Veränderungsbereitschaft voraus. Dabei nehmen besonders Lehrkräfte eine zentrale Rolle in der aktiven Gestaltung und Implementation von Innovationen ein (George, Hall & Stiegelbauer 2006). Organisationsstrukturelle Bedingungen, die eine förderliche Wirkung auf die Innovationsbereitschaft von Kollegien aufweisen stehen hierbei besonders im Fokus der Forschung, (Lossen, Rollett & Willems 2013; Geijssel, Van Den Berg & Slegers 1999) da sie erfolgreichere Schulentwicklungsprozesse, wie beispielsweise den Ausbau zur Ganztagschule, begünstigen. In diesem Zusammenhang stellen Kollegialität und Kooperation im Kollegium relevante Konstrukte dar. Kooperation bezogen auf Zusammenarbeit und Kollegialität im Sinne von Beziehungsqualität (Kelchtermans, 2006) beschreiben verwandte Konstrukte, die oftmals nur unter dem Begriff Kooperation subsumiert werden. Trotz ihrer engen Beziehung zueinander ist es notwendig beide Variablen getrennt zu betrachten. Empirisch finden sich hohe Korrelationen zwischen Kooperation und Kollegialität (Gerecht, Steinert & Döbrich, 2009; Bergmann & Rollett, 2008).

Geijssel, van den Berg und Slegers (1999) konnten in ihren Studien nachweisen, dass in innovationsfähigen Schulen eine konsequente Zusammenarbeit im Kollegium stattfindet. Ebenso fördert eine als kollegial wahrgenommene Atmosphäre die Veränderungsbereitschaft von Grundschullehrkräften (Clement & Vandenberghe 1999). Schulische Verbesserungen werden erreicht, indem Lehrkräfte gemeinsam planen und gestalten, damit sie in ihren Diskussionen bewährte Praktiken teilen und entwickeln können (Chrispeels, Andrews & Gonzales, 2007).

Außerdem gelingt es Schulen besser, Reformen umzusetzen, wenn die Lehrerschaft aktiv kooperiert und gemeinsam dazulernt (Hubbard, Mehan & Stein 2006).

Fragestellungen

- 1) Führen mehr Kooperation und Kollegialität zu einer günstigeren Entwicklung der Innovationsbereitschaft?
- 2) Führt eine bessere Entwicklung von Kooperation und Kollegialität auch zu einer besseren Entwicklung der Innovationsbereitschaft?
- 3) Gibt es hinsichtlich des Einflusses von Kooperation und Kollegialität auf die Innovationsbereitschaft Unterschiede zwischen der Individual- und der Schulebene?

Methoden

Der Beitrag basiert auf längsschnittlichen Daten aus der ‚Studie zur Entwicklung von Ganztagsgrundschulen (StEG)‘ über drei Messzeitpunkte (MZPe). Einbezogen wurden Angaben von 859 Lehrkräften (77 männlich, 633 weiblich) aus 75 Ganztagsgrundschulen in 14 Bundesländern, die zu allen MZPen als Mitglieder des Kollegiums gelistet waren. Fehlende Werte wurden in SPSS mit einem EM-Algorithmus imputiert und anschließend ein Zwei-Ebenen-Cross-Lagged-Panel-Model in Mplus 7 berechnet.

Für die Analysen wurden folgende Skalen verwendet: Die ‚Intensität der Lehrer- Kooperation‘ (20 Items;  $\alpha$ : .85-.87) beschreibt die Häufigkeit verschiedener Aspekte der Kooperation unter Lehrkräften wie Austausch über pädagogische Ziele, gemeinsame Planung von Unterricht oder die Hausaufgabenpraxis. ‚Kollegialität‘ (7 Items;  $\alpha$ : .90-.91) bezieht sich auf das soziale Klima im Kollegium, im Sinne von Gemeinschaftsgeist, Geschlossenheit oder Umgang mit Konflikten. Die ‚Innovationsbereitschaft‘ (6 Items;  $\alpha$ : .87-.88) der Kollegien beinhaltet unter anderem Aufgeschlossenheit für neue pädagogische Ansätze oder das Engagement für Entwicklungen und Erneuerungen.

Ergebnisse

In einem ersten Schritt ergaben die Analysen, dass 14-20% der Varianz der Intensität der Lehrerkoooperation auf die Schulzugehörigkeit zurückzuführen sind. Für die Kollegialität waren es um 45% und bezüglich der Innovationsbereitschaft 32-36%. Im Zwei-Ebenen-Cross-Lagged-Panel-Model zeigt sich, dass eine sich besser entwickelnde Kooperationskultur die Innovationsbereitschaft der Lehrkräfte fördert. Auf Individualebene geht eine bessere Entwicklung von Kooperation und Kollegialität mit einer günstigeren Entwicklung der Innovationsbereitschaft einher. Auf Schulebene finden sich diese Zusammenhänge positiver Entwicklungen hauptsächlich als Effekt der Kollegialität auf die Innovationsbereitschaft. Bessere Ausgangsbedingungen eines sich verbessernden sozialen Klimas führen dagegen zu keiner positiveren Entwicklung von Innovationsbereitschaft und Kooperationsintensität: Unter Kontrolle der positiven Effekte lassen sich sogar nachteilige Effekte des kollegialen Zusammenhalts auf die Innovationsbereitschaft und - auf lange Sicht - auch auf die Lehrerkoooperation identifizieren.

## A10

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus S107

### Videogestützte Unterrichtsreflexion im Praxissemester

**Andreas Gösch<sup>1</sup>, Kerstin Göbel<sup>2</sup>**

<sup>1</sup>Universität Duisburg-Essen, Deutschland; <sup>2</sup>Universität Duisburg-Essen, Deutschland; [andreas.goesch@uni-due.de](mailto:andreas.goesch@uni-due.de)

#### Theoretischer Hintergrund

Die Ermöglichung situativer Lerngelegenheiten in der berufspraktischen Lehrerbildung stellt eine wichtige Motivationsquelle und ein Qualitätsmerkmal in der Lehrerbildung dar (KMK, 2004). Die videogestützte Reflexion von Unterricht kann eine Ergänzung zu berufspraktischen Lerngelegenheiten darstellen, die das professionelle Wissen angehender Lehrkräfte aktiviert und Transferperspektiven in Anwendungskontexte ermöglicht (Blomberg et al., 2013). Der allgemeine Nutzen von Unterrichtsvideos in der Aus- und Weiterbildung von Lehrkräften konnte bereits empirisch bestätigt werden (u.a. Miller & Zhou, 2007). So kann videogestützte Unterrichtsreflexion dazu beitragen, theoretische Konzepte zum Unterricht zu aktivieren und die Integration von Praxis und Theorie zu fördern (Kleinknecht & Schneider, 2013). Dabei leisten sowohl die Reflexion von eigenen wie auch von fremden Unterrichtsvideos jeweils ihren eigenen Beitrag zur Reflexionsunterstützung (Seidel, 2011). Studierende wünschen sich von Begleitveranstaltungen zu den Praxisphasen in der Lehrerbildung, dass diese die nötigen Kenntnisse und Fähigkeiten zur Vorbereitung und Durchführung von Unterricht vermitteln (Gröschner & Seidel, 2012; Brouwer & Korthagen, 2005), die Integration videogestützter, situativer Lernangebote mit Bezug zu eigenen Unterrichtserfahrungen erscheint in diesem Kontext vielversprechend. Der Rückgriff auf kollegiale Reflexionsstrukturen hat sich in bisherigen Forschungsansätzen als besonders ertragreich für die Reflexion eigenen Unterrichts erwiesen, da durch den Austausch über unterschiedliche Einschätzungen und Beobachtungen neue Anregungen und Impulse entstehen können (Wyss, 2013). In diesem Zusammenhang erscheint das aus der systemischen Beratung entstammende Prinzip des ‚Reflecting Teams‘ als kollegiales Reflexionssetting interessant. Hierbei wird durch strukturierende Vorgaben zum gemeinsamen Reflektieren auf Handlungs- und auf Metaebene die Eröffnung neuer Perspektiven auf das eigene Denken und Handeln angeregt (Andersen, 1990). Diese Methode wird zunehmend in der Aus- und Weiterbildung von Lehrpersonen eingesetzt (Kricke, Rohr & Schindler, 2012; Göbel & Neuber, angenommen).

#### Fragestellung

Der Beitrag stellt die Konzeption des Projekts „Videogestützte Unterrichtsreflexion im Praxissemester“ sowie die Ergebnisse für die erste Erhebungswelle vor. In den Blick genommen wird dabei, ob das angebotene Setting dazu beitragen kann, die allgemeine Bereitschaft zur Unterrichtsreflexion bei Praxissemesterstudierenden zu steigern.

#### Methode

In einem quasi-experimentellen Design haben innerhalb der ersten von vier Erhebungswellen 34 Studierende im Praxissemester der Universität Duisburg-Essen an der Studie teilgenommen, wobei die Experimentalgruppe (n=15) Gelegenheit zur systematisierten Reflexion von eigenen und fremden Unterrichtsvideos in Kleingruppen erhielt. Die Kontrollgruppe (n=19) erhielt keine vergleichbare Reflexionsgelegenheit. Der strukturelle Ablauf des Reflexionssettings erfolgte in Anlehnung an das systemische Beratungsgespräch des Reflecting Teams. Die inhaltliche Ausrichtung der Reflexion orientierte sich an einer zuvor selbstständig von den Studierenden gewählten Facette der Unterrichtsqualität (Helmke, 2009). Im Reflexionsprozess wurden die Studierenden durch Prompts geleitet und unterstützt, sodass sie die Eigen- und Fremdreiflexion weitgehend selbstständig durchführen konnten. Die Gruppenmitglieder sorgten innerhalb der ihnen zugewiesenen Rollen im Reflexionsprozess durch gezielte Interviewfragen und durch verbalisierte Analysen der Selbstreflexionsphase für Reflexionsanregung. Die allgemeine Reflexionsbereitschaft und der empfundene Nutzen spezifischer Reflexionsstrukturen wurden durch einen standardisierten Fragebogen vor und nach der Reflexionsphase im Praxissemester erhoben (Göbel & Neuber, 2016), zudem wurde eine Selbsteinschätzung zum zweiten Messzeitpunkt durch die Studierenden über die Veränderung ihrer Reflexionsbereitschaft und Fähigkeit zum zweiten Erhebungszeitpunkt erfasst.

#### Ergebnisse

Der strukturierte Ablauf der videogestützten Unterrichtsreflexion im Sinne des Reflecting Teams wird von den Studierenden nur zum Teil als hilfreich für die eigene Reflexion empfunden. Dennoch hat sich deren Einschätzung des generellen Nutzens von videogestützter Unterrichtsreflexion nicht verringert. Weiterhin zeigt sich, dass in der Experimentalgruppe insbesondere die Bereitschaft zur kollegialen Reflexion von Unterricht anstieg und die kritische Haltung gegenüber der Aufnahme, Sichtung und Reflexion eigener Unterrichtsvideos abnahm. Die Selbsteinschätzung zum zweiten Messzeitpunkt ergab zudem, dass sich die Studierenden der Experimentalgruppe im Vergleich zur Kontrollgruppe stärker motiviert fühlen, ihren eigenen Unterricht zukünftig systematisch zu reflektieren. Die Ergebnisse weisen auf den Nutzen, aber auch die Notwendigkeit der Weiterentwicklung des videogestützten Reflexionssettings hin.

### Lernumgebung Schulpraktikum – Förderung der Selbstwirksamkeitsüberzeugung von Lehramtsstudierenden durch Unterrichtsbesprechungen

**M.Sc. Daniela Rupp, Dr. Eva Becker, Prof. Dr. Fritz Staub**

Universität Zürich, Schweiz; [daniela.rupp@ife.uzh.ch](mailto:daniela.rupp@ife.uzh.ch)

Theoretischer Hintergrund: Schulpraktika haben das Ziel, die professionellen Handlungskompetenzen zu fördern, wobei Studierende im Schulpraktikum die Gelegenheit erhalten, erste Handlungserfahrungen in Ihrer zukünftigen Rolle als Lehrperson zu sammeln. Schulpraktika können daher auch eine wichtige Phase in der Entwicklung der lehrbezogenen Selbstwirksamkeit angehender Lehrpersonen darstellen (Flores, 2015; Gorski, Davis & Reiter, 2012), da sowohl eigene (Unterrichts-)Erfahrungen (mastery experiences) gemacht werden als auch stellvertretende Erfahrungen (vicarious experiences) durch Hospitationen und Beobachtungen stattfinden können und in Rahmen von Unterrichtsbesprechungen (verbal persuasion) das eigene Handeln reflektiert und rückgemeldet werden kann (Quellen der Selbstwirksamkeit nach Bandura, 1997). Die lehrbezogene Selbstwirksamkeitsüberzeugung ist nach dem Modell zur Entwicklung professioneller Kompetenzen von Kunter, Baumert & Blum (2011) dem Aspekt der motivationalen Orientierung zuzuordnen und verschiedene Studien konnten wichtige Zusammenhänge zur Unterrichtseffektivität (z.B. Dicke et al., 2014; Holzberger, Philipp & Kunter, 2013) als auch zur Gesundheit (z.B. Skaalvik & Skaalvik, 2010; Schwarzer & Hallum, 2008) finden. Die Befundlage zur Wirksamkeit von Schulpraktika ist jedoch inkonsistent (Hascher, 2006, 2011) und neben positiven Effekten (von Felten, 2005; König et al., 2012), gibt es auch Studien die kaum Belege für Lern- und Kompetenzzuwächsen durch das Schulpraktikum

zeigen (Hascher et al., 2004; Müller, 2010). Dies kann dadurch erklärt werden, dass die Art und der Umfang der Lernbegleitung von Lehramtsstudierenden (und damit auch die möglichen Quellen der Selbstwirksamkeitsentwicklung) während des Unterrichtspraktikums sowohl innerhalb als auch zwischen den Ausbildungsstandorten erheblich variieren (Arnold et al., 2010). Zudem fehlt es an Studien, welche die Entwicklung der Selbstwirksamkeitsüberzeugung und deren Wechselwirkung zur erhaltenen Lernbegleitung über einen Prä-Post Vergleich hinaus untersuchen (Klassen et al., 2011). Ziel der vorliegenden Studie ist es daher, die Rolle von eigenen Unterrichtserfahrungen (mastery experiences) sowie die Rolle von Unterrichtsbesprechungen, in denen sich die Studierenden mit einer erfahrenen Praktikumslehrperson austauschen (vicarious experiences, verbal persuasion), auf die Entwicklung der Selbstwirksamkeit systematisch über den Verlauf des Schulpraktikums zu untersuchen.

Fragestellungen: Die Studie geht der Fragen nach, ob das Schulpraktikum einen förderlichen Einfluss auf die Entwicklung der lehrbezogenen Selbstwirksamkeit hat (Prä-Post-Vergleich). Zudem werden intraindividuelle Entwicklungsverläufe analysiert, in denen die eigenen Unterrichtserfahrungen sowie die Lernbegleitung in Form von Unterrichtsbesprechungen gleichermaßen berücksichtigt werden (Prozessbefragungen).

Methode: Aktuell wurden 46 Studierende (MAlter = 27.98, 73.9% weiblich) an 39 Schulen in der Schweiz in ihrem Schulpraktikum begleitet (Ziel: 80-100 Studierende). Für eine Vergleichbarkeit zwischen den Praktika, führen alle Studierenden eine Unterrichtseinheit im Umfang von sechs Unterrichtsstunden zum Thema „schriftliches Argumentieren“ im Fach Deutsch in Klassen der Sekundarstufe und des Gymnasiums mit Schüler/innen der siebten bis neunten Jahrgangstufe durch (nSuS = 834, MAlter = 14.4, 51.3% weiblich). Dabei werden die Studierenden gebeten, die Durchführung der Unterrichtseinheit zu dokumentieren, indem für die Einzel/Doppelstunden die dazu allenfalls durchgeführten Vor- und/oder Nachbesprechungen mit der Praktikumslehrperson bezüglich Häufigkeit, Dauer und Qualität eingeschätzt werden. Zudem werden nach drei der sechs Unterrichtsstunden Beurteilungen zur Unterrichtsqualität (auch aus Sicht der Schüler/innen) abgegeben. Vor und nach der Unterrichtseinheit wird die Selbsteinschätzungen der Studierenden hinsichtlich der wahrgenommen lehrbezogenen Selbstwirksamkeit erfasst (Skalen nach Tschannen-Moran, Woolfolk & Hoy, 1998). Mit Hilfe von Latent-Change Modellen (berechnet mit Mplus, Version 7.11; Muthén & Muthén, 1998-2012), wird die Veränderung der Selbstwirksamkeit modelliert und mögliche Wechselwirkungen mit der Qualität der eigenen Unterrichtserfahrungen sowie den durchgeführten Unterrichtsbesprechungen analysiert.

(Erwartete) Ergebnisse: In den untersuchten Schulpraktika zeigte sich ein Zuwachs der Selbstwirksamkeit der Studierenden, insbesondere in Hinblick auf Instruktionsstrategien, Lernunterstützung der Schülerinnen und Schüler als auch Unterrichtsplanung. Zuwächse in der Selbstwirksamkeit im Bereich Klassenmanagement fielen hingegen kleiner aus. In einem weiteren Schritt werden die intraindividuellen Verläufe in Hinblick auf die erhaltene Lernbegleitung in Form von Unterrichtsbesprechungen sowie die (wahrgenommene) Qualität der durchgeführten Unterrichtsstunden analysiert. Es wird erwartet, dass der Kompetenzzuwachs bei einer intensiveren und qualitativ höherwertigen Lernbegleitung im Schulpraktikum stärker ausfällt.

## **Peer Learning im Praktikum. Unterstützung von Lernprozessen?**

**Minh-Ly Do, Lea De Zordo, Tina Hascher**

Universität Bern, Deutschland; [minh-ly.do@students.unibe.ch](mailto:minh-ly.do@students.unibe.ch)

### Theoretischer Hintergrund und Fragestellung

Dem an der Schulpraxis orientierten Berufsbild des Einzelkämpfers im Klassenzimmer ste-hen viele Befunde über die Bedeutsamkeit von Lehrerkooperationen gegenüber (vgl. Bonsen & Rolff, 2006; Gräsel et al., 2006; Terhart & Klieme, 2006). Eine spezielle Kooperationsform in der schulpraktischen Phase bildet das Teamteaching, in dem Studierende gemeinsam für den Unterricht verantwortlich sind. Während im anglo-amerikanischen Raum die Forschung um kooperative Peer-Lernprozesse auf eine lange Tradition zurückblicken kann, stellt das Teamteaching im Praktikum im deutschsprachigen Raum eine bis dato noch relativ junge Forschungsrichtung dar. Aktuelle Untersuchungen zeigen, dass Studierende der Beziehungskonstellation mit dem/der Praktikumpartner/in einen hohen Stellenwert zum Gelingen des Tandempraktikums beimessen (vgl. de Zordo/Hagenauer & Hascher 2017; de Zordo, Bisang & Hascher 2018). Dies hängt auch damit zusammen, dass Gleichaltrige Expertinnen und Experten für ihre Lebenswelt, ihre Probleme und ihre Lösungswege sind (vgl. Schröder 2003).

Es ist jedoch weitgehend offen, was im und durch Teamteaching gelernt wird. Hinzu kommt, dass individuelle Lernprozesse im Praktikum bisher unzureichend erforscht wurden (vgl. Hascher 2012). Dieses Forschungsdesiderat aufgreifend, wird in unserem Projekt untersucht, welche individuellen Lernprozesse sich im Praktikum vollziehen und wie Studierende als Peers mit-/voneinander lernen. Folgende Fragestellungen sind forschungsleitend: (1) Was verstehen Studierende unter Kooperation? (2) Was begünstigt kooperative Lernprozesse im Praktikum? (3) Welche Gewichtung kommt diesen Aspekten aus der Sicht der Studierenden zu?

### Methode

Lernprozesse können nicht unmittelbar beobachtet werden. Das diesem Projekt zugrunde-liegende qualitative Forschungsdesign ist explorativ ausgelegt. Auf Basis von teilstrukturier-ten leitfadengestützten Interviews wird untersucht, welche Lernprozesse sich in verschiedene-nen Praktikumssettings (Einzel- versus Tandempraktikum) ergeben. Hierbei erfolgt eine dif-ferenzierte Betrachtung zwischen Lernprozessen hinsichtlich der beiden Lernsettings der Peers als Lernende mit- oder voneinander.

Im Rahmen des Forschungsprojektes KiP („Kooperation im Praktikum“, ein gemeinsam ge-führtes Projekt der Universität Bern und der Pädagogischen Hochschule Bern) wird genauer untersucht, welche Lernprozesse sich in kooperativen Praktika vollziehen, da vier von fünf Praktika in der Ausbildung für die Lehrbefähigung auf der Vorschul- und Primarstufe im Tan-dem organisiert sind. Im Rahmen eines Mixed-Method-Designs wurden N = 30 Studierende zu ihren Erfahrungen aus dem Tandempraktikum interviewt (Herbst/Winter 2016/2017), da-von 20 Studierende ein halbes Jahr später (Frühjahr 2017) zu ihren Erfahrungen nach einem Einzelpraktikum. Die Interviews wurden inhaltsanalytisch nach Mayring (2008) ausgewertet. Die computergestützte Auswertung mit MAXQDA erfolgt zunächst deduktiv geleitet nach den drei Kooperationsformen (Austausch, arbeitsteilige Kooperation und Kokonstruktion) von Gräsel et. al (2006). Zur Beantwortung der o.g. Forschungsfragen wurden folgende Thesen formuliert: Um erfolgreich kooperieren zu können, bedarf es a) positiver Einstellungen zur Kooperation, b) einer positiven Kompetenzeinschätzung, c) einer positiven Beziehungsquali-tät unter den Kooperationspartner\_innen und d) positiver Erfahrungen im Prozess der Ko-operation.

### Ergebnisse und Diskussion

Erste Auswertungen der Interviews zeigen, dass die Studierenden das Tandempraktikum als Lerngelegenheit unter statusgleichen Peers wahrnehmen. Hierbei wird der Peer-Status explizit als positive Rollenkonstellation beim Co-Planning (vgl. Pratt et al. 2017) benannt, wohingegen die Studierenden bei der Unterrichtsdurchführung das Modell Assistenz (vgl. Baeten & Simon 2014) bevorzugen. Zukünftige Auswertungen richten sich auf die Frage, was kooperative Lernprozesse begünstigt. Hierzu werden im nachfolgenden Schritt verschiedene Phasen der Zusammenarbeit (Vorbereitung, Durchführung und Nachbereitung) differenzierter analysiert. Ob die Studierenden nach dem Tandempraktikum das Einzelpraktikum als Fortführung von Lernprozessen ansehen, wird Gegenstand nachfolgender Untersuchungen darstellen. Die Ergebnisse können Aufschluss über relevante Faktoren für lernwirksame Peer-Prozesse in Praktika geben. Nicht zuletzt können die Ergebnisse im Tandempraktikum als Anregung für die Gestaltung von Lehr-Lernprozessen für mögliche andere Praxisformate (z.B. Praxissemester) diskutiert werden.

## **Unterrichtsplanung im Praxissemester - zwischen Intuition, Wissen, Akteuren und Systembedingungen**

**Silvia Greiten**

Bergische Universität Wuppertal, Deutschland; [greiten@uni-wuppertal.de](mailto:greiten@uni-wuppertal.de)

Fragestellungen im Kontext des theoretischen Rahmens und Forschungsstandes

Unterrichtsplanung kann als mentales Handeln bezeichnet werden, das theoretisches Wissen voraussetzt (Bromme, 1981). Studien belegen, dass Lehrkräfte in ihrer Unterrichtsplanung meist kategorial (Bsp. Phasierung, Aufgaben) vorgehen (vgl. Gassmann, 2012). Vorgaben zum Praxissemester in NRW weisen dem Unterrichten eine große Bedeutung zu und Ansprüche an den Unterricht für heterogene Lerngruppen bringen Studierende in komplexe Planungsprozesse. Die Kooperation mit Akteuren aus Universität, ZfSLs und Schulen soll die Theorie-Praxis-Verbindung unterstützen. Bislang ist zu wenig erforscht, inwiefern Lehramtsstudierende Kategorien der Unterrichtsplanung nutzen und theoretisch fundiert verwenden. Die Studie untersucht drei Forschungsfragen: Wie gehen Studierende nach eigenen Angaben im Praxissemester bei der Unterrichtsplanung vor? Auf welches Wissen greifen Studierende im Praxissemester bei der Unterrichtsplanung zurück? Wie begründen Studierende im Praxissemester ihre Unterrichtsplanung?

Methodik

Zur Untersuchung mit Interviews wurde eine Stichprobe von 10 Studierenden aus unterschiedlichen Fachkombinationen im Master der Sekundarstufen am Ende des Praxissemesters gebildet. Alle haben Vorbereitungsseminare an der Universität und mindestens 60 Unterrichtsstunden und drei Unterrichtsbesuche durch das ZfSL absolviert. Der Interviewleitfaden umfasste die Themenfelder Beschreibung der eigenen Unterrichtsplanung, Nutzung des Wissens zu didaktischen Modellen und aus Bildungswissenschaften, begleitende Akteure und Rahmenvorgaben zum Praxissemester in NRW. Die Theorie-Praxis-Verbindung und individuelle Argumentationen wurden sowohl durch offene Fragen als auch konkrete Fragen nach der Anwendung theoretischen Wissens untersucht. Die methodische Auswertung erfolgt mit der qualitativen Inhaltsanalyse und dem Schwerpunkt der inhaltlichen Strukturierung mit deduktiv und induktiv gebildeten Kategorien; unter Nutzung der Software MAXQDA werden Typenbildungen vorgenommen (Kuckartz, 2016, 97, 152).

Ergebnisse

Innerhalb der Stichprobe wurden im analytischen Prozess kontrastierende Typiken zur Unterrichtsplanung im Praxissemester herausgearbeitet: intuitiv ohne Theorie-Praxis-Verbindung versus phasenstrukturiert mit Begriffswissen. Dazwischen finden sich Mischformen. Alle Interviews verweisen darauf, dass die Unterrichtsplanung im Praxissemester maßgeblich durch die Akteure bestimmt wird, vorrangig durch das ZfSL und die Intuition der Studierenden, nachrangig durch Mentoren. Viele Studierende greifen nur auf im Studium erworbenes Wissen zurück, andere nutzen Wissen mehr auf der Begriffsebene. Ihre Begründungen bewegen sich zwischen einem technokratischen Modell in Form von abzuarbeitenden Kriterien guten Unterrichts und standardisierten Phasenrastern zum Aufbau von Unterrichtsstunden. Die Befunde provozieren Fragen, welches Wissen zur Unterrichtsplanung die beteiligten Akteure anbieten und welches implizite Wissen der Akteure, vor allem der Studierenden, in die Ausbildungspraxis einwirkt (Greiten, 2017 i.D.).

# A11

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus, Aula 33

## Konstruktion von Tests und Testaufgaben zur Erfassung von Lernzuwächsen

Chair(s): Johannes Hartig (DIPF)

DiskutantIn(nen): Marcus Hasselhorn (DIPF)

Für die Schulleistungsdiagnostik sowohl auf Individualebene als auch auf der Ebene von Lerngruppen stellt die reliable und valide Erfassung von Lernzuwächsen eine wichtige Anforderung dar. Individuelle Lernverlaufsdagnostik kann sichtbar machen, inwieweit einzelne Schülerinnen und Schüler von gleichem Unterricht in unterschiedlichem Maße profitieren und hierdurch dabei helfen, mit angemessenen adaptiven Förderkonzepten zu reagieren. Darüber hinaus dienen Testergebnisse und die Veränderungen in Testergebnissen von Schülerinnen und Schülern häufig als zentrales Kriterium für die Beurteilung der Effektivität von schulischem Unterricht. Diese Verwendung von Testergebnissen setzt voraus, dass Testinstrumente die durch den Unterricht verursachten Lernzuwächse auch tatsächlich abbilden, also instruktionssensitiv sind.

Inwieweit Testitems und Tests tatsächlich geeignet sind, Zuwächse in den erfassten Fähigkeiten auf Individual- und Klassenebene abzubilden, ist jedoch häufig unklar. Gängige Qualitätskriterien bei der Testentwicklung wie interne Konsistenz von Skalen und Item-Test-Korrelationen basieren auf Querschnittsdaten und berücksichtigen die Zugehörigkeit zu Lerngruppen in der Regel nicht. Nach diesen Kriterien optimierte Tests müssen weder optimal für die Erfassung von Lernzuwächsen noch für die Erfassung von Unterrichtseffekten sein. Die Beiträge dieses Symposium befassen sich mit der Konstruktion und Evaluation von Messinstrumenten, die Lernzuwächse auf Individual- und Klassenebene abbilden sollen. Hierbei werden verschiedene Methoden und Kriterien veranschaulicht, mit denen die Erreichung dieses Anspruchs evaluiert werden kann.

Der erste Beitrag von NN et al. fokussiert die Erfassung von individuellen Lernverläufen anhand eines Mathematiktests für Primarschüler. Es wird der Einfluss verschiedener Scoring-Methoden sowie von verschiedenen Itemmerkmalen auf die Änderungssensibilität der Testergebnisse untersucht. Im zweiten Beitrag NN et al. wird anhand einer Testreihe zur Lernverlaufsdagnostik im Lesen untersucht, ob sich Facetten der Instruktionssensitivität durch Itemmerkmale vorhersagen lassen, die die Schwierigkeit der Items im Querschnitt beeinflussen. Der dritte und vierte Beitrag bauen aufeinander auf, hier ist die Instruktionssensitivität eines Mathematiktests für die Schweizer Primarstufe Untersuchungsgegenstand. Der Beitrag NN et al. untersucht, inwieweit Lehrkräfte und Fachdidaktiker/innen anhand verschiedener Kriterien in der Lage sind, die vermutete Instruktionssensitivität von Testitems konsistent einzuschätzen. In den von NN et al. vorgestellten Analysen wird geprüft, ob die von den Experten/innen vorgenommenen Urteile geeignet sind, Facetten der Instruktionssensitivität, die anhand von längsschnitlichen Testdaten geschätzt wurden, auf Itemebene vorherzusagen. Die vier Beiträge werden abschließend von NN diskutiert.

### Beiträge des Symposiums

#### Auf Goldmünzenjagd: Psychometrische Kennwerte verschiedener Scoringansätze bei computergestützter Lernverlaufsdagnostik im Bereich Mathematik

**Christin Schwenk<sup>1</sup>, Jörg-Tobias Kuhn<sup>1</sup>, Daniela Gühne<sup>2</sup>, Philipp Doeblner<sup>2</sup>, Heinz Holling<sup>1</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, <sup>2</sup>Technische Universität Dortmund

Im Beitrag wird ein computergestützter Lernverlaufstest („Goldmünzenjagd“) für Grundschul Kinder mit Rechenschwierigkeiten vorgestellt. Der Test bildet die Lernentwicklung in zwei grundlegenden mathematischen Basiskompetenzen ab: dem arithmetischen Faktenwissen (Addition bzw. Subtraktion bis 20) und der ordinalen Zahlenverarbeitung (Bewertung der Ordnung von Zahlenreihen mit drei zweistelligen Elementen). Der Test wurde nach dem robust indicator-Ansatz konstruiert und basiert damit auf elementaren Kompetenzen, die sich empirisch als valide Prädiktoren für die arithmetische Leistungsentwicklung erwiesen haben. Damit wird eine fortlaufende Differenzierung vor allem im unteren Leistungsbereich ermöglicht.

Im Beitrag werden die psychometrischen Kennwerte sowie die Änderungssensibilität des Tests anhand von drei geläufigen Scoringvarianten verglichen: Geschwindigkeit (Anzahl bearbeiteter Items), Präzision (Anteil korrekt bearbeiteter Items) und Effizienz. Die Effizienz wird hierbei über ein High speed, high stakes-Scoring abgebildet, das Geschwindigkeit und Präzision zu einem Gewinn oder Verlust von Münzen auf Itemebene verknüpft, wobei schnelle Antworten stärker gewichtet werden als langsamere. Die Vergleiche der Scorings erfolgten sowohl für die drei Untertests (Addition, Subtraktion, Zahlenordnen) als auch für den Gesamttestscore.

In einer Feldstudie (N = 241 Grundschul Kinder, Klassenstufe 2 bis 4) wurde der Lernverlaufstest alle 5 Tage während einer 30-tägigen computergestützten Mathematikintervention administriert. Vor (Prätest) und nach (Posttest) dieser Intervention wurde der statusdiagnostische Mathematiktest CODY-M 2-4 eingesetzt. Das Münzscoring zeigte sich psychometrisch, sowohl in der Reliabilität ( $r = .87-.93$ ) als auch in der Kriteriumsvalidität ( $r = .51$ ), den klassischen Geschwindigkeits- und Präzisions-Scorings überlegen. Für alle drei Scorings im Lernverlaufstest wurden zur Prüfung der Änderungssensibilität lineare hierarchische Regressionen angepasst, um das individuelle Ausgangsniveau (random intercept) und die Lernentwicklung (random slope) zu modellieren. Varianz im statusdiagnostischen Posttest konnte durch diese Parameter individueller Lernverläufe auch nach Kontrolle der Prätestwerte aufgeklärt werden. Ein zusätzlicher Analyseschritt beleuchtet zudem, inwieweit die Änderungssensibilität der drei Untertests von „Goldmünzenjagd“ durch die Itemkomplexität moderiert wird (Addition bzw. Subtraktion: Items mit vs. Items ohne Zehnerübergang, Zahlenordnen: kongruente vs. inkongruente Items).

Der vorgestellte Lernverlaufstest eignet sich zur formativen Evaluation der Leistungsentwicklung von Grundschulkindern in basalen mathematischen Kompetenzbereichen sowie im unteren Leistungsspektrum. Insbesondere für rechenschwache Kinder bietet das Goldmünzen-Scoring eine direkt ersichtliche Anreizstruktur, die schlechter Performanz aufgrund von Motivationsdefiziten vorbeugen kann sowie die Entwicklung von zählenden hin zu abrufbasierten Rechenstrategien fördert.

#### Globale und differenzielle Sensitivität von Lesetests zur Lernverlaufsdagnostik und der Einfluss schwierigkeitsgenerierender Aufgabenmerkmale

**Natalie Förster<sup>1</sup>, Naumann Alexander<sup>2</sup>, Elmar Souvignier<sup>1</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, <sup>2</sup>DIPF

Um Leistungsveränderungen über die Zeit zu erfassen, müssen die verwendeten Messinstrumente sensitiv für derartige Veränderungen sein. Dies gilt umso mehr, wenn Leistungsentwicklungen in kurzen zeitlichen Abständen im Sinne eines formativen Assessments gemessen werden und Rückschlüsse über den Unterricht geben sollen. Insofern stellt die



Instruktionssensitivität (IS), d.h. das Ausmaß, in dem der Test sensitiv auf Instruktion reagiert (Polikoff, 2010), eine zentrale psychometrische Eigenschaft von Tests zur Lernverlaufdiagnostik dar. Naumann et al. (2014) unterscheiden zwischen globaler und differenzieller IS, wobei sich die globale IS anhand der durchschnittlichen und die differenzielle IS anhand der klassenspezifischen Veränderungen der Itemschwierigkeiten in einem längsschnittlichen Mehrebenen-DIF Modell bestimmen lassen. Eine aktuelle Weiterentwicklung des Modells von Naumann et al. (2017) zu einem längsschnittlichen Mehrebenen-IRT Modell (LMLIRT) bietet die Grundlage, darüber hinaus zu analysieren, inwiefern bestimmte Itemmerkmale die globale und differenzielle IS beeinflussen und somit Ansatzpunkte für die gezielte Entwicklung instruktionssensitiver Items zu identifizieren (z.B. Hochweber et al., 2017).

Im Rahmen eines DFG-Projektes wurde eine Testreihe zur computerbasierten Lernverlaufdiagnostik im Lesen bei Zweitklässlern entwickelt, wobei gezielt schwierigkeitsgenerierende Merkmale bei der Testkonstruktion berücksichtigt wurden. In der vorliegenden Studie wird nun untersucht, inwiefern sich diese neu entwickelte Testreihe als global und als differenziell sensitiv erweist. Darüber hinaus wird geprüft, ob sich die globale und die differenzielle IS durch die bei der Testkonstruktion berücksichtigten Itemmerkmale erklären lassen.

Unter Berücksichtigung schwierigkeitsgenerierender Merkmale (z. B. Worthäufigkeit, Satzstruktur) wurden vier äquivalente Lesetests entwickelt, die die Effizienz der Teilprozesse des Lesens auf Wort-, Satz- und Textebene erfassen. N = 1917 Zweitklässler aus 104 Klassen bearbeiteten in Abständen von jeweils ca. drei Wochen diese vier Lesetests insgesamt zweimal in der Abfolge A-B-C-D-A-B-C-D am Computer. Erfasst wurde neben der Akkuratheit der Antworten jeweils die Antwortlatenz. Die Analyse der globalen und differenziellen Sensitivität der Items erfolgte mittels des längsschnittlichen Mehrebenen-IRT Modells (LMLIRT; Naumann, Hartig & Hochweber, 2017). Um den Einfluss der Aufgabenmerkmale auf die IS der Items zu untersuchen, reparametrisierten wir das LMLIRT Modell analog zum linear-logistischen Testmodell mit Fehlerterm (LLTMe). Der Einfluss der Aufgabenmerkmale auf die Veränderung der Itemschwierigkeiten über die Zeit gibt dabei Effekte auf die globale Sensitivität an, während die Varianz des Einflusses der Merkmale in den verschiedenen Klassen Auskunft über Effekte auf die differenzielle Sensitivität gibt.

Bislang liegen die Ergebnisse für die Akkuratheit vor. Mit mittleren Itemschwierigkeitsveränderungen von -0.17 und -0.25 erwiesen sich die Wort- und Text-Items hinsichtlich der Akkuratheit als global sensitiv, während die Satz-Items im Verlauf des Schuljahres im Mittel nicht leichter wurden. Die Unterschiede in der Veränderung der Itemschwierigkeiten zwischen den Klassen und somit die differenzielle Sensitivität der Items war vergleichsweise gering ( $M = .11$ ,  $SD = .07$ ). Es zeigte sich ein Trend dahingehend, dass die differenzielle IS umso höher ausfiel, je komplexer die Leseanforderung war. Die Analyse mittels des LLTMe-basierten Modells verdeutlichte, dass die Aufgabenmerkmale der Wort-, Satz- und Textitems zwar prädiktiv für die Itemschwierigkeit zum ersten Messzeitpunkt waren, jedoch nicht im Zusammenhang mit den globalen oder klassenspezifischen Veränderungen der Itemschwierigkeiten zu den späteren Messzeitpunkten stehen. Entsprechend niedrig war auch der Zusammenhang zwischen der vorhergesagten und im Vergleichsmodell geschätzten Veränderung der Itemschwierigkeiten ( $r = .42$ ).

Insgesamt zeigt sich, dass die Lesetestreihe global und differenziell sensitiv für Leistungsveränderungen ist. Ein systematischer Einfluss der Aufgabenmerkmale auf die IS konnte allerdings nicht gefunden werden, worin ein Unterschied zu Befunden von Hochweber et al. (2017) liegt, die einen substantiellen Anteil globaler und differenzieller Sensitivität von Items anhand von Itemmerkmalen erklären konnten. Eine mögliche Begründung ist, dass sowohl die Aufgabenformate als auch die Itemmerkmale bei Hochweber et al. (2017) deutlich komplexer waren als in der vorliegenden Studie.

## Expertenratings – Ein Verfahrenvergleich zur Evaluation der Instruktionssensitivität von Testitems

**Stephanie Musow<sup>1</sup>, Alexander Naumann<sup>2</sup>, Johannes Hartig<sup>2</sup>, Jan Hochweber<sup>1</sup>**

<sup>1</sup>PH St. Gallen, <sup>2</sup>DIPF

### Theoretischer Hintergrund

Das Gütekriterium der Instruktionssensitivität erfasst, inwieweit ein Test bzw. eine Testaufgabe in der Lage ist, Effekte von Unterricht – der Unterrichtsinhalte und Unterrichtsqualität – auf Schülerleistungen abzubilden (Polikoff, 2010). Ist die Instruktionssensitivität eines Tests nicht gewährleistet, kann dies zu fehlerhaften Schlussfolgerungen auf Eigenschaften des Unterrichts führen (Popham et al., 2014). Um die Instruktionssensitivität zu überprüfen, stehen drei Verfahrenswesen zur Verfügung (Polikoff, 2010): (1) Verfahren, welche auf reinen Testdaten basieren, (2) Verfahren, welche sich auf Testdaten unter Hinzuziehung von Unterrichtsvariablen gründen und (3) Expertenratings.

Während es zu den beiden erstgenannten Verfahrenswesen zahlreichen Studien gibt (z.B. Ing, 2017; Naumann, Hochweber, & Klieme, 2016), kamen Expertenratings zur Evaluation der Instruktionssensitivität bislang kaum zur Anwendung. Dabei dürften Expertenratings im Vergleich zu testdatenbasierten Verfahren deutlich ökonomischer anzuwenden sein (siehe Popham, 2007). Die bislang einzige empirische Studie zu Expertenratings erfolgte durch Chen (2012). Die Autorin ließ US-Lehrpersonen Mathematikaufgaben auf einer 10-stufigen Likert-Skala beurteilen. Die Ergebnisse des Ratings waren jedoch ernüchternd, da sich mitunter sehr niedrige Werte in den Interraterreliabilitäten zeigten.

Ziel des vorliegenden Beitrags ist es, die Bedingungen für Expertenurteile zur Instruktionssensitivität von Items tiefergehend zu untersuchen. Wir untersuchen, (a) inwiefern sich die Befunde aus Chens (2012) Globalrating auf Expertenratings durch schweizerische Lehrpersonen übertragen lassen, (b) inwiefern sich differenzierte Ratings zur Instruktionssensitivität von Testitems (basierend auf mehreren Kriterien) von einem Globalrating unterscheiden, und (c) ob sich Lehrpersonen und / oder Fachdidaktiker/innen als Expert/innen eignen.

### Methode

Die Analysen erfolgen mit Daten von 18 Lehrpersonen (LP) und sechs Fachdidaktiker/innen (FD) aus der deutschsprachigen Schweiz, die 32 Arithmetikaufgaben aus einem schweizerischen Mathematiktest für die Jahrgangsstufe 5 beurteilten. Neun LP und drei FD nahmen am globalen Rating teil, während die anderen neun LP und drei FD am differenzierten Rating partizipierten. Gemäß dem Rating nach Chen (2012) wurden die Expert/innen um eine globale Einschätzung zur Instruktionssensitivität gebeten. Beim differenzierten Urteil erfolgte das Rating vierstufig anhand von vier Kriterien hinsichtlich verschiedener Aspekte von Instruktionssensitivität: die Bedeutung von a) Unterrichtsinhalt und b) Unterrichtsqualität für die Lösung eines Items, sowie c) die Eignung eines Items zur Erfassung von Lernfortschritten der Schüler/innen und d) die Vertrautheit der Schüler/innen der Jahrgangsstufe 5 mit der Aufgabe (Beispiel für Unterrichtsinhalt: „Die Schüler/innen können die Aufgabe erst lösen, nachdem das Thema unterrichtet wurde“). Die Ratings innerhalb jeder Bedingung (global vs. differenziert; FD vs. LP) werden auf ihre Übereinstimmung hin überprüft und die jeweiligen Einschätzungen zur Instruktionssensitivität zwischen den beiden Ratingverfahren und unter Berücksichtigung der Expert/innen-Gruppen miteinander verglichen. Berichtet werden die ICC1 (Beurteilerübereinstimmung), die ICC2 (Reliabilität des aggregierten Urteils

einer Gruppe von Ratern) als auch Mittelwerte (M) und Standardabweichungen (SD) bezogen auf die Einschätzung zur Instruktionssensitivität von Testitems.

#### Ergebnisse

Die Beurteilerübereinstimmung fällt bei allen Lehrpersonen-Ratings mit einer ICC1 von  $<.35$  gering aus. Der Vergleich zwischen dem globalen und dem differenzierten Rating der Lehrpersonen lässt erkennen, dass (mit einer Ausnahme) die ICC2-Werte zum globalen Rating niedriger ausfallen (ICC2=.66) als zum differenzierten Rating (Unterrichtsinhalt: ICC2=.75; Unterrichtsqualität: ICC2=.64; Lernzuwachs: ICC2=.72; Vertrautheit: ICC2=.71). Das differenzierte Rating ist damit leicht im Vorteil. Auf Basis der aggregierten Urteile geht hervor, dass die Items insgesamt als (eher) instruktionssensitiv eingestuft werden (globales Rating:  $M=7.33$ ,  $SD=1.17$ ; differenziertes Rating zu Unterrichtsinhalt:  $M=2.99$ ,  $SD=.55$ ; zur Unterrichtsqualität:  $M=2.65$ ,  $SD=.40$ ; zum Lernzuwachs:  $M=3.21$ ,  $SD=.52$ ; zur Vertrautheit:  $M=3.21$ ,  $SD=.57$ ).

Abschließend kann konstatiert werden, dass eine Durchführung von Expertenratings mit wenigen Lehrpersonen fragwürdig erscheint, die aggregierten Ratings jedoch eine zufriedenstellende Messgenauigkeit aufweisen. Weiterführende Untersuchungen mit testdatenbasierten Verfahren sind notwendig, um die Ergebnisse aus den Expertenratings mit den Ergebnissen aus den statistischen Analysen zu vergleichen. Darüber hinaus werden im Beitrag die Resultate zu den Einschätzungen der Fachdidaktiker/innen präsentiert, welche bis Mitte Oktober vollständig vorliegen werden.

### Messung der Instruktionssensitivität von Items: Übereinstimmung von Expertenurteilen und statistischen Indikatoren

**Alexander Naumann<sup>1</sup>, Stephanie Musow<sup>2</sup>, Jan Hochweber<sup>2</sup>, Johannes Hartig<sup>1</sup>**  
<sup>1</sup>DIPF, <sup>2</sup>PH St. Gallen

#### Theoretischer Hintergrund

Ergebnisse von Schülerinnen und Schülern in curriculums- oder standardbasierten Leistungsmessungen dienen regelmäßig als Grundlage für Rückschlüsse über die Effektivität von Unterricht (Hartig, Klieme, & Leutner, 2008). Gültige Rückschlüsse über die Effektivität von Unterricht anhand der individuellen Schülerantworten erfordern jedoch, dass die eingesetzten Instrumente potentiell dazu in der Lage sind, Effekte des Unterrichts aufzufangen. Polikoff (2010) bezeichnet diese psychometrische Eigenschaft eines Tests oder einzelnen Items als Instruktionssensitivität.

Zentral für die Messung von Instruktionssensitivität ist die Beobachtung von Veränderung im Antwortverhalten von Schülerinnen und Schülern in Abhängigkeit vom erlebten Unterricht (Burstein, 1989). Ansätze zur Messung von Instruktionssensitivität greifen auf drei Datenquellen zurück (Polikoff, 2010): (a) empirische Testdaten (Testwerte oder Itemantworten), (b) empirische Maße über den Inhalt und die Qualität des implementierten Unterrichts, oder (c) Expertenurteile über einzelne Items. Allerdings ist die Gültigkeit von Expertenurteilen über Instruktionssensitivität bisher wenig untersucht. Expertenaussagen erscheinen in der Untersuchung von Instruktionssensitivität vorteilhaft, da sie im Vergleich zu auf Itemantworten basierenden Verfahren einen wesentlich geringeren Aufwand erfordern. Ziel des vorliegenden Beitrags ist daher die Untersuchung, ob die von Expertinnen und Experten vorgenommenen Urteile geeignet sind, Facetten der Instruktionssensitivität, die anhand von längsschnittlichen Testdaten geschätzt wurden, auf Itemebene vorherzusagen.

#### Methode

Für unsere Untersuchung nutzen wir Daten zweier Messzeitpunkte aus einer längsschnittlichen Leistungsmessung mit fünfzehn Mathematikitems aus dem Klassenscockpit-Test. Die Erhebungen wurden im Schuljahr 2016/17 bei Schweizer 791 Grundschüler/-innen aus 47 Klassen der Jahrgangsstufe 5 im Kanton St. Gallen durchgeführt. Wir bestimmen die Instruktionssensitivität der Klassenscockpit-Items (a) basierend auf Testdaten mittels des längsschnittlichen Mehrebenen-IRT Modells (LMLIRT Modell; Naumann, Hartig, & Hochweber, 2017) sowie (b) basierend auf Expertenurteilen.

Im LMLIRT Modell wird die klassenspezifische Veränderung der Itemschwierigkeit geschätzt, wobei Mittelwert und Varianz der Itemschwierigkeitsveränderung als Maße für die Instruktionssensitivität eines Items dienen, nämlich für a) globale Sensitivität, die angibt, inwiefern sich die Itemschwierigkeit im Mittel über Klassen verändert, und b) differenzielle Sensitivität, die angibt, inwiefern die Schwierigkeitsveränderung über Klassen hinweg variiert (Naumann, Hochweber, & Hartig, 2014). Die Instruktionssensitivitätsmaße aus dem LMLIRT Modell werden hinsichtlich ihrer Übereinstimmung mit einem globalen Rating sowie einem differenzierten Rating zur Instruktionssensitivität der Items verglichen.

Globale und differenzierte Ratings lagen von jeweils neun Lehrpersonen vor. Beim globalen Rating erfolgte die Beurteilung in Anlehnung an Chen (2012) zehnstufig auf einem einzelnen Kriterium („Wie instruktionssensitiv ist die Aufgabe?“). Beim differenzierten Urteil erfolgte das Rating vierstufig anhand von vier Kriterien, die sich auf verschiedene Aspekte von Instruktionssensitivität beziehen, die Bedeutung von a) Unterrichtsinhalt und b) Unterrichtsqualität für die Lösung eines Items, sowie c) die Eignung eines Items zur Erfassung von Lernfortschritten der Schüler/-innen und d) die Passung der Aufgabe zur Jahrgangsstufe 5. Zur Überprüfung der Übereinstimmung mit den statistischen Instruktionssensitivitätsmaßen wurden jeweils Mittelwerte über die Einschätzungen der Expert/-innen gebildet.

#### Ergebnisse

Statistischen Indikatoren aus dem LMLIRT Modell zufolge reichte die mittlere Itemschwierigkeitsveränderung von  $-.88$  bis  $.07$ , wobei neun der fünfzehn Items global sensitiv waren mit dem Wert 0 außerhalb des 95%-Kreditabilitätsintervalls. Die Varianz der Itemschwierigkeitsveränderung über Gruppen reichte von  $0.14$  bis  $1.53$ , wobei vier Items nicht differenziell sensitiv waren. Weder das globale Expertenurteil noch die einzelnen Kriterien des differenzierten Expertenurteils korrelierten mit dem Ausmaß globaler und differenzieller Itemsensitivität (alle  $p > .20$ ). Auch der Mittelwert über die vier Kriterien des differenzierten Ratings zeigte keinen Zusammenhang mit den statistischen Indikatoren (jeweils  $p > .70$ ).

Zusammenfassend legen die Ergebnisse zur Übereinstimmung nahe, dass die von Lehrpersonen als Expert/-innen vorgenommenen Urteile nicht geeignet scheinen, Facetten der Instruktionssensitivität, die anhand von längsschnittlichen Testdaten geschätzt wurden, auf Itemebene vorherzusagen. Zwar stellen Expertenurteile ein vergleichsweise ökonomisches Verfahren dar, offenbar können sie jedoch die Analyse der Instruktionssensitivität anhand von Testdaten nicht ersetzen.

## A12

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H119

### **KREATIVE POTENZIALE ERKENNEN UND FÖRDERN IM KINDES- UND JUGENDALTER – Empirische Zugänge und Implikationen für die schulische und außerschulische Kulturelle Bildung**

Chair(s): **Nicole Berner** (Pädagogische Hochschule Fachhochschule Nordwestschweiz, Schweiz), **Wida Rogh** (Pädagogische Hochschule Fachhochschule Nordwestschweiz, Schweiz)

DiskutantIn(nen): **Dagmar Widorski** (Pädagogische Hochschule Fachhochschule Nordwestschweiz, Schweiz)

Kreativität kann als Fähigkeit verstanden werden auf zielgerichtete Art und Weise, neuartige Ideen oder Produkte hervorzubringen (Drevdahl, 1956). Dabei erfordert Kreativität divergentes und konvergentes Denken, Fachwissen, fachspezifische Fertigkeiten und Motivation. Zudem stellen Offenheit für neue Erfahrungen, Risikobereitschaft und Konflikttoleranz wichtige individuelle Voraussetzungen für kreatives Verhalten dar (Berner, Theurer & Lipowsky, 2012; Urban, 2004). Der kreative Prozess zeichnet sich durch selbstständiges und ausdauerndes Arbeiten aus und erfordert den Umgang mit Problemlösestrategien.

Damit erweist sich Kreativität als eine Kernkompetenz für die einzelne Person, um individuelle Lernprozesse zu stärken sowie für die moderne Gesellschaft, um diese mit innovativen Errungenschaften voranzutreiben (Runco, 2004). Es resultiert die Anforderung an die pädagogische Praxis, im schulischen und außerschulischen Bereich kreative Potenziale zu erkennen und zu fördern.

In diesem Kontext können insbesondere kulturelle Bildungsangebote einen wichtigen Beitrag zur Kreativitätsförderung leisten (Bielenberg, 2006; Rogh, Berner & Theurer, 2017). Dies wird sowohl in politischen Positionen (EDK, 2016; KMK, 2013) als auch in wissenschaftlichen Beiträgen (u.a. Winner, Goldstein & Vincent-Lancrin, 2013) betont. Empirische Studien zeigen beispielsweise, dass die Kreativitätsentwicklung von Heranwachsenden durch kulturell-ästhetische Angebote in der Schule oder durch kulturelle Freizeitkurse begünstigt werden kann (RfKB, 2017). Die Abbildung der Effekte wird von der Komplexität des Kreativitätskonstrukts und entsprechender Verfahren zur Messung der Kreativitätsentwicklung herausgefordert. Diesen Herausforderungen widmet sich die Kreativitätsforschung, die sich unter anderem damit auseinandersetzt, wie kreative Potenziale im Kindes- und Jugendalter erkannt und gefördert werden können.

Das Verstehen solcher Lern- und Entwicklungsprozesse ist für das professionelle Handeln in der pädagogischen Praxis von besonderer Relevanz, um die Entwicklung entsprechender Unterrichts- und Förderkonzepte zu ermöglichen (Kirchner & Kirschenmann, 2015; Schulz, 2009; Wichelhaus, 2006) und damit eine hohe Qualität künstlerischer und kultureller Bildungsangebote sicherzustellen.

Das Symposium möchte hierzu einen Beitrag leisten. Anhand verschiedener Positionen wird gefragt: Wie kann Kreativität erfasst werden? Wie entwickelt sich Kreativität im Kindes- und Jugendalter? Wie wirken sich schulische und außerschulische kulturelle Angebote auf Kreativität aus? Welche Implikationen ergeben sich für das professionelle Handeln im künstlerischen und kulturellen Bildungskontext?

Die Forschungsgruppen widmen sich diesen Fragen mit unterschiedlichen thematischen und methodischen Zugängen: Rogh et al. bilden die domänenunspezifische Kreativitätsentwicklung im fünften Schuljahr ab, während sich Pürgstaller et al. mit der Erfassung der domänenspezifischen Kreativität im Bereich Tanz- und Bewegung auseinandersetzen. Theurer et al. untersuchen, wie sich kreativitätsförderndes Verhalten von Lehrpersonen im Kunstunterricht auf die bildnerische Schülerkreativität auswirkt. Knoll analysiert, inwieweit kulturelle Angebote durch ein kreatives Klima gekennzeichnet sind.

Die vorgestellten Studien werden am Ende des Symposiums zusammenfassend diskutiert.

#### *Beiträge des Symposiums*

### **Kreativitätsentwicklung im fünften Schuljahr**

**Wida Rogh<sup>1</sup>, Nicole Berner<sup>1</sup>, Caroline Theurer<sup>2</sup>, Frank Lipowsky<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Fachhochschule Nordwestschweiz, Schweiz, <sup>2</sup>Universität Kassel, Deutschland

Die Förderung des kreativen Potenzials von Schülerinnen und Schülern gehört zu den aktuellen Aufgaben, die von der Gesellschaft an Schule herangetragen werden. Diese Forderung zur Kreativitätsförderung im schulischen Kontext wird dadurch bestärkt, dass klassen- und schulspezifische Merkmale die Kreativitätsentwicklung von Heranwachsenden beeinflussen können. Empirische Studien zeigen, dass unter anderem die Schulklassenzugehörigkeit, der Schultyp oder das Klassenklima die Schülerkreativität bestimmen können (u.a. Beghetto & Kaufman, 2014; Besançon & Lubart, 2008; Heise, Böhme & Körner, 2010; Theurer, 2014). Zur gezielten Förderung von Kreativität und differenzierten Bestimmung kreativitätsförderlicher Merkmale innerhalb von Schulklassen und Schulen, ist es wichtig zu wissen, wie sich die Kreativität von Schülerinnen und Schülern entwickelt. Bisherige Studien zeigen, dass die Kreativitätsentwicklung im Kindes- und Jugendalter diskontinuierlich verläuft (zsf. Theurer, 2014) und Entwicklungsverläufe im Laufe der Schulzeit durch mehrere Einbrüche gekennzeichnet sind (Camp, 1994; Charles & Runco, 2001; Smith & Carlsson, 1990; Torrance, 1968). Darüber hinaus zeigt sich, dass die abgebildete Kreativitätsentwicklung je nach eingesetzten Messinstrumenten oder erfassten Kreativitätsfacetten variiert (Runco, 1999).

Das Forschungsprojekt KuBiK5 knüpft an diese Ausgangslage an und nimmt verschiedene domänenunspezifische Merkmale von Kreativität in den Blick, um der Frage nachzugehen, wie sich verschiedene Kreativitätsfacetten über das fünfte Schuljahr entwickeln.

Zur Beantwortung der Fragestellung werden Daten aus Befragungen im Herbst 2015 und Herbst 2016 herangezogen. Rund 1.000 Schülerinnen und Schüler aus 54 Schulklassen in Nordrhein-Westfalen, Hessen und Sachsen wurden zu Beginn des fünften und sechsten Schuljahres unter anderem zu ihrer Kreativität befragt. Die Schülerkreativität wurde zu beiden Erhebungszeitpunkten mit zwei Verfahren getestet. Neben der Testbatterie zum Divergenten Denken von Berner und Kollegen (2012) haben die Kinder den Test zum schöpferischen Denken – Zeichnerisch (TSD-Z) von Urban und Jellen (1995) bearbeitet.

Die Testbatterie zum Divergenten Denken fokussiert die kognitive Ebene von Kreativität und gibt Aufschluss über das flüssige und flexible Denken im verbalen und figuralen Bereich. Die Schülerinnen und Schüler mussten sechs Schreibaufgaben und zwei Zeichenaufgaben lösen. Für jede Testaufgabe hatten sie nur begrenzt Zeit zur Verfügung. Die Testaufgaben wurden dann von unabhängigen Ratern nach der Anzahl der genannten Ideen (flüssiges Denken) und der Verschiedenheit der genannten Ideen (flexibles Denken) ausgewertet. Die Übereinstimmung zwischen den Ratern wurde regelmäßig geprüft ( $.77 \leq gT1 \leq 1.00$ ;  $.86 \leq gT2 \leq 1.00$ ). Mithilfe der einzelnen Flüssigkeits- und Flexibilitätswerte lassen sich für beide Erhebungszeitpunkte sowohl

das bereichsübergreifende divergente Denken ( $\alpha T1 = .87$ ;  $\alpha T2 = .87$ ) als auch das bereichsspezifische verbale ( $\alpha T1 = .85$ ;  $\alpha T2 = .84$ ) und figurale ( $\alpha T1 = .78$ ;  $\alpha T2 = .78$ ) divergente Denken faktoranalytisch abbilden.

Der TSD-Z liefert einen Schätzwert zum schöpferischen Potenzial der Schülerinnen und Schüler. Die Kinder sollten innerhalb einer begrenzten Zeit eine angefangene Zeichnung mit sechs Fragmenten fertig zeichnen. Das schöpferische Potenzial wird über die Bewertung verschiedener kreativitätsrelevanter Merkmale wie z.B. Elaboration, Sinngebung, Risikobereitschaft, Humor oder Unkonventionalität erschlossen. Die auszuwertenden Kategorien sind im Manual definiert und können mit insgesamt bis zu 72 Punkten bewertet werden. Die Kinderzeichnungen wurden von zwei unabhängigen Ratern mit zufriedenstellender Übereinstimmung ( $.93 \leq gT1 \leq .99$ ;  $.90 \leq gT2 \leq .97$ ) ausgewertet.

Im Vortrag werden die Messinstrumente zur Kreativitätserfassung vorgestellt und die Kreativitätsentwicklung im Verlauf des fünften Schuljahres mittels Varianzanalysen mit Messwiederholung abgebildet. Erste Ergebnisse zeigen, dass im Verlauf des fünften Schuljahres das divergente Denken signifikant zunimmt ( $MT1 = 27.54$ ,  $SDT1 = 7.38$ ;  $MT2 = 31.36$ ,  $SDT2 = 8.21$ ), während das schöpferische Potenzial signifikant abnimmt ( $MT1 = 19.08$ ,  $SDT1 = 9.34$ ;  $MT2 = 17.43$ ,  $SDT2 = 8.42$ ). Korrelationsanalysen zeigen geringe, aber bedeutsame Zusammenhänge zwischen den verschiedenen erfassten Kreativitätsfacetten und den Erhebungszeitpunkten ( $.134 \leq r \leq .211$ ,  $p < .001$ ).

Die Entwicklungsverläufe und Zusammenhangsstrukturen werden mit Blick auf weitere Analysen sowie Implikationen für die Kreativitätsförderung im fünften Schuljahr diskutiert.

## **Tanz wirkt – wirkt Tanz? Der Einfluss eines künstlerisch-pädagogischen Tanzunterrichts auf die Kreativitätsentwicklung von Grundschulkindern**

**Esther Pürgstaller, Nils Neuber**

Westfälische Wilhelms-Universität Münster, Deutschland

Kulturelle Bildung und Kreativität gelten in der Schule als implizites Bildungs- und Unterrichtsziel und sind in vielen Lehrplänen fest verankert (Bamford, 2006; Serve, 2000). Dennoch wird die Förderung von Kreativität im Schulkontext vernachlässigt (Schacter et al., 2006). Gründe dafür liegen u.a. an einem Mangel an Empfehlungen für die konkrete Umsetzung im Unterricht (Theurer, 2014). Um kreative Bildungsprozessen zu initiieren, sind Erfahrungs- und Möglichkeitsräume erforderlich, in denen (un)bekannte Muster entdeckt, entwickelt und durchbrochen werden können (Klinge, 2010). Diese werden im Kreativen Kindertanz mittels Aufgabenstellungen zur Improvisation, Gestaltung, Imitation und Reflexion bereitgestellt (Neuber, 2000). Daher bildet Tanz einen besonderen Anknüpfungspunkt für die Förderung von Kreativität. Die allgemeine These, Tanz fördere Kreativität, konnte allerdings bisher nur vereinzelt belegt werden (Minton, 2003; Bournelli & Mountakis, 2008). Vor dem Hintergrund der wenigen Studien fällt zudem auf, dass keine der Studien die methodisch-didaktischen Handlungsweisen des Unterrichts als Einflussgröße berücksichtigen (Behrens, 2013). Dabei wird jedoch darauf verwiesen, dass besonders bestimmte Persönlichkeitsmerkmale und Einstellungen der Lehrpersonen sowie die Unterrichtsgestaltung einen zentralen Einfluss auf die Entwicklung der Schülerinnen und Schüler nehmen können (Hattie, 2009; Helmke 2010; Cropley, 2001; de Souza Fleith, 2000).

Auf dieser Grundlage wird im Beitrag folgenden zentralen Fragestellungen nachgegangen:

1. Welche methodisch-didaktischen Handlungsweisen kommen im Tanzunterricht vor?
2. Inwiefern verändert sich die Kreativitätsentwicklung von Kindern im Grundschulalter während eines dreimonatigen künstlerisch-pädagogischen Tanzangebots?

Das Forschungsvorhaben ist eingebettet in die Studie „Tanz- und Bewegungstheater – Ein künstlerisch-pädagogisches Projekt zur Kulturellen Bildung in der Ganztagsgrundschule“, bei der Indikatoren des Selbstkonzepts, der emotionalen Kompetenz sowie der Kreativität bei Grundschulkindern erfasst wurden (Stern et al., 2017). Den Kern der quasi-experimentellen Längsschnitt-Studie im Versuchs-/Kontrollgruppendesign bildete ein dreimonatiges künstlerisch-pädagogisches Tanz- und Bewegungstheater-Angebot, das während des zweiten Schulhalbjahres 2015/2016 von Tanz- und Theaterpädagogen im Ganztagsangebot an zehn Mainzer Grundschulen durchgeführt wurde. Die Kinder ( $N = 277$ ) der ersten bis vierten Jahrgangsstufe wurden vor (Prä-Test), direkt im Anschluss (Post-Test) sowie drei Monate nach Abschluss des Projektes (Follow-Up) mittels des Motorischen Kreativitätstests 9 - 11 (MKT 9-11; Neuber, 2000) in den Kreativitätsfacetten der Produktivität, Problemlösungsfähigkeit und Originalität getestet. Für die vorliegende Studie wurden während des Tanzprojektes zudem sechzehn Unterrichtseinheiten bei vier Tanzpädagogen videographiert. Darauf aufbauend analysierten geschulte Kodierer den Unterricht über ein eigens entwickeltes Kategoriensystem und Kodiermanual mittels der Software Interact hinsichtlich seiner kreativitätsfördernden methodisch-didaktischen Handlungsweisen (Lektionsdauer, Sozialform, Unterrichtsmethoden, Freiheitsgrad der Aufgabenstellung, Kreativitätsfördernde Aufgabenstellungen). Die Codierungen enthalten die Häufigkeit und Dauer der einzelnen methodisch-didaktischen Handlungsweisen der Tanzpädagogen. Die Kreativitäts- und Unterrichtsdaten wurden schließlich mithilfe von mehrfaktoriellen Varianzanalysen mit Messwiederholung analysiert.

Die Ergebnisse zeigen, dass ein künstlerisch-pädagogisches Tanzprojekt die Entwicklung von Grundschulkindern und -schülern in den Kreativitätsfacetten der Produktivität ( $p = .019$ ;  $\eta^2 = .064$ ) und Originalität ( $p = .028$ ;  $\eta^2 = .038$ ) positiv beeinflusst. Die deskriptive Auswertung der Unterrichtsgestaltung lässt erkennen, dass der Unterricht in Bezug auf die Sozialformen lehrerzentriert ausgerichtet ist und die Unterrichtsmethoden der Gestaltung, Imitation, Improvisation und des Lehrervortrags den höchsten zeitlichen Anteil einnehmen. Der Freiheitsgrad der Aufgabenstellungen zeigt einen deutlich höheren Anteil an teiloffenen Aufgabenstellungen gegenüber geschlossenen Aufgabenstellungen. Dabei zielen weniger als die Hälfte aller gestellten Aufgabenstellungen auf die Förderung von Ideenreichtum (Produktivität), das Entdecken und Lösen von Problemen (Problemlösungsfähigkeit) und das Hervorbringen von unkonventionellen Ideen (Originalität).

Auf dieser Grundlage werden Überlegungen zum Zusammenhang zwischen der methodisch-didaktischen Unterrichtsgestaltung und der Kreativitätsentwicklung von Kindern angestellt und diskutiert.

## **Kreativitätsförderung im Kunstunterricht – eine Tautologie?**

**Caroline Theurer<sup>1</sup>, Nicole Berner<sup>2</sup>**

<sup>1</sup>Universität Kassel, Deutschland, <sup>2</sup>Pädagogische Hochschule Fachhochschule Nordwestschweiz, Schweiz

Kreativitätsförderung stellt allgemeindidaktisch ein Unterrichtsziel dar: Nicht nur in der Bildungstheorie, sondern auch in der Bildungsforschung rückt Kreativitätsförderung immer mehr in den Fokus. Dies drückt sich u. a. dadurch aus, dass Kreativität im internationalen Kontext als Kernkompetenz bezeichnet wird, wenn es darum geht, junge Menschen auf die Herausforderungen des modernen Zeitalters vorzubereiten (OECD, 2008).

Auch im deutschsprachigen Raum sehen die Bildungspolitik und die pädagogische Praxis Kreativität als förderungswürdig an. So wird Kreativitätsförderung in Schule und Unterricht seit einigen Jahren nicht nur implizit, sondern auch explizit in Bildungs- und Rahmenlehrplänen gefordert (zsf. Theurer, 2014). Damit erhalten Lehrpersonen den Auftrag zur Kreativitätsförderung, ohne dass hierzu genügend verlässliche Evaluationsstudien vorlägen. Die Frage, wie Kreativität gefördert werden kann, wird also gemeinhin als wichtig angesehen, ist aber nicht trivial, da Kreativität ein facettenreiches Konstrukt ist, welches durch vielfältige Faktoren beeinflusst wird (Preiser, 2006; Urban, 1993).

Künstlerischen Schulfächern wird dabei besonderes Potenzial zugesprochen. Sie bieten vergleichsweise viele Möglichkeiten, ästhetische Erfahrungen zu sammeln und sich gestaltend auszuprobieren. Besonders im Kunstunterricht können die zu be- und erarbeitenden Gegenstände vielfältige sinnliche Zugänge bieten und damit ästhetisches Erfahren erleichtern (z. B. Dewey, 1988; Schäfer, 1999; zsf. Berner, 2016a). Solche künstlerischen Auseinandersetzungen und ästhetischen Erfahrungen dürften aus theoretischer Sicht kreative Prozesse auslösen (Rogh et al., 2017). Allerdings gelingt dies nicht immer und von den Gegenständen ausgehend allein: Lehrpersonen sollten hier als aktive Begleiter fungieren, damit Kunstunterricht tatsächlich positiv auf kreative Leistungen der Schüler wirken kann (Berner, 2016b).

Dass sich das Verhalten von Lehrpersonen in verschiedenen Unterrichtsfächern positiv auf die Entwicklung von domänenunspezifischer Kreativität auswirken kann, wurde von Theurer (2014) untersucht. Inwiefern sich auch fachspezifische Wirkungen zeigen, wurde in dieser Studie nicht untersucht, obwohl ein domänenspezifischer Blick auf Verhalten von Lehrpersonen im Kunstunterricht ein spezifisches Erkenntnisinteresse birgt.

Im Vortrag soll deshalb der Frage nachgegangen werden, inwieweit sich kreativitätsförderndes Verhalten von Lehrpersonen im Grundschul-Kunstunterricht auf die bildnerische Kreativität auswirkt.

Bearbeitet wird die Fragestellung unter Rückgriff auf Daten aus der Grundschulstudie PERLE (Lipowsky et al., 2013), in deren Rahmen u. a. Kunstunterricht unter curricular vergleichbaren Bedingungen videografiert wurde (Berner et al., 2013). Zudem wurde die Kreativität zu mehreren Messzeitpunkten erfasst.

Zum Zeitpunkt der Videostudie im Fach Kunst waren die Grundschul Kinder durchschnittlich rund siebeneinhalb Jahre alt. In den Doppelstunden Kunst (N = 33) produzierten die Grundschul Kinder (N = 612) Plastiken aus Modelliermasse, welche von Berner (2013) in Hinblick auf bildnerische Kreativität analysiert wurden. Faktorenanalytisch wurden der zielorientierte Einsatz bildnerischer Mittel (ZO;  $\alpha = .83$ ) und die Unkonventionalität im bildnerischen Gestalten (UK;  $\alpha = .57$ ) als Facetten bildnerischer Kreativität extrahiert.

Die videografierten Unterrichtseinheiten wurden für die Fächer Deutsch, Mathematik und Kunst u. a. systematisch nach kreativitätsfördernden Elementen untersucht (Theurer, 2014). Betrachtet wurden sowohl kognitive als auch emotionale Aspekte des Unterrichtsgeschehens: So wurde nicht nur erfasst, welche Materialien und Methoden zum Einsatz kamen, sondern auch, wie vertrauensvoll und wertschätzend die Interaktionsstrukturen zwischen den Beteiligten wirkten bzw. inwieweit Unabhängigkeit im Denken und Eigentätigkeit gefördert wurden.

Neben deskriptiven Befunden zu den Schülerplastiken und Aspekten kreativitätsfördernden Lehrerhandelns werden im Vortrag Zusammenhänge zwischen diesen Merkmalen berichtet. Mehrebenenanalytisch kann gezeigt werden, dass die erfassten Unterrichtsmerkmale die plastische Kreativität der Grundschul Kinder unterschiedlich beeinflussen. Kognitive Aspekte des kreativitätsfördernden Lehrerhandelns wirken sich positiv auf die Unkonventionalität im bildnerischen Gestalten aus ( $\beta = .425$ ,  $p < .05$ ; ZO:  $\beta = -.135$ ns) und emotionale Aspekte kreativitätsfördernden Lehrerhandelns begünstigen den zielorientierten Einsatz bildnerischer Mittel ( $\beta = .362$ ,  $p < .05$ ; UK:  $\beta = .157$ ns). Kinder in Klassen, deren Lehrpersonen z. B. offenere Aufgaben stellen und häufiger unkonventionelle Ideen der Schüler für den Unterrichtsverlauf nutzen, formen unkonventionellere Plastiken.

## **Wie lassen sich Best-Practice-Beispiele Kultureller Bildung identifizieren?**

**Katrin Knoll**

Justus-Liebig-Universität Gießen, Deutschland

Kulturelle Bildung wird als Schlüssel für die Förderung von Kreativität betrachtet. Betrachtet man allerdings die zahlreichen Veröffentlichungen, die sich inzwischen mit der Qualität Kultureller Bildung beschäftigen (u. a. Eger, 2014; Seidel et al., 2009; Unterberg, 2014; Wimmer, Schad & Nagel, 2013) so wird deutlich, dass Kulturelle Bildung – wie auch andere Bildungsangebote – bestimmte Kriterien erfüllen muss, um wirksam zu sein.

Merkmale, um Qualität zu bestimmen, sind beispielsweise ästhetische Gesichtspunkte, strukturelle Aspekte oder der Grad der Professionalisierung von Lehrkräften (Unterberg, 2014; Wimmer, Schad & Nagel, 2013). Ferner ergibt sich eine pädagogische Perspektive auf die Förderung von Kreativität (Runco, 2004) – insbesondere dann, wenn Kulturelle Bildung im Kontext Schule stattfindet.

Kreativität und kreative Prozesse finden im Zusammenspiel von Individuen in sozialen Kontexten statt (Meusburger, 2010). Kreative Prozesse gelten als ein Spezifikum von künstlerischen Projekten, die innerhalb eines kreativen Klimas (Isaksen, 2010) begünstigt werden. Ein kreatives Klima zeichnet sich unter anderem durch die Autonomie der Beteiligten, die Unterstützung und Aktivierung durch Lehrpersonen, sowie Freude am Tun, Motivation und die Abwesenheit von Konflikten aus (Isaksen, 2010; Preiser, 2011).

Definiert man die Förderung von Kreativität als pädagogisches Ziel, dann ergibt sich daraus, dass gute kulturelle Bildungsangebote durch ein kreatives Klima und kreative Arbeitsprozesse gekennzeichnet sind. Deshalb sollten kreative Outcomes in Verbindung mit kreativitätsförderlichen Merkmalen untersucht werden.

Die vorliegende Studie erfolgt im Rahmen der Begleitforschung zum Modellprogramm „Kulturagenten für kreative Schulen“ (Abs et al., 2013). Im Modellprogramm initiieren Lehrpersonen künstlerische Projekte in Kooperation mit Künstler/innen. Als Datengrundlage dient die Befragung von Schüler/innen und Künstler/innen. Die Angebote differieren u.a. sowohl in Angebotsformat und Dauer, als auch in der Kunstform. Aus diesem Grund ist es erforderlich, sowohl kreative Outcomes, wie auch kreative Prozesse kunstformübergreifend zu erfassen.

Die Perspektive der Künstler/innen wurde zum Projektende anhand eines Fragebogens, der in Anlehnung an das Creative Learning Assessment (CLA) (Ellis & Lawrence, 2009) entwickelt wurde, erhoben. Die entwickelten Fragen erfassen Prozesse innerhalb der Schüler/innengruppe während der Projektarbeit.

Die Schüler/innenperspektive auf die Qualität der Angebote wurde ebenfalls mittels eines Fragebogens untersucht, der in Anlehnung an die Beurteilung außerunterrichtlicher Angebote in der Ganztagschule (Quellenberg, 2009) entwickelt und um Fragen, die Spezifika eines kreativen Klimas erfassen (Isaksen, 2010), ergänzt wurde.

Kreative Outcomes bei Schüler/innen werden durch ein standardisiertes Testinstrument (TSD-Z, Urban & Jellen, 1995; 2010) und den selbsteingeschätzten kreativen Lernstil in Anlehnung an das ELLI-Inventory (Deakin-Crick, 2007; Deakin-Crick & Yu, 2008) erfasst.

Die Angebote werden mittels Einschätzung der Künstler/innen und der Schüler/innen auf ihre Kreativitätsförderlichkeit hin untersucht. Von Interesse ist deren (gemeinsames) Potenzial zur Identifikation von Best-Practice-Beispielen eines kreativen Klimas und kreativer Arbeitsprozesse.

Zwar korrelieren die beiden Perspektiven auf die Qualität Kultureller Bildung deutlich miteinander, verhalten sich jedoch unterschiedlich in Bezug auf kreative Outcomes der Schüler/innen. Damit zeigt sich einerseits, dass unterschiedliche Facetten der Prozessqualität gemessen werden und andererseits, dass die Förderung von Kreativität auf unterschiedlichen Ebenen möglich ist.

## A13

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H118

### **Kognitive und überzeugungsbezogene Aspekte der Rezeption wissenschaftlicher Evidenz durch Laien**

Chair(s): **Eva Thomm** (Universität Erfurt), **Johannes Bauer** (Universität Erfurt)

DiskutantIn(nen): **Robin Stark** (Universität des Saarlandes)

Im Alltag, aber auch im Beruf, müssen sich Laien in vielfältiger Weise mit wissenschaftlicher Evidenz auseinandersetzen – sei es um eine persönliche Entscheidungen zu treffen oder um ihr professionelles Handeln zu informieren (Bromme & Kienhues, 2017). Fähigkeiten, wissenschaftliche Evidenz zu rezipieren und kritisch zu bewerten, sind in verschiedenen formellen und informellen Lernkontexten von hoher Bedeutung und werden sowohl als schulisches Bildungsziel (OECD, 2013) als auch im Kontext evidenzbasierten professionellen Handelns (Sacket et al., 1996) als zentral angesehen. Dies erfordert neben kognitiven Fähigkeiten wissenschaftlichen Denkens und Argumentierens, auch die Entwicklung relevanter motivationaler Orientierungen und Überzeugungssysteme (Fischer et al., 2014). Dabei stellt die Unsicherheit, Konflikthaftigkeit und Komplexität wissenschaftlichen Wissens Laien, die selbst keine Wissenschaftler/-innen sind, vor große Herausforderungen: Sie verfügen nur über begrenzte Urteilsfähigkeit, die Validität wissenschaftlicher Erkenntnisse direkt zu bewerten und sind deshalb auch auf vertrauensbasierte Urteile angewiesen (Bromme & Goldman, 2014). Die Nutzung wissenschaftlichen Wissens erfordert somit eine basale Rezeptionskompetenz, um sich auf dem Kontinuum zwischen „blindem“ Vertrauen in die Wissenschaft und einem tiefgehenden Verständnis wissenschaftlicher Befunde verorten zu können. In diesem Kontext stellen sich zahlreiche Fragen für die Bildungsforschung: Unter welchen Umständen sehen Laien es als erstrebenswert an, wissenschaftliche Evidenz für alltägliche oder berufliche Fragen heranzuziehen? Welche Prozesse laufen beim evidenzbasierten Denken ab und wie sind diese durch Vorannahmen, Einstellungen und Überzeugungen beeinflusst? Welche Fähigkeiten können und müssen Laien auf welchen Niveaus erwerben, um Evidenz angemessen bewerten und nutzen zu können? Und wie kann der Erwerb dieser Fähigkeiten instruktional unterstützt werden?

Das Symposium befasst sich vor diesem Hintergrund mit der zentralen Frage, wie Laien im Beruf sowie im Alltag mit wissenschaftlicher Evidenz umgehen und diese erfolgreich nutzen können. Die vier Beiträge betrachten das Zusammenspiel und Verhältnis hierfür relevanter kognitiver und überzeugungsbezogener Faktoren. Dabei beziehen sich die ersten beiden Beiträge auf professionsbezogene Kontexte, zwei weitere auf die Rezeption wissenschaftlicher Evidenz im Alltag. Beitrag 1 analysiert anhand eines skripttheoretischen Ansatzes, inwieweit Expertise und Einstellung gegenüber Bildungswissenschaften evidenzorientiertes Denken von Lehramtsstudierenden und die Anwendung bildungswissenschaftlicher Befunde beeinflusst. Beitrag 2 untersucht, inwiefern Einstellungen und Überzeugungen von Lehrkräften gegenüber evidenzbasierter Praxis von Zugriffsmöglichkeiten auf Forschungsressourcen, Selbstkonzepten zu forschungsmethodischen Kompetenzen sowie schulischen Kontextvariablen abhängen. Beitrag 3 betrachtet unter Berücksichtigung persönlicher Einstellungen, wie Laien unterschiedliche Positionen in evidenzbasierten Expertendiskussionen rezipieren, um einen Standpunkt zu einer wissenschaftsbezogenen Fragestellung zu entwickeln. Beitrag 4 schließt hier an, indem er Bedingungen analysiert, unter denen eine angemessene Rezeption wissenschaftlicher Evidenz im Alltag gelingen kann, auch wenn diese persönlich motiviert ist.

#### *Beiträge des Symposiums*

### **Evidenzorientiertes Denken von Lehramtsstudierenden bei der Bearbeitung pädagogischer Problemstellungen: Ein Vergleich von Studienanfänger(inne)n und fortgeschrittenen Studierenden**

**Katharina Kiemer, Ingo Kollar**  
Universität Augsburg

Sowohl von bildungspolitischer (KMK, 2014) als auch von bildungswissenschaftlicher Seite (Bromme, Prenzel & Jäger, 2014) wird gefordert, dass Lehrkräfte pädagogische Problemsituationen nicht auf der Grundlage von Erfahrungswissen, sondern auf Basis von bildungswissenschaftlichen Theorien und empirischen Forschungsergebnissen lösen. Dies sicherzustellen ist eine genuine Aufgabe der universitären Lehramtsausbildung. Studien zeigen jedoch, dass angehende wie auch praktizierende Lehrkräfte häufig Probleme damit haben, diesem Anspruch nachzukommen (z.B. Wagner, Klein, Klopp & Stark, 2014). Als Gründe hierfür werden u.a. das mangelnde bildungswissenschaftliche Vorwissen (Star & Strickland, 2008) sowie ungünstige Einstellungen gegenüber der Nützlichkeit von in den Bildungswissenschaften erbrachten Erkenntnissen angesehen (z.B. Parr & Timperley, 2008). Im vorliegenden Beitrag wird der Frage nachgegangen, inwiefern für diese Probleme möglicherweise auch potenziell defizitäre Skripts zum Umgang mit pädagogischen Problemsituationen verantwortlich sind. Skripts bezeichnen prozedurales Wissen, das Lehrkräfte und Lehramtsstudierende darin anleitet, wie sie entsprechende Probleme analysieren (Fischer, Kollar, Stegmann & Wecker, 2013). Aus der Forschung (z.B. Sherin & van Es, 2009) kann abgeleitet werden, dass Lehrende während der Analyse pädagogischer Problemsituationen idealerweise fünf kognitive Prozesse durchlaufen sollten: (1) Problemidentifikation, (2) Problemrekonstruktion, (3) Aufstellen eines kausalen Erklärungsmodells, (4) eigene Handlungsziele formulieren und (5) Handlungsoption(en) auswählen.

#### Fragestellung

Die Studie geht folgenden Fragen nach:

1. Inwiefern unterscheiden sich beginnende von fortgeschrittenen Lehramtsstudierenden hinsichtlich ihrer Skripts zum evidenzorientierten Denken?
2. Inwiefern werden die Skripts zum evidenzorientierten Denken von Lehramtsstudierenden durch deren Einstellungen gegenüber den Bildungswissenschaften beeinflusst?
3. In welchem Zusammenhang stehen die Skripts zum evidenzorientierten Denken von Lehramtsstudierenden mit deren Fähigkeit zur Anwendung bildungswissenschaftlicher Theorien und Evidenzen?

#### Methode

Versuchsteilnehmer(innen) waren N = 339 Lehramtsstudierende (N = 103 Studienanfänger(innen), N = 236 fortgeschrittene Studierende). Diese wurden gebeten, eine per Fallvignette vorgelegte pädagogische Problemsituation schriftlich zu analysieren. Die Messung der Skripts zum evidenzorientierten Denken erfolgte zunächst über eine Analyse der Art und Reihenfolge der Aktivitäten, von denen die Teilnehmer(innen) anschließend angaben, sie während der Fallbearbeitung durchgeführt zu haben. Diese Messung wird derzeit über eine objektive Kodierung der tatsächlichen Fallbearbeitungen validiert. Die Fallbearbeitungen werden zudem momentan darauf hin analysiert, inwiefern die Versuchspersonen

bildungswissenschaftliche vs. subjektive Theorien sowie bildungswissenschaftliche Evidenzen vs. Erfahrungswissen zur Analyse der Situation anwendeten. Die Einstellungen gegenüber den Bildungswissenschaften wurde per Likert-skaliertem Fragebogen erhoben (eigene Entwicklung; Cronbachs Alpha = .88).

#### Ergebnisse

Hinsichtlich Forschungsfrage 1 wurde deutlich, dass sich Studienanfänger(inne)n bezüglich der Abdeckung von drei der fünf theoretisch angenommenen kognitiven Prozesse von den fortgeschrittenen Studierenden unterschieden: So gaben fortgeschrittene Studierende zwar signifikant häufiger als Studienanfänger(innen) an, Probleme zu identifizieren ( $X^2(1) = 4.00, p < .05, d = .14$ ), zeigten aber gleichzeitig ein signifikant geringeres Engagement bzgl. der Formulierung eigener Handlungsziele ( $X^2(1) = 7.27, p < .01, d = .27$ ) sowie ein zumindest tendenziell niedrigeres Engagement hinsichtlich der Problemrekonstruktion ( $X^2(1) = 2.86, p = .09, d = .16$ ). Hinsichtlich der übrigen beiden Prozesse ergaben sich keine signifikanten Unterschiede ( $X^2(1) > 1.82, n.s.$ ). Bezüglich Forschungsfrage 2 ergab sich, dass die Einstellungen gegenüber den Bildungswissenschaften einen schwachen, aber signifikanten Prädiktor für die Gesamtzahl der oben beschriebenen idealtypischen Aktivitäten darstellten ( $b = .24, p = .05, R^2 = .02$ ). Die Analysen zu Forschungsfrage 3 laufen derzeit. Ergebnisse werden auf der GEBF-Konferenz 2018 verfügbar sein.

Die Ergebnisse zeigen, dass sich expertisebedingte Unterschiede zwischen Studierenden hinsichtlich des evidenzorientierten Denkens auch auf Basis ihrer Skripts zum evidenzorientierten Denken identifizieren lassen. Dass die Einstellungen gegenüber den Bildungswissenschaften zumindest einen kleinen Teil der Varianz in den Skripts erklären, lässt es notwendig erscheinen, im Rahmen der universitären Lehramtsausbildung die Relevanz und Anwendbarkeit bildungswissenschaftlicher Erkenntnisse für professionelles Handeln explizit zu diskutieren und zu reflektieren. In welchem Zusammenhang die Skripts zum evidenzorientierten Denken zudem mit der Anwendung bildungswissenschaftlicher Theorien und Evidenzen stehen, werden die derzeit laufenden Analysen zeigen.

### **Eine Frage des Zugangs? Einstellungen und Überzeugungen von Lehrkräften gegenüber evidenzbasierter Praxis**

**Johannes Bauer<sup>1</sup>, Manfred Prenzel<sup>2</sup>, Christine Sälzer<sup>2</sup>, Eva Thomm<sup>1</sup>**  
<sup>1</sup>Universität Erfurt, <sup>2</sup>TU München

#### Theoretischer Hintergrund

Im Kontext bildungspolitischer und wissenschaftlicher Diskussionen um Standards von Lehrerprofessionalität wird argumentiert, dass Lehrkräfte in der Lage sein sollten, Forschungsbefunde aus ihren Unterrichtsfächern, sondern auch aus den Bildungswissenschaften zu rezipieren, kritisch zu bewerten und diese im Sinne einer evidenzbasierten Praxis für ihr professionelles Handeln zu nutzen (European Commission, 2007; KMK, 2014). Dies setzt voraus, dass Lehrkräfte Zugang zu entsprechenden Ressourcen haben, die primäre oder kompilierte Evidenz bereitstellen (Seidel et al., 2017). Zudem erfordert die Bewertung und Nutzung von Evidenz neben kognitiven Fähigkeiten (inhaltliches und methodisches Fachwissen) günstige motivationale Orientierungen, Einstellungen und Überzeugungen (Bauer et al., 2017; Heffer et al., 2017; Parr & Timperley, 2008). Allerdings existieren nur wenige größer angelegte Studien, die die Einstellung von Lehrkräften gegenüber evidenzbasierter Praxis und bildungswissenschaftlicher Forschung untersuchen. Es bleibt folglich unklar, welche Bedeutung Lehrkräfte evidenzbasierter Praxis und bildungswissenschaftlicher Forschung für das eigene professionelle Handeln beimessen, und welche individuellen und kontextuellen Faktoren ihre Einstellungen und Überzeugungen beeinflussen (Niemi, 2008; Parr & Timperley, 2008).

#### Fragestellung

Die vorliegende Studie prüfte ein adaptiertes theoretisches Modell von Niemi (2008). Diesem Modell lag die Annahme zugrunde, dass der Zugriff auf Forschungsressourcen, unterstützende Aspekte der Schulkultur sowie Selbstkonzepte zu forschungsmethodischen Kompetenzen die Einstellungen und Überzeugungen von Lehrkräften gegenüber evidenzbasierter Praxis präzidieren. Hierzu wurde untersucht, (1) welche relativen Varianzanteile in Einstellungen und Überzeugungen zu evidenzbasierter Praxis auf Schul- und Individualebene aufgeklärt werden, und (2) welche Modellvariablen diese Einstellungen und Überzeugungen vorhersagen.

#### Methode

Im Rahmen der Lehrerbefragung zu PISA 2012 in Deutschland (Prenzel, Sälzer, Klieme & Köller, 2013) wurden N = 2549 Mathematiklehrkräfte aus 272 Schulen u.a. zu folgenden Skalen befragt (vgl. Jette et al., 2003; Johnston et al., 2003; Young & Ward, 2001): Zugang zu Forschungsliteratur (3 Items,  $\alpha = .71$ ); Selbstkonzepte zu Forschungsmethoden und Statistik (7 Items,  $\alpha = .91$ ); positive Einstellungen gegenüber evidenzbasierter Praxis (2 Items,  $\alpha = .82$ ); Überzeugungen hinsichtlich subjektiver Barrieren: fehlende Zeit und Kenntnisse (3 Items,  $\alpha = .77$ ) und Irrelevanz bildungswissenschaftlicher Forschung (3 Items,  $\alpha = .72$ ).

#### Ergebnisse

Hinsichtlich der Relevanz von Faktoren auf der Schul- und Individualebene wurden in einem Zweiebenen-Strukturgleichungsmodell latente ICCs für die Einstellungen gegenüber wissenschaftlichen Befunden sowie subjektive Barrieren für deren Nutzung berechnet. Diese fielen insgesamt gering (Einstellungen ICC = .035; Barrieren: fehlende Zeit und Kenntnisse ICC = .016; Barrieren: Irrelevanz ICC = .037) und nur für die subjektive Irrelevanz bildungswissenschaftlicher Forschung statistisch signifikant aus. Unterschiede auf der Schulebene klären also kaum substantielle Varianz in Einstellungen und Überzeugungen durch auf (Fragestellung 1).

Auf individueller Ebene zeigte sich, dass der Zugang zu Forschungsressourcen und methodische Selbstkonzepte signifikante Prädiktoren einer positiven Einstellung zu evidenzbasierter Praxis und subjektiven Barrieren sind (alle Pfade  $p < .05$ ; Fragestellung 2). Dabei konnten insbesondere Unterschiede in der Wahrnehmung subjektiver Barrieren aufgeklärt werden (fehlende Zeit und Kenntnisse  $R^2 = .41$ ; Irrelevanz bildungswissenschaftlicher Forschung  $R^2 = .16$ ). Lehrkräfte mit einem besseren Zugang zu Forschungsressourcen und positiveren forschungsmethodischen Selbstkonzepten wiesen hier niedrigere Werte auf (Range der beta-Gewichte  $-.50 < -.13$ ). Für die positiven Einstellungen gegenüber evidenzbasierter Praxis fiel die aufgeklärte Varianz geringer aus ( $R^2 = .08$ ). Hier zeigten sich kleine positive Effekte beider Prädiktoren in der erwarteten Richtung (Zugang beta = .17; Selbstkonzepte beta = .16).

Diese Befunde untermauern Annahmen von Niemi (2008), indem sie die Relevanz von Variablen auf individueller Ebene bestätigen. Sie legen aber auch nahe, dass Variablen auf Schulebene eine weniger bedeutsame Rolle spielen als im Modell angenommen. Eine positive Einstellung scheint demnach weniger durch das schulische und kollegiale Umfeld als vielmehr



durch den Zugang zu Wissensressourcen und individuellen Voraussetzungen bedingt zu sein. Dieser Befund legt nahe, dass potentielle Interventionen zur Förderung evidenzbasierter Einstellungen und Überzeugungen hier Ansatzpunkte finden können (vgl. Hefter, et al., 2017).

## **Der Einfluss von Kommunikationsstil und Voreinstellungen auf die Verarbeitung widersprüchlicher Befunde in Expertendiskussionen**

**Elisabeth Mayweg-Paus<sup>1</sup>, Friederike Hendriks<sup>1</sup>, Regina Jucks<sup>2</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, <sup>2</sup>Wahrgenommene Interessenkonflikte bei der Rezeption wissenschaftlicher Evidenz: Wenn motivierte Informationsverarbeitung sinnvoll ist

Theoretischer Hintergrund & Fragestellung

Ein Format, welchem Laien bei der Recherche zu wissenschaftlichen Informationen im Internet begegnen, ist das der Expertendiskussion. Hier argumentieren Experten mit unterschiedlichen – und sich widersprechenden – Evidenzen. Im Unterschied zu Texten, die durch einen Autor formuliert sind, werden in Diskussionen Perspektiven direkt durch die Experten repräsentiert und im Diskurs entwickelt. Dabei können die Experten sowohl einseitig, als auch zweiseitig kommunizieren (Mayweg-Paus & Jucks, 2017). Die vorliegende Arbeit untersucht folgende Fragestellungen:

1. Wie wirkt sich die Rezeption einer Expertendiskussion, die entweder einseitige oder zweiseitige Expertenstandpunkte verfolgt, auf die Einstellung der Probanden und die Argumentation zum Thema aus?
2. Welchen Einfluss hat die Voreinstellung auf die Argumentation der Probanden?

Methode

Wir berichten eine Studie mit  $N = 157$  Studierenden, in der die Verteilung von Pro- und Kontrainformationen zwischen zwei Experten innerhalb einer schriftlichen Diskussion manipuliert wurde. In zwei Experimentalgruppen wurden Pro- und Kontra-Argumente zum Thema „Computernutzung im Grundschulalter“ entweder zwischen Experten aufgeteilt (d.h. ein Experte vertrat nur Pro oder nur Kontra Argumente, Bedingung „einseitig“;  $n = 74$ ) oder innerhalb der Experten formuliert (Bedingung „zweiseitig“;  $n = 83$ ). Die Einstellung der Probanden wird durch die Variable „Kinder im Grundschulalter sollen den Computer nutzen“ auf einer 7-Punkte Skala erfasst. Zudem wurde die Zustimmung zu vier zentralen Argumenten (2 x Kontra- und 2x Pro) des Diskurses auf einer 7-Punkte Skala vor und nach dem Lesen der Expertendiskussion erhoben. Die Probanden wurden außerdem gebeten einen eigenen Standpunkt zu formulieren. Anschließend wurde die Entscheidungssicherheit mit ausgewählten und adaptierten Items der Decisional Conflict Scale (nach O'Connor, 1995) erfasst.

Ergebnisse

Die Auswertung erfolgte mithilfe von Kovarianzanalysen, wobei die Einstellung als zentrierte Kovariable einbezogen wurde (Richter, 2006). Die Einstellung vor dem Lesen der Expertendiskussion hatte einen signifikanten Beitrag zu der Zustimmung einzelner zentraler Argumente nach dem Lesen – allerdings war die Zustimmung zu allen Argumenten auch beeinflusst durch das Lesen der Expertendiskussion (die Zustimmung zu allen Argumenten war bis auf eine Ausnahme nach dem Lesen der Diskussion höher; sowohl zu Pro- als auch Kontraargumenten). Nur das Pro-Argument „Die Nutzung von Computern hilft Kindern im Grundschulalter beim Lernen“ resultierte in höherem Zuwachs an Zustimmung in der Experimentalgruppe, die eine zweiseitige Expertendiskussion gelesen hatte (Interaktion:  $F(1,154) = 5.31, p = .02$ ).

Hinsichtlich der Argumentation wurde ausgewertet, wie viele Argumente aus dem Text die Probanden in ihre Stellungnahme einfließen ließen. Hierbei ergab sich, dass die Nutzung von Kontra-Argumenten durch die Einstellung prädiert wurde;  $F(1,154) = 16.68, p < .001$ , aber nicht durch die Bedingungszugehörigkeit;  $F(1,154) = .67, p = .42$ . Die Anzahl der genannten Pro-Argumente war weder durch die Einstellung;  $F(1,154) = .05, p = .82$ , noch durch die Bedingungszugehörigkeit beeinflusst;  $F(1,154) = .47, p = .49$ . Im Mittel wurden  $M = 1.15, SD = 1.24$  Kontra- und  $M = 0.94, SD = 0.86$  Pro-Argumente genannt. Probanden, die eine einseitige Expertendiskussion gelesen hatten, waren mit ihrem Standpunkt zufriedener als solche, die eine zweiseitige gelesen hatten;  $F(1,154) = 5.00, p = .03$ , auch hier hatte die Voreinstellung einen Einfluss;  $F(1,154) = 7.27, p < .01$ .

Hieraus lässt sich schließen, dass die Voreinstellung der Probanden zu dem Thema einen Einfluss auf die Auswahl der Argumente in einer persönlichen Stellungnahme hatte. So nutzten Probanden mit einer negativen Einstellung mehr Kontraargumente. Der Kommunikationsstil in der Expertendiskussion (einseitig vs. zweiseitig) hatte hierauf keinen Einfluss, allerdings konnte eine höhere Entscheidungszufriedenheit in der Experimentalgruppe festgestellt werden, die eine einseitige Expertendiskussion gelesen hatte.

Basierend auf den dargestellten Ergebnissen sollen Ansätze diskutiert werden, wie Rezipienten bei der Verarbeitung von Expertendiskussionen unterstützt werden. Ebenfalls soll überlegt werden, wie Experten ihre Kommunikationsstrategien verbessern können. Dabei wird der Fokus nicht nur auf die Inhalte, sondern auch auf die Rolle der rhetorischen Anforderungen gesetzt.

## **Wahrgenommene Interessenkonflikte bei der Rezeption wissenschaftlicher Evidenz: Wenn motivierte Informationsverarbeitung sinnvoll ist**

**Lukas Gierth<sup>1</sup>, Rainer Bromme<sup>2</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, <sup>2</sup>Westfälische Wilhelms-Universität Münster

Theoretischer Hintergrund

Wissenschaftliche Evidenz ist auch für Laien, z.B. in Gesundheitsfragen, von alltäglicher Relevanz, weshalb sie meistens mit gewissen Motivationen an die Suche und Verarbeitung wissenschaftlicher Informationen herantreten. Den meisten Menschen fehlt dabei das nötige Vorwissen, um wissenschaftliche Informationen auf inhaltlicher Ebene zu bewerten. Stattdessen richten sie sich in der Bewertung wissenschaftlicher Evidenz nach persönlichen Vorlieben, Meinungen und bereits bestehenden Überzeugungen zum jeweiligen Thema (Sinatra, Kienhues, & Hofer, 2014). Diese Form der motivierten Informationsverarbeitung ist oft Ursache von unbalancierter Verarbeitung wissenschaftlicher Evidenz und Teil des sogenannten confirmation bias (Nickerson, 1998). Jedoch ist motivierte Informationsverarbeitung der Rezeption und Verarbeitung wissenschaftlicher Evidenz nicht immer abträglich. So gibt es Hinweise, dass Menschen zum Beispiel bessere Leistungen in der Analyse von Evidenz zeigen, die ihr Selbstbild bedroht (Mata, Sherman, Ferreira, & Mendonça, 2015). Auch logisch-deduktive Fähigkeiten können sich bessern, wenn Menschen motiviert sind, einen Betrüger zu finden (Cosmides, 1989). Laien können also unter bestimmten Umständen auf Grundlage persönlicher Motivationen, die wissenschaftliche Evidenz genauer evaluieren und somit eher zu richtigen Urteilen gelangen.

Eine Motivation bei der Rezeption wissenschaftlicher Evidenz ist es, zuverlässiges und gesichertes Wissen zu erhalten. Eine Grundvoraussetzung dafür ist, dass die Quelle von wissenschaftlichen Informationen vertrauenswürdig ist. So können Quelleninformationen gemeinsam mit inhaltlichen Bewertungen einer wissenschaftlichen Position in ein Urteil einfließen (Hahn, Harris, & Corner, 2009) und, wenn ein Interessenkonflikt der Experten wahrgenommen wird, sogar vor Fehlinformationen schützen (Van Der Linden et al., 2017).

#### Fragestellung

Mit dieser Studie, suchen wir herauszufinden, ob Quelleninformationen, die auf einen möglichen Interessenkonflikt von Experten hinweisen, sich auf kognitive Prozesse bei der Auseinandersetzung mit Evidenz auswirken. Dabei vermuten wir, dass etablierte kognitive Verzerrungen im logisch-deduktiven Schlussfolgern und auch bei numerischen Fähigkeiten verringert werden, wenn Laien durch die Wahrnehmung eines Interessenkonflikts motiviert sind, die ihnen vorliegende Information genauer zu prüfen.

#### Methode

Die Studie besteht aus zwei aufeinanderfolgenden Experimenten, die jeweils einem Zwischensubjektdesign folgen. Im ersten Experiment erhalten die Probanden einen kurzen Online-Artikel, in dem gesagt wird, dass Zuckerkonsum die Erinnerungsleistung fördert. Die Daten einer Tabelle im Text widerlegen zwar die Behauptung, das war aber nur bei reflektierter Verarbeitung zu erkennen. Als abhängige Variable wird erfasst, ob sie den tatsächlichen Zusammenhang aus der Tabelle (einer Kontingenztafel) ablesen können. Die Analyse derartiger Kontingenztafeln ist für Fehler anfällig (West, Toplak, & Stanovich, 2008). Dabei wird experimentell variiert, ob die Wissenschaftler, die die Daten in der Kontingenztafel sammelten, einem Interessenkonflikt unterlagen. Die Probanden werden einer von drei Bedingungen randomisiert zugewiesen: (1) die Daten sprechen für die Interessen der Wissenschaftler, (2) die Daten sprechen gegen die Interessen der Wissenschaftler und (3) kein Interessenkonflikt lag vor (Kontrollbedingung). Als Manipulationscheck erheben wir mit dem Muenster Epistemic Trustworthiness Inventory (METI; Hendriks, Kienhues, & Bromme, 2015) die wahrgenommene Vertrauenswürdigkeit der jeweiligen Quelle.

Im zweiten Experiment wird Wasons (1968) Karten-Selektion-Aufgabe im Kontext einer kurzen Geschichte über medizinische Expertenaussagen vor Gericht genutzt. Probanden müssen dabei Falsifikationsregeln korrekt anwenden, um zweifelhafte Expertenaussagen zu überprüfen. In der Experimentalbedingung waren diese zweifelhaften Aussagen auf einen Interessenkonflikt zurückzuführen, in der Kontrollbedingung nur auf den Arbeitsort. Als abhängige Variable wird die korrekte Anwendung der Falsifikationsregel gewertet.

#### Ergebnisse

Da die Erhebung noch nicht beendet ist, können hier noch keine Ergebnisse berichtet werden. Für die Auswertung des ersten Experiments erwarten wir, dass Probanden bei einem wahrgenommenen Interessenkonflikt nur dann eine bessere Leistung als die Kontrollgruppe zeigen, wenn die Daten für die Interessen der Wissenschaftler sprechen. Wir erwarten also: Nur wenn Anlass für die Vermutung besteht, dass die Wissenschaftler die Daten der Tabelle im Sinne ihrer Interessen falsch gedeutet haben, aktiviert dies eine kritische Prüfung. Für die Auswertung der Karten-Selektionsaufgabe erwarten wir, dass Probanden in der Interessenkonflikt-Bedingung die Wason-Aufgabe eher lösen, als Probanden in der Kontrollbedingung.

## A14

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H117

### Studien mit Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an inklusiven Schulen - Welchen Beitrag können Large-Scale-Assessments leisten?

Chair(s): **Aleksander Kocaj** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB)), **Lena Nusser** (Leibniz-Institut für Bildungsverläufe (LifBi)), **Christine Sälzer** (Technische Universität München (TUM))

DiskutantIn(nen): **Christine Sälzer** (Technische Universität München (TUM))

Das deutsche Bildungssystem steht vor der Herausforderung, für SchülerInnen mit sonderpädagogischem Förderbedarf (SPF) inklusive Bildungsangebote bereitzustellen, die den Zielen der UN-Behindertenrechtskonvention gerecht werden. Dabei tragen verschiedene Faktoren zum Gelingen einer inklusiven Schule bei. Zur Erforschung der zugrundeliegenden Prozesse und Einflüsse müssen valide und vergleichbare Daten zu diesen SchülerInnen erhoben werden. Deshalb werden SchülerInnen mit SPF verstärkt in Large-Scale-Assessments (LSA) im Bildungsbereich einbezogen. Allerdings wurden die eingesetzten Testinstrumente und Fragebögen i.d.R. nicht für diese Schülergruppe entwickelt. Es besteht Forschungsbedarf dahingehend, ob die Testinstrumente für SchülerInnen mit SPF geeignet sind und welche Anpassungen gegebenenfalls notwendig sind (Kuhl et al., 2015). Um diese methodischen und inhaltlichen Aspekte zu untersuchen, hat sich 2012 eine interdisziplinäre Gruppe von KollegInnen zum Netzwerk NELSEN (NETwork of Large-scale-studies including students with Special Educational Needs) zusammengeschlossen.

In diesem Symposium werden vier Beiträge aus dem NELSEN-Netzwerk vorgestellt. Zwei der vier Beiträge (Böhme, Kohrt, Haag & Weirich; Kohrt, Gresch & Haag) basieren auf einem aktuellen LSA, dessen Ergebnisse zum Zeitpunkt der Einreichung dieses Symposiums noch nicht berichtet werden dürfen. Für die beiden weiteren Beiträge (Lüke & Grosche; Ahders & Kuhl) ist die Datenerhebung noch nicht gänzlich abgeschlossen, so dass hier vorerst nur teilweise Ergebnisse berichtet werden können. In allen Fällen würden die Präsentationen auf der GEBF-Konferenz belastbare Ergebnisse beinhalten.

Den Schwerpunkt dieses Symposiums bearbeiten die vier Beiträge aus unterschiedlichen Perspektiven. Neben zwei methodischen sind auch zwei inhaltlich ausgerichtete Beiträge vorgesehen. Der erste Beitrag von Lüke und Grosche stellt die Entwicklung und Validierung eines Instrumentes zur Erfassung der professionsunabhängigen Einstellung zum inklusiven Schulsystem (PREIS) vor. Somit werden die gesamtgesellschaftlichen Anforderungen von Inklusion aufgegriffen und geprüft, ob und inwieweit mit diesem Instrument eine vergleichende Erfassung des relevanten Einflussfaktors „Einstellung gegenüber Inklusion“ bei verschiedenen AkteurlInnen im Bildungsbereich in LSAs umgesetzt werden kann.

Die Arbeit von Böhme et al. fokussiert auf die Kompetenztestung von SchülerInnen mit SPF. Vorgestellt wird ein Ansatz zur Entwicklung zugänglicher Aufgaben, der eine valide und vergleichbare Erfassung domänenspezifischer Kompetenzen für die Schülergruppen mit SPF und ohne SPF zum Ziel hat. Die psychometrischen Eigenschaften und Messäquivalenz dieser angepassten Testaufgaben werden im Vergleich zu den regulären Testaufgaben untersucht.

Das Thema der Umsetzung von schulischer Inklusion im Schulalltag und den damit verbundenen Herausforderungen für Lehrkräfte und Schulleitungen, diese Prozesse angemessen zu begleiten, werden im dritten Beitrag von Ahders und Kuhl aufgegriffen. Sie stellen das Fortbildungsprogramms „Jede(r) ist besonders (JIB)“ und vor und evaluieren deren Wirksamkeit. Mit Blick auf die schulische Lebenswelt, die evidenzbasiert in Richtung Inklusion im Alltag weiterentwickelt werden soll, ergänzt dieser Beitrag das Symposium um den Aspekt der Praxisrelevanz.

Der Beitrag von Kohrt et al. untersucht, inwieweit sich die soziale Integration der SchülerInnen mit SPF an allgemeinen Schulen von jenen Lernenden unterscheidet, die an Förderschulen unterrichtet werden und in welchem Zusammenhang die Klassenzusammensetzung mit der sozialen Integration der SchülerInnen mit SPF steht. Der vierte Beitrag komplettiert das Symposium, indem ein Blick auf soziale Prozesse im Klassenverbund gerichtet wird. Abschließend werden die Beiträge diskutiert, wobei ein Schwerpunkt auf den methodischen und inhaltlichen Herausforderungen beim Einbezug von SchülerInnen mit SPF in LSAs liegt.

*Beiträge des Symposiums*

### Konstruktion und Validierung einer Kurzska zur Erfassung der Professionsunabhängigen Einstellungen zum Inklusiven Schulsystem (PREIS-K)

**Timo Lüke, Michael Grosche**

Bergische Universität Wuppertal

Die Umsetzung von Inklusion im Schulsystem ist eine Reform, die nicht nur von Lehrkräften, sondern auch von Eltern und der Gesellschaft getragen werden muss. Bisher konzentriert sich die auf Inklusion bezogene Einstellungsforschung jedoch fast ausschließlich auf Lehrkräfte, obwohl die Einstellungen anderer Menschen ebenfalls hohe Relevanz für die inklusive Schulentwicklung haben (Schwab, Tretter & Gebhardt, 2013). Dies liegt vermutlich auch daran, dass die meisten Erhebungsinstrumente pädagogisches Fachwissen voraussetzen und daher eher auf Lehrkräfte zugeschnitten waren. Bisher fehlt ein valides Instrument, das für alle Zielgruppen von Inklusion in gleicher Weise verständlich und einsetzbar ist, um die Einstellungen der verschiedenen Akteure direkt und auf derselben verzerrungsfreien Skala miteinander zu vergleichen. Wir haben ein solches Instrument in einem mehrstufigen Verfahren entwickelt: Die Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS, Lüke & Grosche, under review), die sich durch ihre inklusionstheoretische Fundierung und inhaltliche Validität auszeichnet. Zur Überprüfung der Testgüte (Konfirmatorische Faktorenanalyse, Retest-Reliabilität, Extremgruppenvergleich) führten wir vier konsekutive Studien mit verschiedenen Stichproben durch. Insgesamt konnten wir zeigen, dass mit der PREIS erstmalig ein Instrument vorliegt, das inhaltswidrig, unter Verzicht auf Formulierungen der Zweigruppen-Theorie und ohne professionsspezifisches Fachwissen vorauszusetzen, faktoriell valide und reliabel das einfaktorielle Konstrukt „Einstellungen zum inklusiven Schulsystem“ misst. Das Instrument ist frei und offen unter einer CC-BY-SA-Lizenz verfügbar (Lüke & Grosche, 2017). Die PREIS besteht aus 14 Items, die anhand einer fünfstufigen Likert-Skala („stimme nicht zu“ – „stimme zu“) beantwortet werden und eine recht umfassende Instruktion benötigen. Die Gesamtskala ist mit einer Bearbeitungszeit von 5 bis 10 Minuten zwar recht ökonomisch, für den Einsatz in Large-Scale-Assessments – angesichts der Vielzahl zu erfassender Konstrukte – trotzdem zu lang. Daher haben wir basierend auf den Evaluationsstudien eine Kurzska (PREIS-K) mit lediglich sechs Items entwickelt. Zu den Daten der ersten vier Studien (N = 674) passt das Messmodell mit den ausgewählten sechs Items gut ( $\chi^2 = 47.72$ ,  $df = 9$ ,  $p < .001$ ,  $CFI = .98$ ,  $TLI = .97$ ,  $SRMR = .02$ ,  $RMSEA = .05$ ). In zwei weiteren konsekutiven Studien testen wir dieses Modell auf Passung an zwei unabhängigen Stichproben und untersuchen die Übereinstimmung zwischen der Gesamtskala und der Kurzska sowie die Auswirkung des Verzichts auf die ausführliche

Instruktion. Die Datenauswertungen beider Studien werden bis zur Konferenz abgeschlossen sein. Wir können dann spezifische Empfehlungen zum Einsatz des Instruments in Large-Scale-Assessments geben. Damit könnten die Einstellungen zum inklusiven Schulsystem valide und professionsübergreifend erfasst werden.

## **Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf**

**Katrin Böhme<sup>1</sup>, Pauline Kohrt<sup>2</sup>, Nicole Haag<sup>2</sup>, Sebastian Weirich<sup>2</sup>**

<sup>1</sup>Universität Potsdam, <sup>2</sup>Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Theoretischer und empirischer Hintergrund

Studien zum nationalen und internationalen Bildungsmonitoring dienen dazu, den Erfolg von Bildungsangeboten auf der Basis empirischer Daten bewerten zu können. Bildungserfolg bemisst sich unter anderem an den von Schülerinnen und Schülern erworbenen Kompetenzen. Auch Schülerinnen und Schüler mit sonderpädagogischen Förderbedarfen (SPF) sollten Zielgruppe des Bildungsmonitorings sein (Elliot, Davies & Kettler, 2012), um beurteilen zu können, welche Qualität Bildungs- und Förderangebote für diese Schülerschaft haben. Empirische Studien belegen aber, dass reguläre Testaufgaben, die für die Kompetenzmessung von Schülerinnen und Schülern ohne besondere Förderbedarfe entwickelt wurden, für die Messung der Kompetenzen von Schülerinnen und Schülern mit SPF nur bedingt geeignet sind (vgl. Kocaj et al., 2016; Müller, 2014; Südkamp et al., 2015). Die regulären Testaufgaben sind in der Regel deutlich zu schwer, was zu einer unzureichenden Differenzierung führt. Kritisch ist ferner, dass der Anteil fehlender Werte häufig sehr groß ist. Um auch über Schülerinnen und Schüler mit SPF differenzierte Aussage treffen zu können, sollten daher in Studien des Bildungsmonitorings angepasste Testaufgaben eingesetzt werden, die den besonderen kognitiven und emotionalen Bedürfnissen dieser Zielgruppe besser gerecht werden können als die bislang zur Kompetenzmessung verwendeten Aufgaben (Elliot, Davies & Kettler, 2012; Weigert, 2012).

Abgeleitete Handlungsimplicationen

Für den IQB-Bildungstrend 2016 im Primarbereich wurden angepasste Testaufgaben konstruiert, die für Grundschul Kinder mit SPF besser zugänglich und insgesamt eher leicht sein sollten. Die bei der Testkonstruktion verwendeten Strategien orientierten sich unter anderem an Erkenntnissen zur Entwicklung zugänglicher Leistungstests (Elliot, Kettler, Beddow & Kurz, 2011) und greifen auf die Theorie der kognitiven Belastung (Cognitive Load Theory, CLT; Chandler & Sweller, 1991) und die Theorie der Zugänglichkeit (Accessibility Theory; Beddow, 2012) zurück. Die Umsetzung dieser Prinzipien soll im Vortrag illustriert werden.

Fragestellung

Auf Basis der Daten des IQB-Bildungstrends 2016 untersuchen wir, ob die angepassten Testaufgaben die Kompetenzen von Grundschulkindern mit SPF besser erfassen als reguläre Testaufgaben. Dies sollte sich sowohl anhand günstigerer psychometrischer Eigenschaften als auch anhand geringerer Anteile fehlender Werte in den angepassten Testaufgaben zeigen. Ferner soll die Frage beantwortet werden, ob beide Aufgabensets die in den Bildungsstandards beschriebenen Kompetenzkonstrukte äquivalent erfassen.

Methode

Die Entwicklung und Pilotierung der angepassten Testaufgaben erfolgten im Jahr 2015. Im IQB-Bildungstrend 2016 wurden im Fach Deutsch 6 angepasste Testaufgaben im Kompetenzbereich Lesen und 7 angepasste Testaufgaben im Kompetenzbereich Zuhören eingesetzt. Für das Fach Mathematik wurden 37 angepasste Testaufgaben ausgewählt. In einer Gesamtstichprobe von 29 000 Schülerinnen und Schülern bearbeiteten in einem Multi-matrix-Design zwischen 243 und 5531 Schülerinnen und Schülern diese angepassten Aufgaben gemeinsam mit ausgewählten regulären Testaufgaben. Beide Aufgabensets (angepasste und reguläre Testitems) wurden für drei Schülergruppen jeweils separat analysiert. Diese Gruppen sind (1) Kinder ohne SPF, (2) Kinder mit SPF an allgemeinen Schulen und (3) Kinder mit SPF an Förderschulen. Als Messmodell wurde jeweils ein eindimensionales Rasch-Modell zugrunde gelegt. Folgende psychometrische Eigenschaften wurden untersucht: die mittlere Schwierigkeit der Items pro Gruppe (p-Werte), ihre Trennschärfen und ihre Passung zum Rasch-Modell (Infit). Zusätzlich wurde der Anteil fehlender Werte für beide Aufgabensets geprüft. Ferner wurde anhand latenter Korrelationen untersucht, ob beide Aufgabensets dasselbe Konstrukt erfassen (Kolen & Brennan, 2004).

Ergebnisse

Ergebnisse können zum aktuellen Zeitpunkt nicht vorgestellt werden, da die Datengrundlage der Analysen – der IQB-Bildungstrend 2016 in der Primarstufe – erst im Oktober 2017 der Öffentlichkeit vorgestellt wird.

## **Jede(r) ist besonders: Evaluation einer Lehrerprofessionalisierungsmaßnahme zur Unterstützung inklusiver Schulentwicklung**

**Lea Ahders, Poldi Kuhl**

Leuphana Universität Lüneburg

Im Zuge der zunehmenden schulischen Inklusion in Deutschland werden viele Lehrkräfte mit Herausforderungen konfrontiert, auf die sie im Rahmen ihres Lehramtsstudiums nur selten vorbereitet wurden. Dem großen subjektiv geäußerten Fortbildungsbedarf von Lehrkräften in den Bereichen „Sonderpädagogische Förderung“, „Umgang mit Heterogenität“ und „Individualisierung/ Binnendifferenzierung“ steht bislang jedoch kein angemessenes Fortbildungsangebot gegenüber (Richter, Kuhl & Pant, 2012). Die Wirkung von Fortbildungsangeboten, die sich an einzelne Lehrkräfte richten, ist oft begrenzt, da sich der Transfer der Fortbildungsinhalte in die Schulen als problematisch erweist (z.B. Burke & Hutchins, 2007). Insbesondere erfolgreiche inklusive Schulentwicklung zeichnet sich jedoch dadurch aus, dass nicht die einzelne Lehrkraft mit ihrem Unterricht, sondern die Schule als komplexes System in den Blick genommen wird (Prenzel, 2013).

Diesen Befunden Rechnung tragend, hat das Fortbildungsprogramm „Jede(r) ist besonders (JIB)“ der Deutschen Schulakademie zum Ziel, den je drei Mitgliedern der Schulleitungsteams aus 25 teilnehmenden Schulen (N=75) aus verschiedenen Bundesländern gute Praxis zum Umgang mit Inklusion und Heterogenität im Unterricht zu vermitteln. Über einen Zeitraum von zwei Jahren und bestehend aus insgesamt acht Modulen, beispielsweise zur Unterrichtsgestaltung, zum Schulleben und zur außerschulischen Vernetzung, nehmen die Teams aus den teilnehmenden Schulen an der Fortbildungsmaßnahme teil. Langfristiges Ziel ist es, in den Schulen Maßnahmen für eine inklusive Schulentwicklung anzustoßen, die der Ausgangslage und den Bedürfnissen der jeweiligen Schule entsprechen.

Begleitet wird die Fortbildungsmaßnahme durch eine wissenschaftliche Evaluation, die folgende Ziele verfolgt: (1) die Wirksamkeit der Fortbildungsmaßnahme für Schulentwicklungsprozesse an den teilnehmenden Schulen empirisch zu

überprüfen sowie (2) die Schulentwicklungsprozesse an diesen Schulen nachzuzeichnen und so Gelingensbedingungen und auch Barrieren erfolgreicher inklusiver Schulentwicklung abzuleiten. Um dem nachgehen zu können, wird im Rahmen der wissenschaftlichen Begleitevaluation ein längsschnittlicher Mixed-Methods-Ansatz verfolgt, der Online-Befragungen der Kollegien, systematische Beobachtungen ausgewählter JIB-Module sowie ggf. Besuche und Gruppendiskussionen an ausgewählten Teilnehmerschulen umfasst.

Im Beitrag wird das Konzept der wissenschaftlichen Begleitevaluation zu JIB vorgestellt. Zudem werden Ergebnisse aus der ersten Online-Befragung berichtet, die im September 2017 stattfand. Diese erste Erhebung beinhaltet einen Schulleiterfragebogen zur Erfassung der strukturellen Bedingungen an den der teilnehmenden Schulen und darüber hinaus einen Fragebogen für die Kolleg/innen an den Schulen. Dieser erfragt verschiedene Aspekte, bspw. Kooperation, Innovationsbereitschaft, Selbstwirksamkeit, die sich im Hinblick auf gelingende Schulentwicklungsprozesse als bedeutsam erwiesen haben. Die Ergebnisse der ersten Befragung werden vor dem Hintergrund der Herausforderungen schulischer Inklusion diskutiert.

## **Soziale Integration von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an allgemeinen Schulen und Förderschulen: Die Rolle individueller Einflussfaktoren und der Klassenzusammensetzung**

**Pauline Kohrt, Cornelia Gresch, Nicole Haag**

Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) werden zunehmend auch an allgemeinen Schulen gemeinsam mit Schülerinnen und Schülern ohne SPF unterrichtet. Ein Ziel dieses gemeinsamen Unterrichts ist die soziale Integration der Schülerinnen und Schüler mit SPF, die wiederum relevant für die psychische Gesundheit und erfolgreiches Lernen ist (Becherer, Köller & Zimmermann, 2017). Inwiefern dies tatsächlich gelingt, wird allerdings kontrovers diskutiert (siehe auch Huber, 2009): In der US-amerikanischen und kanadischen Literatur wird von einer höheren sozialen Integration von Schülerinnen und Schülern mit SPF in inklusiven Settings im Vergleich zu separierten Formen des Unterrichts berichtet (Oh-Young & Filler, 2015, Wiener & Tardif, 2014), während Bakker et al. (2007) in den Niederlanden keine Unterschiede feststellen konnten. In Deutschland besteht zur sozialen Integration von Schülerinnen und Schülern mit SPF an allgemeinen Schulen verglichen mit jenen an Förderschulen noch Forschungsbedarf; gleichfalls ist wenig über die Hintergründe erfolgreicher oder misslingender sozialer Integration im Klassenverband bekannt.

Folgt man sozialpsychologischen Theorien zum Entstehungskontext sozialer Integration, spielen auf der einen Seite individuelle Merkmale eine zentrale Rolle: Schülerinnen und Schüler mit besseren schulischen Leistungen (Haeberlin, 1999; Huber & Wilbert, 2012), normkonformem Verhalten und positiven Interaktionen mit den Klassenkameraden (Coie, Dodge & Kupersmidt, 1990) zeigen eine höhere soziale Integration. Nicht-normkonformes Verhalten, wie beispielsweise Aggressivität, hängt hingegen negativ mit der sozialen Integration zusammen (Huber, 2006; Newcomb, Bukowski & Pattee, 1993).

Auf der anderen Seite ist auch der Kontext, in dem Prozesse der sozialen Integration stattfinden, bedeutsam. Der Theorie der sozialen Vergleichsprozesse (Festinger, 1954) folgend vergleichen Menschen ihre Leistungen und Fähigkeiten mit ähnlichen Menschen, um ein akkurates Selbstbild zu erhalten (Frey et al., 1993). Im Schulalltag finden diese Vergleichsprozesse innerhalb der Lerngruppe statt (Bezugsgruppentheorie; Merton & Rossi, 1968). Schülerinnen und Schüler vergleichen sich mit ihren Klassenkameradinnen und -kameraden und sind bestrebt, eventuell bestehende Diskrepanzen aufzulösen. Misslingt dies, könnte das zu einer geringeren sozialen Integration in ihrer Lerngruppe führen. Übertragen auf die Beschulung in Förderschulen oder allgemeinen Schulen ist anzunehmen, dass Schülerinnen und Schüler mit SPF in Förderschulen eine höhere Ähnlichkeit zu ihren Mitschülerinnen und Mitschülern aufweisen, entsprechend günstigere Vergleiche anstellen und eine höhere soziale Integration aufweisen als in allgemeinen Schulen.

Diese theoretischen Annahmen werden im vorliegenden Beitrag betrachtet: Im Vergleich zwischen Förderschulen und allgemeinen Schulen wird untersucht, ob sich die soziale Integration von Schülerinnen und Schülern mit SPF nach Schulform unterscheidet und ob diese Unterschiede durch die Interaktion zwischen individuellen Merkmalen und der Klassenzusammensetzung erklärt werden können. Als Datengrundlage wird der bundesweit in der 4. Jahrgangsstufe durchgeführte IQB-Bildungstrend 2016 (Stanat et al. 2017) genutzt. Die soziale Integration wurde im Selbstbericht von den Schülerinnen und Schülern mit SPF an allgemeinen Schulen (N = 1260) und an Förderschulen (N = 1196) erhoben und es werden mit standardisierten Leistungstests erhobene Kompetenzdaten der Bereiche Lesen und Mathematik in die Analysen einbezogen. Anhand von Mehrebenenmodellen wird analysiert, inwiefern sich die wahrgenommene soziale Integration von Schülerinnen und Schülern mit SPF zwischen allgemeinen Schulen und Förderschulen unterscheidet. Auf Klassenebene werden verschiedene Merkmale und Heterogenitätsmaße berücksichtigt: die individuelle Abweichung von der mittleren Klassenleistung, der Anteil der Schülerinnen und Schüler mit SPF und die individuelle Abweichung vom mittleren sozioökonomischen Status in der Klasse. Auf Schülebene werden zudem die kognitiven Grundfähigkeiten und der sozioökonomische Status kontrolliert.

Ergebnisse können zum aktuellen Zeitpunkt nicht vorgestellt werden, da die Datengrundlage der Analysen – der IQB-Bildungstrend 2016 in der Primarstufe – erst im Oktober 2017 der Öffentlichkeit vorgestellt wird.

Die Studie leistet einen weiterführenden Beitrag zu der Frage nach den Folgen inklusiver Beschulung für die soziale Integration von Schülerinnen und Schülern mit SPF.

## A15

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H116

### **Sozio-emotionale Prozesse von Schülerinnen und Schülern: Eine intraindividuelle Perspektive auf Emotionen im Schulkontext**

Chair(s): **Melanie Keller** (IPN Kiel, Deutschland), **Judith Dirk** (Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a. M.)

DiskutantIn(nen): **Anne Frenzel** (Ludwig Maximilians Universität München)

Klassenzimmer sind Orte intensiver sozialer Interaktionen, die auch vom Affekt und den Emotionen ihrer jeweiligen Akteure bestimmt sind. Zunehmende Forschungsarbeiten der letzten Jahre untermauern die Relevanz sozio-emotionaler Faktoren z.B. für Lernerfolge von Schülerinnen und Schülern (z.B. Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017).

Obwohl hierbei zugrundeliegende theoretische Modelle zu einem großen Teil auf intraindividuellen Annahmen beruhen (z.B. Kontroll Wert-Theorie für Lern- und Leistungsemotionen, Pekrun, 2006), werden sie in empirischen Studien oftmals ausschließlich interindividuell überprüft. Dabei lassen sich interindividuell gefundene Zusammenhänge zwischen Variablen nicht notwendigerweise auf intraindividuelle Prozesse übertragen (z.B. Molenaar & Campbell, 2009; Schmitz, 2006). In klassischen längsschnittlichen Studien können zwar intraindividuelle Entwicklungsverläufe modelliert werden, jedoch werden in derartigen Studien oftmals Messzeitpunkte auf großen Zeitskalen (einmal pro Jahr in fünf aufeinanderfolgenden Schuljahren, Pekrun et al., 2017) realisiert, was wiederum die Modellierung kurzfristiger Interaktions- und Entwicklungsprozesse, die z.B. innerhalb einer Schulstunde oder über einen Schultag hinweg geschehen, ausschließt. Demgegenüber steht, dass insbesondere Affekt einerseits von Situation zu Situation, d.h. innerhalb von Personen schwankt, was in der Forschung seit langem bekannt ist (Wessman & Ricks, 1966), und andererseits, dass retrospektiv erfasste Emotionen durch u.a. Erinnerungsverzerrung (recall bias; Robinson & Clore, 2002) in ihrer Validität eingeschränkt sein können.

Eine Möglichkeit, sozio-emotionale Prozesse intraindividuell valide zu modellieren ist durch sogenannte Experience Sampling oder Tagebuchstudien. Hierbei werden ähnlich zu Längsschnittstudien aber auf kürzeren Zeitskalen mehrere Messzeitpunkte pro Person realisiert; die zu messenden Variablen allerdings werden situativ, d.h. im Moment (Experience Sampling), oder mit Bezug auf das Erleben eines Tages (Tagebuchstudie) erfasst. Dadurch (1) ist die Validität bei der Messung der Konstrukte potentiell erhöht, (2) können relevante sozio-emotionale Prozesse, die z.B. von Schulstunde zu Schulstunde oder Schultag zu Schultag ablaufen, erfasst werden, sowie (3) anschließend intraindividuelle von interindividuellen Zusammenhängen getrennt statistisch modelliert werden.

Das vorliegende Symposium bietet prototypische und innovative Studien, die jeweils intraindividuelle Prozesse in sozio-emotionalen Variablen im Schulkontext mithilfe der Experience Sampling Methode bzw. in Tagebuchstudien untersuchen. Die Studien basieren auf Selbstberichten von Schülerinnen und Schülern und widmen sich entweder deren Wahrnehmung der Lernumwelt (Schmidt, Dirk, & Schmiedek: Soziale Eingebundenheit; Keller & Becker: Lehreremotionen), oder individueller motivationaler Faktoren (Dirk & Schmiedek: Soziale Ziele), oder beidem (Götz & Lütke: Unterrichtsgestaltung, individuelle Appraisal) und untersuchen jeweils den Einfluss dieser Faktoren auf affektive Variablen. Methodisch wurden die Studien entweder mithilfe der Experience Sampling Methode und einer momentanen Erfassung im Unterricht (Keller & Becker, Götz & Lütke) oder als Tagebuchstudie mit ein- oder mehrmaliger Erfassung pro Schultag (Schmidt et al., Dirk & Schmiedek) durchgeführt und die gewonnenen Daten anschließend auf intraindividuelle Ebene analysiert.

Zusammenfassend berücksichtigt das Symposium somit sowohl die wahrgenommene Lernumwelt als auch individuelle motivationale Faktoren, die zur Entstehung von Emotionen bei Schülerinnen und Schülern beitragen. Darüber hinaus zeigen die Beiträge das Potential intraindividuelle Analysen für das Verständnis sozio-emotionaler Prozesse auf. Nicht zuletzt können die alltagsnahen Untersuchungen sozio-emotionaler Prozesse im Schulkontext als Grundlage für spezifische Interventionen genutzt werden: Das Wissen um diese sozio-emotionalen Prozesse kann professionelles Handeln unterstützen, indem Lehrkräfte Einflussfaktoren auf emotionales Erleben von Schülerinnen und Schülern (er)kennen und ihren Unterricht entsprechend adaptiv gestalten.

*Beiträge des Symposiums*

### **Das Zusammenspiel von sozialer Eingebundenheit, Stimmung und Selbstwert: Eine intraindividuelle Perspektive auf das Erleben von Schülerinnen und Schülern**

**Andrea Schmidt<sup>1</sup>, Judith Dirk<sup>1</sup>, Florian Schmiedek<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a. M., <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a. M., Goethe Universität Frankfurt a. M.

Zahlreiche Studien haben die Bedeutsamkeit der sozialen Eingebundenheit für das affektive Wohlbefinden und Selbstwertgefühl von Kindern aufgezeigt (z.B. Holder & Coleman, 2007; Reynolds & Repetti, 2008). Dabei handelt es sich jedoch meist um querschnittliche Befunde zu interindividuellen Unterschieden der sozialen Eingebundenheit oder längsschnittliche Befunde zu längerfristigen Veränderungen. In welchem Ausmaß tägliche Schwankungen in sozialer Eingebundenheit von Grundschulern vorliegen und wie diese mit dem aktuellen Affekt und Selbstwert im Alltag zusammenhängen, ist hingegen bisher kaum untersucht worden. So konnten Reynolds & Repetti (2008) in einer Tagebuchstudie zeigen, dass Kinder, die mehr Probleme mit Mitschülern berichteten, im Mittel einen niedrigeren Selbstwert und einen höheren negativen Affekt angaben, als Kinder, die weniger Probleme berichteten. Unklar bleibt jedoch, ob diese interindividuellen Zusammenhänge intraindividuelle Prozesse widerspiegeln: Berichten Kinder an Tagen, an denen sie sich stärker sozial eingebunden fühlen als für sie üblich, auch ein höheres Selbstwertgefühl und einen niedrigeren negativen Affekt? Zudem gibt es bisher kaum Studien, die das Zusammenspiel von sozialer Eingebundenheit und Affekt bei der Vorhersage des Selbstwertgefühls betrachten.

Ziel der vorliegenden Studie war es, Zusammenhänge zwischen der sozialen Eingebundenheit in die Klassengemeinschaft, dem Affekt und dem Selbstwertgefühl sowohl auf interindividueller als auch intraindividuelle Ebene bivariat und multivariat zu untersuchen. In einem ersten Schritt wurde geprüft, ob die Voraussetzungen für die Untersuchung intraindividuelle Prozesse gegeben sind und die drei Konstrukte Schwankungen innerhalb von Kindern aufweisen, die sich reliabel erfassen lassen. In weiteren Schritten wurde untersucht, ob Schwankungen in der sozialen Eingebundenheit, im Affekt und im Selbstwertgefühl systematische tägliche Zusammenhänge aufweisen

Im Rahmen einer Smartphone-basierten Experience Sampling Studie bearbeiteten 55 Kinder (9-12 Jahre) an 9 aufeinanderfolgenden Schultagen einmal täglich Kurzskaleten zu ihren Erfahrungen der sozialen Eingebundenheit in die Klassengemeinschaft (z.B. „Heute habe ich mich gut mit den anderen Kindern in meiner Klasse verstanden“), ihrem Selbstwertgeföhl (z.B. „Heute war ich völlig zufrieden mit mir.“) sowie ihrem positiven Affekt (z.B. „Ich fühle mich gerade gut gelaunt“) und negativen Affekt (z.B. „Ich fühle mich gerade unglücklich“; Leonhardt, Könen, Dirk & Schmiedek, 2016).

Ergebnisse einer konfirmatorischen Mehrebenenfaktorenanalyse zeigten, dass die Kurzskaleten sowohl interindividuelle Unterschiede als auch intraindividuelle Schwankungen in der sozialen Eingebundenheit, im Affekt und im Selbstwertgeföhl reliabel erfassen können. Die mittleren Ausprägungen des positiven Affekts, des Selbstwertgeföhls und der sozialen Eingebundenheit waren hoch, die mittlere Ausprägung des negativen Affekts war niedrig. Die Kinder fühlten sich im Durchschnitt somit eher gut gelaunt, berichteten eher einen hohen Selbstwert und fühlten sich eher gut eingebunden in die Klassengemeinschaft.

Mehrebenenanalysen zeigten weiterhin systematische Zusammenhänge zwischen täglichem negativen und positiven Affekt und sozialer Eingebundenheit: An Tagen, an denen Kinder sich stärker sozial eingebunden fühlten als für sie üblich, gaben sie einen höheren positiven Affekt an und in der Tendenz einen niedrigeren negativen Affekt. Analog hingen tägliches Selbstwertgeföhl und soziale Eingebundenheit systematisch zusammen: An Tagen, an denen Kinder sich stärker sozial eingebunden fühlten als für sie üblich, gaben sie auch ein höheres Selbstwertgeföhl an. Bezüglich des Zusammenspiels von sozialer Eingebundenheit und Affekt konnte gezeigt werden, dass Affekt und soziale Eingebundenheit sowohl Varianz teilen als auch bedeutsame unique Beiträge bei der Vorhersage des Selbstwertgeföhls leisten.

Ergebnisse bisheriger Studien zum mittleren Zusammenhang von Geföhlen der sozialen Eingebundenheit mit Selbstbewertungen und Affekt konnten repliziert werden. Darüber hinaus konnte gezeigt werden, dass tägliche Variationen in der sozialen Eingebundenheit in die Klassengemeinschaft bei Kindern der vierten bis sechsten Klassenstufe positiv mit selbstberichteten Affekt und Selbstwertgeföhl an demselben Tag zusammenhängen. Die vorliegende Studie betont die Bedeutsamkeit der wahrgenommenen sozialen Eingebundenheit im Schulkontext für das Wohlbefinden von Kindern. Lehrern wird empfohlen, die Klassengemeinschaft in sozialer Hinsicht, z.B. durch kooperative Spiele, zu fördern und darauf zu achten, dass im Schulalltag alle Kinder sozial eingebunden werden.

### **Authentizität von Lehreremotionen: Intraindividuelle Zusammenhänge mit Schüleremotionen und emotionaler Ansteckung im Unterricht**

**Melanie Keller<sup>1</sup>, Eva Becker<sup>2</sup>**

<sup>1</sup>Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel, <sup>2</sup>Universität Zürich

Emotionen von Schülerinnen und Schülern im Unterricht hängen von vielfältigen Einflussfaktoren ab (Pekrun, 2006). Dabei spielen auch die Emotionen der Lehrpersonen eine Rolle: Es wurde vielfach gezeigt, dass Emotionen zwischen Interaktionspartnern „überspringen“ können (emotional crossover: Bakker & Demerouti, 2013; emotional contagion: Hatfield, Cacioppo, & Rapson, 1993). Derartige Transmissionsprozesse von Lehrpersonen auf Schülerinnen und Schüler konnten bereits für verschiedene Emotionen gezeigt werden (Becker, Goetz, Morger, & Ranellucci, 2014; für Angst: Beilock, Gunderson, Ramirez, & Levine, 2010; für Freude: Frenzel, Becker-Kurz, Pekrun, Goetz, & Lüdtke, im Druck).

Es ist gut belegt, dass Lehrpersonen ihre Emotionen bzw. ihren Emotionsausdruck regulieren (z.B. Jiang, Vauras, Volet, & Wang, 2016). Es ist jedoch anzunehmen, dass neben der gezeigten Emotion, auch die emotionale Authentizität einen Einfluss auf die emotionalen Reaktionen der Schülerinnen und Schülern hat. Aus experimenteller Laborforschung ist bekannt, dass inauthentische von authentischen Emotionsausdrücken unterschieden werden können (McLellan, Johnston, Dalrymple-Alford, & Porter, 2010) und auch, dass wahrgenommene emotionale Authentizität das Verhalten und die Emotionen des Rezipienten beeinflusst (Johnston, Miles, & Macrae, 2010; Surakka & Hietanen, 1998). Im Lehr-Lernkontext gibt es erste Befunde, die darauf hinweisen, dass inauthentisch gezeigter Lehrereithusiasmus Schülerinnen und Schüler in ihren emotionalen Reaktionen beeinflusst (Keller, Becker, Frenzel, & Taxer, eingereicht). Im Bereich diskreter Emotionen gibt es jedoch bisher fast ausschließlich Forschung aus der Perspektive der Lehrperson (z.B. Taxer & Frenzel, 2015), so dass die Frage offen ist, inwieweit emotionale Authentizität für Schülerreaktionen eine Rolle spielt.

Die vorliegende Studie hat zum Ziel, emotionale Authentizität von Lehrpersonen in der Schülerwahrnehmung und ihren Einfluss auf die Schüleremotionen sowie den emotionalen Ansteckungsprozess zwischen Lehrer- und Schüleremotionen zu untersuchen. Die Zusammenhänge werden dabei auf Situationsebene, d.h. innerhalb von Schülerinnen und Schülern über die Zeit analysiert, da Emotionen zeitlich stark fluktuierende Phänomene sind (z.B. Becker et al., 2014) und da die Sensitivität bei der Erkennung von Emotionen anderer bzw. deren emotionaler Authentizität interindividuelle Unterschiede aufweist (McLellan et al., 2010).

Die Studie wurde mithilfe der Experience Sampling-Methode (Bolger & Laurenceau, 2013) durchgeführt. Drei bis vier zufällig ausgewählte Schülerinnen und Schüler (N = 149) pro Klasse berichteten über 10 Schultage hinweg zu randomisierten Zeitpunkten im Unterricht (N = 2668 Messzeitpunkte) über eigene Emotionen und wahrgenommene Lehreremotionen (jeweils Freude, Ärger, und Angst) sowie über die wahrgenommene emotionale Authentizität. Hierbei wurde auf Einzelitems bei der Erfassung der Konstrukte zurückgegriffen (z.B. Schülerfreude: „Im Moment freue ich mich“ oder Authentizität: „Die Geföhle, die meine Lehrperson im Moment zeigt, wirken echt.“). Die Items wurden jeweils am Personenmittelwert zentriert und schrittweise Regressionsmodelle auf intraindividuelle Ebene mithilfe des Statistikprogramms Mplus (Version 7.11; Muthén & Muthén, 1998-2012) modelliert.

Deskriptiv zeigte sich, dass die wahrgenommene emotionale Authentizität der Lehrperson hoch war, jedoch stark von Unterrichtsstunde zu Unterrichtsstunde, d.h. innerhalb von Schülerinnen und Schülern, schwankte. Situationen, in denen Schülerinnen und Schüler mehr Authentizität berichteten, gingen einher mit mehr Freude und weniger Ärger und Angst von Schülerinnen und Schülern. Obwohl Authentizität jenseits wahrgenommener Lehreremotionen Varianz in Schüleremotionen aufklärte, nahm sie keine Moderatorfunktion für den Prozess der emotionalen Ansteckung ein.

Unsere Befunde zeigen, dass die Emotionen von Schülerinnen und Schülern auf intraindividuelle Ebene von der emotionalen Authentizität abhängen; die vorherrschende Prämisse auf Lehrerseite „Don't show them!“ (Sutton, 2004, S. 379) wenn es um die Regulierung negativer Emotionen geht, ist also einerseits aus Gründen der emotionalen Ansteckung zu begrüßen. Andererseits muss diese Prämisse aber mit Vorsicht betrachtet werden, wenn derart regulierte Emotionen von den Schülerinnen und Schülern als inauthentisch wahrgenommen werden. Da die vorliegende Studie ausschließlich auf Schüler selbstberichte zurückgriff, sollte zukünftige Forschung das von Lehrpersonen selbst eingeschätzte emotionale Erleben und den Emotionsausdruck inkludieren, um so die Antezedenzen für Authentizität auf Lehrerseite sowie deren Wahrnehmung und Konsequenzen auf Schülerseite bestimmen zu können.

## **Die Erfassung sozialer Ziele im Schulkontext: Faktorstruktur inter- und intraindividuelle Unterschiede und Zusammenhänge mit Affekt**

**Judith Dirk<sup>1</sup>, Florian Schmiedek<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a. M., <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a. M., Goethe Universität Frankfurt a. M.

Soziale Ziele motivieren und steuern Verhalten in sozialen Situationen und beziehen sich auf die Erreichung erwünschter sozialer Zielzustände, wie z.B. eine Freundschaft zu vertiefen oder mit anderen Menschen in einer bestimmten Art und Weise zu interagieren, beispielsweise einen guten Eindruck bei jemandem zu hinterlassen. Die sozialen Ziele von Kindern und Jugendlichen sind ein wichtiger Prädiktor ihrer sozialen Anpassung und beeinflussen ihr Sozialverhalten (Ryan & Shim, 2008) und ihren Affekt (Shim, Wang & Cassidy, 2013). Basierend auf Arbeiten von Ryan und Shim (2008) sowie in Anlehnung an Arbeiten zu Lern- und Leistungszielen (Elliot & McGregor, 2001) kann zwischen drei Dimensionen von sozialen Zielen unterschieden werden: soziale Entwicklungsziele, soziale Annäherungs- und soziale Vermeidungszielen. Trotz einiger Arbeiten zu sozialen Zielen bei Kindern und Jugendlichen sind diese im deutschen Sprachraum noch wenig untersucht. Zudem liegt der Fokus bei den meisten Arbeiten bisher auf der Untersuchung von Zielen im Sinne von traits und somit der Annahme, dass soziale Ziele eher situationsunabhängig und zeitlich stabil sind. Einige wenige Arbeiten zeigen jedoch, dass die sozialen Ziele von Kindern in Abhängigkeit des Kontexts variieren können (z.B. Ojanen, Aunola, & Salmivalli, 2007) und somit im Sinne von states intraindividuelle Variabilität aufweisen. Offen ist, inwiefern sich intraindividuelle Variabilität von sozialen Zielen auch alltagsnah im Schulkontext erfassen lässt und inwiefern bereits Kinder am Ende der Grundschule über ihre sozialen Ziele Auskunft geben können. Wir nehmen an, dass Kinder am Ende der Grundschule soziale Ziele berichten können und dass diese Ziele intraindividuelle Variabilität aufweisen, d.h. dass sie tagtäglichen Schwankungen unterliegen. Des Weiteren erwarten wir, dass die verschiedenen sozialen Ziele differenzierte Zusammenhänge mit Affekt zeigen. Zum Beispiel sollten Kinder mehr positiven und weniger negativen Affekt berichten, wenn sie soziale Entwicklungsziele anstreben während soziale Vermeidungsziele eher mit mehr negativem und weniger positivem Affekt einhergehen sollten.

In der aktuellen Studie haben 112 Kinder in der vierten Klasse (9-11 Jahre, 40% Mädchen) zweimal täglich (am Vormittag in der Schule und am Nachmittag in der Freizeit) in Tagebüchern ihre aktuellen sozialen Ziele in Bezug auf ihre Gleichaltrigen (z.B. Heute Vormittag war es mir wichtig, dass mich andere Kinder mögen) sowie ihre momentane Stimmung (z.B. Heute Vormittag fühle ich mich gut gelaunt) berichtet. Die sozialen Ziele wurden mit insgesamt 30 Items erfasst, die die verschiedenen sozialen Ziele repräsentieren. Der Affekt wurde mit 12 Items erfasst, die in Tagebuchstudien mit Kindern der Altersgruppe entwickelt wurden (Leonhardt, Könen, Dirk & Schmiedek, 2016). Die Ergebnisse zeigen, dass Kinder am Ende der Grundschule mit den ausgewählten Items ihre sozialen Ziele berichten können und dass diese intraindividuelle Variabilität aufweisen. Konfirmatorische Mehrebenenfaktorenanalysen zeigten, dass die Items sowohl interindividuelle Unterschiede als auch intraindividuelle Variabilität abbilden können und die drei Zieldimensionen soziale Entwicklungsziele, soziale Annäherungs- und soziale Vermeidungsziele repliziert werden konnten. Des Weiteren berichteten Kinder in Situationen, in denen ihnen besonders soziale Entwicklungsziele wichtig waren, auch mehr positiven und weniger negativen Affekt.

Diese Ergebnisse replizieren Befunde von Ryan und Shim (2008) zur interindividuellen Struktur sozialer Ziele bei Kindern und Jugendlichen und zeigen, dass soziale Ziele darüber hinaus auch systematische intraindividuelle Variabilität von Tag zu Tag aufweisen, die mit Variation im Affekt einhergeht. Diese intraindividuellen sozioemotionalen Prozesse tragen zum Verständnis kindlichen Sozialverhaltens bei. Das Wissen um Stimmungsschwankungen und kurzfristige Änderungen des Sozialverhaltens bei Kindern ist informativ für Lehrkräfte und kann ihr professionelles Handeln unterstützen.

## **Kognitive Appraisals medieren den Effekt von Unterrichtsgestaltung auf die Emotionen von Schülerinnen und Schülern im Klassenzimmer: Ein intraindividuelle Ansatz**

**Thomas Götz<sup>1</sup>, Lüdtke Oliver<sup>2</sup>**

<sup>1</sup>Universität Konstanz und Pädagogische Hochschule Thurgau, <sup>2</sup>Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel

Kumulative empirische Evidenz deutet darauf hin, dass Emotionen in der Schule zahlreiche positive und negative Effekte auf Lernen und Leistung ausüben (z.B. Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017) – insofern ist Wissen zu den Antezedenzen von Lern- und Leistungsemotionen von hoher Relevanz. An Schulen kann die Unterrichtsgestaltung als eine zentrale Ursache der Emotionen von Schülerinnen und Schülern betrachtet werden. In Pekrun's (2006) Kontroll-Wert-Theorie zu Lern- und Leistungsemotionen wird angenommen, dass die Effekte von Unterrichtsvariablen auf Lern- und Leistungsemotionen durch kognitive Appraisals mediert werden. Als zentrale Appraisals werden hier subjektive Kontrolle und Valenz (intrinsische und extrinsische) genannt.

Obwohl vereinzelte empirische Studien im Einklang mit diesen Mediationsannahmen (z.B. Goetz, Pekrun, Hall, & Haag, 2006) stehen, sind folgende zentrale Fragen ungeklärt: (1) Zeigen sich die postulierten Zusammenhänge auf intraindividuelle Ebene? Während die in der Kontroll-Wert-Theorie beschriebenen Annahmen sich auf intraindividuelle Mechanismen beziehen, wurden diese in bisherigen Studien ausschließlich auf interindividuelle Ebene getestet – mit einer Ausnahme (Goetz, Lüdtke, Nett, Keller, & Lipnevich, 2013), in welcher ausschließlich Unterrichtsvariablen und Emotionen, d.h. keine medierenden Prozesse untersucht wurden. (2) Variiert die Mediationssalienz spezifischer Appraisals über Unterrichtsfacetten und Emotionen hinweg, d.h. sind bestimmte Appraisals für die Mediation spezifischer Unterrichts-Emotions-Konstellationen von besonderer Bedeutung? Die vorliegende Studie greift beide Themen auf, indem auf intraindividuelle Ebene Mediationseffekte unterschiedlicher Appraisals im Hinblick auf den Zusammenhang von Unterrichtsgestaltung und diskreten Emotionen untersucht wurden.

In einer deutschen und einer schweizerischen Stichprobe wurden jeweils 3-4 pro Klasse zufällig ausgewählte Schüler/innen (N = 122/149) mittels der Experience-Sampling-Methode (Bolger & Laurenceau, 2013) zu randomisierten Zeitpunkten während des Unterrichts befragt (N = 1518/2669 Messzeitpunkte). Erfasst wurden dabei wahrgenommene Unterrichtsgestaltung, kognitive Appraisals (Kontrolle, intrinsische und extrinsische Valenz) und emotionales Erleben (u.a. Freude, Ärger, Angst, und Langeweile).

Appraisals und Emotionen wurden dabei jeweils anhand von Einzelitems erfasst (z.B. Kontrolle: „Im Moment bin ich gut im Unterricht“ oder Freude: „Im Moment freue ich mich“). Die Unterrichtsgestaltung wurde mit mehreren Items erfasst, die zunächst mithilfe einer explorativen Faktorenanalyse auf intraindividuelle Ebene zu zwei Faktoren gruppiert wurden; einer der Faktoren repräsentierte eine verständliche und motivierende, und der andere Faktor eine schwierige und überfordernde Unterrichtsgestaltung. Anschließend wurden alle Variablen am Personenmittelwert zentriert und auf intraindividuelle Ebene



und getrennt nach Emotion und Sample Pfadmodelle modelliert, in denen beide Unterrichtsfaktoren jeweils die drei Appraisals und diese wiederum die jeweilige Emotion prädizierten.

In beiden Samples zeigten sich auf intraindividuellem Ebene klare und sehr ähnliche Zusammenhänge zwischen Unterrichtsgestaltung und dem emotionalen Erleben der Schülerinnen und Schülern (Goetz et al., 2013). Dieser Zusammenhang wurde in beiden Samples in Einklang mit den Annahmen partiell über die Appraisals mediiert; so führten beispielsweise – innerhalb eines/r Schülers/in – stärker als motivierend wahrgenommene Unterrichtssituationen zu mehr Kontrolle und hierüber zu mehr Freude als weniger motivierende Unterrichtssituationen.

Diese Mediation muss im Detail allerdings differenziert betrachtet werden: Je nach Emotion fungieren unterschiedliche Appraisals als Mediatoren für die Unterrichtsgestaltung. So zeigte sich, dass für Freude und Langeweile die Mediation im Wesentlichen über Kontrolle und intrinsische aber nicht extrinsische Valenz erfolgt; für Angst hingegen scheint eher extrinsische Valenz als Mediator für die Unterrichtsgestaltung von Bedeutung zu sein. Zusammengefasst zeigen also die Ergebnisse, dass intraindividuelle Prozesse bei Schülerinnen und Schülern mithilfe des Kontroll-Wert Modells zu Lern- und Leistungsemotionen (Pekrun, 2006) gut beschrieben werden können, jedoch eine stärkere theoretische Ausdifferenzierung des angenommenen Mediationsprozesses je nach Emotion und Appraisal notwendig ist.

## A16

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H115

### Lehrer-Schüler-Interaktionen: Erfassung, Korrelate und Intervention

Chair(s): **Bastian Carstensen** (Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel), **Uta Klusmann** (Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel)

DiskutantIn(nen): **Diana Raufelder** (Ernst-Moritz-Arndt-Universität Greifswald)

Lehrer-Schüler-Interaktionen spielen im Schulalltag eine zentrale Rolle: Neben lern- und leistungsbezogenen sowie motivationalen Merkmalen (Cornelius-White, 2007; Roorda et al., 2011) sind gelingende Interaktionen zwischen der Lehrkraft und ihren Schüler/-innen positiv mit der Entwicklung sozialer und emotionaler Kompetenzen von Schüler/-innen assoziiert (Ahnert et al., 2012; Pianta, Hamre & Stuhlman, 2003). Für Lehrkräfte wird angenommen, dass positive Interaktionen mit ihren Schüler/-innen eine Ressource für das subjektive Wohlbefinden darstellen (Spilt, Koomen & Thijs, 2011).

Gute Lehrer-Schüler-Interaktionen, die zu einer positiven Lehrer-Schüler-Beziehung beitragen, zeichnen sich dadurch aus, dass Lehrkräfte durch ihr Verhalten einerseits klare Strukturen schaffen und die Schüler/-innen lenken, dabei andererseits ihre Bedürfnisse berücksichtigen, mit ihnen kooperieren und sie fachlich und persönlich unterstützen (Wubbels & Brekelmans, 2005). Diese Verhaltensweisen spiegeln sich in Modellen der Unterrichtsqualität (Klieme, Schümer & Knoll, 2001; Pianta & Hamre, 2009) wider, insbesondere in den Dimensionen konstruktive Unterstützung und Klassenführung.

Trotz umfangreicher Forschung zur Interaktionsqualität zwischen Lehrkräften und Schüler/-innen bestehen noch einige Forschungsdesiderata: Erstens geschieht die Operationalisierung oftmals über globale Maße, welche konkrete Interaktionsverläufe und den transaktionalen Charakter der Beziehung nicht erfassen können (Knierim, Raufelder & Wettstein, 2017). Zweitens fokussieren Studien oftmals auf die Relevanz gelingender Interaktionen für die Schüler/-innen; Untersuchungen zu Zusammenhängen zwischen der Lehrer-Schüler-Interaktion und Outcomes seitens der Lehrkraft wären wünschenswert. Drittens ist wenig darüber bekannt, inwiefern man die Fähigkeiten von Lehrkräften zur Etablierung eines positiven Interaktionsklimas gezielt schon in der Ausbildung fördern kann.

Die 4 Beiträge dieses Symposiums thematisieren diese Forschungslücken wie folgt: Der erste Beitrag (Scherzinger et al.) widmet sich der Frage, welche konkreten sozialen Interaktionen eine positive Lehrer-Schüler-Beziehung fördern können. Dazu wird die State Space Grid Methode genutzt und mit traditionellen Erfassungsmethoden kombiniert. Der zweite Beitrag (Aldrup et al.) untersucht den Zusammenhang zwischen Lehrer-Schüler-Interaktionen und dem beruflichen Wohlbefinden von Lehrkräften. Der Fokus liegt dabei nicht nur auf Unterrichtsstörungen sondern darüber hinaus auf der emotionalen Qualität der Lehrer-Schüler-Beziehung. Der dritte (Carstensen et al.) und der vierte Beitrag (Hannemann et al.) thematisieren die Förderung lehrerseitiger Voraussetzungen für gelingende Lehrer-Schüler-Interaktionen. Hierbei handelt es sich zum einen um ein universitäres Training zur Förderung der sozial-emotionalen Kompetenz, zum anderen um ein Klassenführungstraining für angehende Lehrkräfte.

#### Beiträge des Symposiums

### Wie können angehende Lehrkräfte unterstützt werden, gelingende Interaktionen und Beziehungen zu den Schülerinnen und Schülern aufzubauen? Die Erfassung von Lehrer-Schüler-Interaktionen mit State Space Grids.

**Alexander Wettstein, Marion Scherzinger, Benjamin Roth**  
Pädagogische Hochschule Bern

#### Theoretischer Hintergrund

Gute Lehrer-Schüler-Beziehungen bilden die Grundlage für erfolgreiche Lehr-Lern-Prozesse. Gelingende Lehrer-Schüler-Beziehungen wirken sich positiv auf die Lernmotivation (Wentzel, 2010), die Schulleistungen (Pianta, Hamre & Stuhlman, 2003), die psychosoziale Entwicklung der Schülerinnen und Schüler und das Wohlbefinden der Lehrperson aus (Davis, 2003; den Brok, Brekelmans & Wubbels, 2004; Hattie, 2009; Hughes, 2012; Klem & Connell, 2004; Mainhard, in press; Martin & Dowson, 2009; Raufelder, 2010; Roorda, Koomen, Spilt, & Oort, 2011; Stipek, 2004, 2006; Wentzel & Wigfield, 2009; Wentzel, 1997, 2010, 2012; Wentzel, 2009, 2010; Wubbels & Brekelmans, 2005; Zimmer-Gembeck, Chipuer, Hanisch, Creed, & McGregor, 2006).

Angehende Lehrkräfte fühlen sich durch die Gestaltung sozialer Interaktionen sowie den Aufbau positiver Lehrer-Schüler-Beziehungen herausgefordert. Sie nennen die Aufrechterhaltung von Disziplin und die Etablierung positiver Lehrer-Schüler-Beziehungen als vordringliche Herausforderungen beim Berufseinstieg (De Jong, Mainhard et al. 2014; Fuller & Bown, 1975; Ghaith & Shaaban, 1999; Liston, Withcomb & Borko, 2006; Veenman, 1984). Dabei empfinden viele Berufsanfängerinnen und Berufsanfänger die Aufrechterhaltung von Disziplin und eine wertschätzende Lehrer-Schüler-Beziehung als Widerspruch (De Jong, Mainhard et al., 2014; McLaughlin, 1991; Weinstein, 1998).

Es stellt sich deshalb die Frage, wie angehende Lehrkräfte unterstützt werden können, gelingende Interaktionen und Beziehungen zu den Schülerinnen und Schülern aufzubauen und wie diese Inhalte in der Ausbildung von Lehrerinnen und Lehrern aufgegriffen werden. Methodisch werden Lehrer-Schüler-Beziehungen meist global über Schüler- oder Lehrerfragebogen auf einer Makroebene erfasst (z.B. „Ich mag meine Lehrperson“). Dabei bleibt oft unklar, wie sich Beziehungen in den sozialen Interaktionen auf einer Mikroebene zeigen und wo Interventionen für eine Verbesserung der Lehrer-Schüler-Beziehung ansetzen können. Zudem wird das Konstrukt der Lehrer-Schüler-Beziehung sehr heterogen gefasst und unterschiedlich operationalisiert (Knierim, Raufelder & Wettstein, 2017). Entsprechend werden mit sehr unterschiedlichen Instrumenten verschiedene Aspekte der Beziehung erfasst. Es besteht insbesondere ein Bedarf an Instrumenten, die 1.) dem transaktionalen Charakter der Lehrer-Schüler-Beziehung gerecht werden, 2.) die gleichzeitig Beziehungen und konkrete Interaktionsverläufe in den Fokus nehmen und so die Ableitung von Interventionen ermöglichen und die 3.) den ökologischen Kontext des Unterrichts (Nickel, 1976; Raufelder & Hoferichter, 2015; Wettstein, 2012) mitberücksichtigen.

#### Fragestellungen

In unserer Studie untersuchen wir, 1.) wie sich Lehrer-Schüler-Interaktionen auf mikrogenetischer Ebene gestalten und inwiefern sich die Interaktionsverläufe zwischen Gruppen mit unterschiedlicher Wahrnehmung der Lehrer-Schüler-Beziehungen auf makrogenetischer Ebene unterscheiden. 2.) Welches Potenzial die State Space Grid Methode aufweist, um Lehrer-Schüler-Interaktionen auf mikrogenetischer Ebene zu erfassen und kritische Interaktionsverläufe zu erkennen.

In unserem Beitrag werden wir zudem auf die Frage eingehen, wie die State Space Grid Methode und die damit gewonnenen Erkenntnisse für die Aus- und Weiterbildung von Lehrkräften genutzt werden können.

#### Methode

In einer multimethodalen Studie wurde an 83 Klassen die Beziehung mit Fragebogen aus Lehrer- und Schülerperspektive erhoben (Wettstein, Ramseier, Scherzinger & Gasser, 2016) und der Unterricht an 18 Klassen videografiert. Die videografierten Lehrer-Schüler-Interaktionen wurden mit der State Space Grid Methode, einem grafischen Ansatz, welcher ordinale Beobachtungsdaten mit Hilfe von zwei Dimensionen quantifiziert (Hollenstein, 2013) analysiert. Dabei wurde untersucht, wie sich Lehrer-Schüler-Interaktionen auf einer mikrogenetischen Ebene gestalten und wie diese Interaktionsverläufe mit den Beziehungseinschätzung im Fragebogen (Makroebene) zusammenhängen.

#### Ergebnisse

Gelingende Lehrer-Schüler-Beziehungen bilden die Grundlage für Lehr-Lern-Prozesse. Die Lehrer-Schüler-Beziehung wird meist aus einer Makroperspektive global erfasst. Dabei bleibt oft unklar, wie sich Beziehungen in den sozialen Interaktionen auf einer Mikroebene widerspiegeln und wo Interventionen für eine Verbesserung der Lehrer-Schüler-Beziehung ansetzen können. Im Beitrag werden die State Space Grid Methode und erste Studienergebnisse präsentiert. Die Erfassung sozialer Interaktionen mittels State Space Grids eröffnet ein vertieftes Verständnis für die Lehrer-Schüler-Beziehung und gibt Hinweise zu kritischen Mustern in Interaktionsverläufen, welche für Interventionen zur Verbesserung der Lehrer-Schüler-Beziehung genutzt werden könnten.

### **Lehrer-Schüler-Interaktionen und berufliches Wohlbefinden von Lehrkräften**

**Karen Aldrup<sup>1</sup>, Uta Klusmann<sup>1</sup>, Oliver Lüdtke<sup>2</sup>, Richard Göllner<sup>3</sup>, Ulrich Trautwein<sup>3</sup>**

<sup>1</sup>Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel, <sup>2</sup>Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; Zentrum für internationale Bildungsvergleichsstudien, <sup>3</sup>Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

#### Theoretischer Hintergrund

Die sozialen Interaktionen mit Schülerinnen und Schülern sind ein wesentlicher Bestandteil des Lehrerberufs. Dabei sind gelingende soziale Interaktionen zum einen durch die Nutzung angemessener Methoden der Klassenführung gekennzeichnet, mit denen die Lehrkraft einen störungsarmen Unterricht gewährleistet (Brophy, 2006; Kunter & Voss, 2011). Zum anderen spiegeln sie sich in emotionaler Nähe, gegenseitigem Vertrauen und persönlicher Wertschätzung wider (Hamre & Pianta, 2007; Tennant et al., 2015).

Eine Vielzahl empirischer Studien konnte zeigen, dass die Qualität der Lehrer-Schüler-Interaktion eng mit dem Lernerfolg und der motivationalen Entwicklung der Schülerinnen und Schüler verbunden ist (Hamre, Hatfield, Pianta & Jamil, 2014; Kunter et al., 2013; Roorda, Koomen, Spilt & Oort, 2011). Mit Blick auf Outcomes auf Lehrerseite lag der Fokus bislang auf Aspekten der Klassenführung und weniger auf der affektiven Qualität der Beziehung. Dabei zeigte sich konsistent, dass Lehrkräfte, die ein höheres Ausmaß an Unterrichtsstörungen wahrnehmen, ein geringeres berufliches Wohlbefinden empfinden (Aloe, Shisler, Norris, Nickerson & Rinker, 2014; Dicke et al., 2014). Wenngleich die Relevanz der affektiven Beziehungsqualität bislang kaum empirisch untersucht wurde, wird ihre Bedeutung aus theoretischer Perspektive zunehmend betont. Spilt, Koomen und Thijs (2011) schlagen beispielsweise vor, dass der Zusammenhang zwischen Unterrichtsstörungen und Wohlbefinden durch die affektive Qualität der Lehrer-Schüler-Beziehung mediiert wird. Ihr Modell basiert auf der Annahme, dass Unterrichtsstörungen den Aufbau einer positiven affektiven Beziehung erschweren. Schließlich können Unaufmerksamkeit und respektloses Verhalten als Ausdruck fehlender persönlicher Wertschätzung interpretiert werden und einen Kreislauf negativer Interaktionen hervorrufen (Buyse, Verschueren, Doumen, van Damme & Maes, 2008; Newberry & Davis, 2008). Dies wiederum reduziert das berufliche Wohlbefinden der Lehrkräfte, da die Etablierung positiver affektiver Lehrer-Schüler-Beziehungen einem zentralen beruflichen Ziel und grundlegendem psychologischen Bedürfnis vieler Lehrkräfte entgegensteht (Butler, 2012; Frenzel, 2014).

#### Fragestellung

Der vorliegende Beitrag untersucht im Rahmen einer Längsschnittstudie Zusammenhänge zwischen Lehrer-Schüler-Interaktionen und dem beruflichen Wohlbefinden von Lehrkräften. Zum einen wird beleuchtet, inwieweit Unterrichtsstörungen mit dem Wohlbefinden assoziiert sind, wobei einerseits Schülerinnen und Schüler und andererseits ihre Lehrkräfte das Ausmaß von Unterrichtsstörungen einschätzten. Dadurch werden vorherige Studien erweitert, da Rückschlüsse über die Zusammenhänge zwischen Unterrichtsstörungen und Wohlbefinden über den reinen Selbstbericht der Lehrkräfte hinaus möglich werden. Zum anderen wird betrachtet, ob die affektive Qualität der Lehrer-Schüler-Beziehung mit dem Wohlbefinden assoziiert ist und den Zusammenhang zwischen Unterrichtsstörungen und Wohlbefinden mediiert.

#### Methode

Die aktuellen Analysen basieren auf Daten von N = 222 Klassenlehrerinnen und -lehrern, die zu zwei Messzeitpunkten im Abstand von einem Jahr Auskunft über Unterrichtsstörungen in ihrer Klasse, die affektive Qualität der Lehrer-Schüler-Beziehung, emotionale Erschöpfung und beruflichen Enthusiasmus gaben. Die Lehrerangaben zu den Unterrichtsstörungen wurden durch Schülereinschätzungen ergänzt (N = 4111). Die Lehrkräfte waren im Mittel 46 Jahre alt und hatten 21 Jahre Berufserfahrung. Sie unterrichteten an Haupt- (39%), Real- (22%) und Mittelschulen (40%) und 30% waren männlich.

Die Daten wurden mittels Pfadanalysen analysiert, wobei separate Modelle für die Lehrer- und Schülerperspektive und für beruflichen Enthusiasmus und emotionale Erschöpfung geschätzt wurden. Um Veränderungen im entsprechenden Outcome untersuchen zu können, wurde der Wert vom vorherigen Messzeitpunkt als Kontrollvariable genutzt. Weiterhin wurden das Geschlecht und Berufserfahrung der Lehrkraft sowie der mittlere sozioökonomische Status, Anteil von Schülerinnen und Schülern mit Migrationshintergrund und Anteil von Jungen kontrolliert.

#### Ergebnisse

Lehrkräfte, die mehr Unterrichtsstörungen berichteten, zeigten einen Anstieg ihrer emotionalen Erschöpfung und einen verringerten beruflichen Enthusiasmus. Obwohl die Lehrer- und die Schülereinschätzung von Unterrichtsstörungen weitgehend konvergierte, war die Schülereinschätzung von Unterrichtsstörungen nicht mit dem Wohlbefinden ihrer Lehrkraft assoziiert. Die affektive Qualität der Lehrer-Schüler-Beziehung mediierte den Zusammenhang zwischen der Lehrerwahrnehmung von Unterrichtsstörungen und ihrem beruflichen Enthusiasmus, wohingegen sich für emotionale Erschöpfung kein statistisch signifikanter indirekter Effekt fand. Die Befunde betonen die Bedeutung der Lehrer-Schüler-Interaktion für das berufliche Wohlbefinden von Lehrkräften. Dabei scheint allerdings vornehmlich die individuelle Lehrerwahrnehmung zentral zu sein.

## **Sozial-emotionale Kompetenz von angehenden Lehrkräften: Konzeption und Evaluation eines Trainings**

**Bastian Carstensen<sup>1</sup>, Michaela Köller<sup>2</sup>, Uta Klusmann<sup>1</sup>**

<sup>1</sup>Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel, <sup>2</sup>Institut für Pädagogisch-Psychologische Lehr- und Lernforschung, Kiel

### Theoretischer Hintergrund

Der Lehrerberuf ist durch eine hohe Dichte sozialer Interaktionen insbesondere mit den Schülerinnen und Schülern gekennzeichnet. Das Gelingen der sozialen Interaktionen ist ein wichtiger Prädiktor für das Wohlbefinden und die Motivation von Lehrkräften (Spilt, Koomen & Thijs, 2011) (Klassen, Perry & Frenzel, 2012). Auch für die Schüler/-innen gehen positiv gestaltete Lehrer-Schüler-Interaktionen mit größerem Engagement, mehr Interesse sowie größeren Lernerfolgen einher (Cornelius-White, 2007; Pianta, Hamre & Allen, 2012; Roorda et al., 2011). Entsprechend wird die Qualität der sozialen Interaktionen zwischen Lehrkräften und Schüler/innen als zentrales Unterrichtsmerkmal verstanden (z.B. Hamre et al., 2007). Die sozial-emotionale Kompetenz von Lehrkräften wurde als wichtige Voraussetzung für die Etablierung und Aufrechterhaltung positiver Interaktionen vorgeschlagen (Jennings & Greenberg, 2009). Sie umfasst Kenntnisse über Emotionen und soziale Interaktionen, die angemessene Regulation eigener und fremder Emotionen sowie den Einbezug der Bedürfnisse der Interaktionspartner in das eigene Handeln. Aus zahlreichen definitorischen Ansätzen zur sozial-emotionalen Kompetenz kristallisieren sich die wissens- und fähigkeitsbasierten Kernbereiche Wissen über Emotionen, Regulation von Emotionen und soziale Fähigkeiten heraus (Mayer & Salovey, 1997; Rubin & Rose-Krasnor, 1992; Zins et al., 2004).

Während bereits positiv evaluierte Trainingsprogramme im sozial-emotionalen Kompetenzspektrum für Lehrkräfte konzipiert wurden (u.a. Jennings et al., 2017; Steins, Bitan & Haep, 2015), mangelt es derzeit an Angeboten, die Lehrkräfte prospektiv auf die spezifischen, sozialen Herausforderungen des Berufs im Studium vorbereiten (Hohenstein et al., 2014).

### Forschungsanliegen

Der vorliegende Beitrag behandelt die Konzeption und Evaluation eines Trainingsprogramms für Lehramtsstudierende zur Förderung der sozial-emotionalen Kompetenz. Das Training umfasst zwölf inhaltliche Sitzungen und kann innerhalb eines Semesters absolviert werden. Die Inhalte beziehen sich auf die oben genannten Kernbereiche der sozial-emotionalen Kompetenz und betreffen u.a. psychologische Emotionstheorien, Bedeutung von Emotionen im Schulkontext, Strategien zur Emotionsregulation, Perspektivenübernahme, Wahrnehmungsverzerrungen, das angemessene Durchsetzen eigener Ziele oder den Umgang mit Konflikten. Neben der Vermittlung theoretischen Wissens umfasst das didaktische Konzept praktische Elemente, welche die Elaboration und den Transfer des Gelernten fördern sollen.

Die Evaluation umfasste zum einen die Reaktion der Teilnehmenden auf das Trainingsangebot. Zum anderen wurde die Effektivität des Trainingsprogramms hinsichtlich der Förderung des Wissens über Emotionen, der Regulation von Emotionen sowie den sozialen Fähigkeiten überprüft.

### Methode

Die Evaluation basiert auf einem quasi-experimentellen Studiendesign mit Kontrollbedingung und drei Messzeitpunkten. Die Reaktion der Teilnehmenden wurde über Angaben zur Zufriedenheit und zur wahrgenommenen Relevanz der Inhalte fortlaufend nach jeder Trainingssitzung erhoben.

Insgesamt N = 207 Studierende schätzten sich in den Kernbereichen der sozial-emotionalen Kompetenz ein. In der Experimentalbedingung (N = 95) wurde der Selbstbericht zu sozialen Fähigkeiten durch einen Fremdbbericht von regelmäßigen Kontaktpersonen ergänzt. Zudem schätzten die Trainingsteilnehmer/-innen ihr Wohlbefinden im Rahmen eines Tagebuches zu Beginn und zum Ende des Trainingsprogramms an insgesamt 21 Tagen ein.

Die Angaben zur Zufriedenheit und wahrgenommenen Relevanz wurden deskriptiv ausgewertet. Fehlende Werte zu t1 und t2 wurden mittels multipler Imputation geschätzt. Die imputierten Datensätze wurden durch t-Tests für abhängige Stichproben (Vergleich t1, t2) getrennt für Experimental- und Kontrollbedingung analysiert. Zur Erfassung längerfristiger Effekte des Trainings wurden die Originaldaten von t2 und t3 auf Unterschiede getestet. Ebenso wurden die aggregierten Tagebuchdaten analysiert.

### Ergebnisse

Die Reaktion der Teilnehmenden auf das Trainingsprogramm fällt mit hohen Zufriedenheitswerten und einer hohen wahrgenommenen Relevanz der Inhalte positiv aus.

Während sich für die Kontrollbedingung zu keinem Messzeitpunkt signifikante Veränderungen ergaben, schätzten die Trainingsteilnehmer/-innen sich nach dem Training besser hinsichtlich der Regulation von Emotionen ein. Zudem zeigten sich Verbesserungen in den selbsteingeschätzten sozialen Fähigkeiten; diese wurden zum Teil durch den Fremdbbericht bestätigt. Die Analyse der Tagebuchdaten zeigte signifikant höhere Werte für das Wohlbefinden zum zweiten Messzeitpunkt. Hinsichtlich des Wissens über Emotionen ergaben sich keine signifikanten Veränderungen. Der dritte Messzeitpunkt weist auf eine Stabilität der Trainingseffekte hin.

## **Evaluation eines Classroom-Management-Trainings für Lehramtsstudierende**

**Lena Hannemann, Gesa Uhde, Barbara Thies, Hannah Perst**  
TU Braunschweig

### Theoretischer Hintergrund

Lehrkräfte fühlen sich zu Beginn ihrer praktischen Tätigkeit oft stark verunsichert und unzureichend vorbereitet (Lubitz, 2007; Melnick & Meister, 2008; Havers, 2010). Sie haben Probleme bei der Anwendung des theoretischen Wissens im praktischen Schulkontext (Klusmann et al., 2012). Als Ursache wird oftmals ein unzureichender Praxisbezug in der universitären Ausbildung angeführt (Combe & Kolbe, 2004; O'Neill & Stephenson, 2012). Als eine der wichtigsten Ressourcen im Umgang mit Belastungen im Lehrberuf gelten Classroom-Management-Kompetenzen (CM; Kiel et al., 2013; Dicke et al., 2015). CM-Kompetenzen gehen auf Schüler/innen-Seiten mit einem Zuwachs an Leistung und Lerninteresse einher und tragen wesentlich zum Aufbau einer positiven Schüler/innen-

Lehrkraft-Beziehung bei (Marzano, 2000; Helmke & Jäger, 2002; Helmke et al., 2008; Schönabächler, 2008).

CM-Kompetenzen setzen sich als Facette des pädagogisch-psychologischen Wissens zusammen aus deklarativem und prozeduralem Wissen (Baumert & Kunter, 2011; Ophardt & Thiel, 2013). Kompetenztrainings setzen beim Aufbau prozeduralen Wissens an (Uhde & Jürgens, 2013). Im deutschen Sprachraum wurden bisher verschiedene Kompetenztrainings für Lehrkräfte konzipiert und evaluiert (Dann & Humpert, 2002; Havers, 2007; Thiel, Ophardt & Piwowar, 2013; Uhde, 2015). Es

fehlen jedoch Kompetenztrainings im Classroom-Management-Bereich, die zu Beginn der Lehramtsausbildung ansetzen und den Theorie-Praxis-Transfer bei Lehramtsstudierenden frühzeitig anbahnen.

Forschungsanliegen

Ziel der Studie ist die Überprüfung der Effektivität eines Trainings zur Förderung von CM-Kompetenzen sowie eines Literaturkurses zum Aufbau von deklarativem Wissen im Classroom-Management-Bereich. Selbsteingeschätztes Wissen und selbsteingeschätzte Kompetenzen im Classroom-Management-Bereich werden erhoben. Bei der Intervention >CM-Training< steht die Anbahnung prozeduralen Handlungswissens im Fokus. Im Rahmen der Intervention >CMLiteraturkurs< wird der Aufbau der theoretischen Wissensstruktur zum Thema Classroom-Management unterstützt. Die Fragestellungen lauten:

Schreiben sich Lehramtsstudierende der Trainingsgruppe höhere CM-Kompetenzen zu als Studierende, die einen Literaturkurs oder keine Intervention erhalten haben? Schreiben sich Studierende der Literaturkursgruppe höheres CM-Wissen zu als Studierende, die am CM-Training oder an keiner Intervention teilgenommen haben? Bleiben diese Veränderungen über die Zeit stabil?

Methode

Die Teilnehmenden (Lehramtsstudierende des dritten Semesters) wurden zufällig einer CM-Trainingsgruppe, einer CM-Literaturkurs-Gruppe oder einer Kontrollgruppe zugewiesen. Für die Daten der Prä- und Postmessung lagen vor: CM-Trainingsgruppe (nTG=96), CM-Literaturkurs-Gruppe (nLG=76), Kontrollgruppe (nKG=55). Zur Erfassung der abhängigen Variablen wurden Skalen zum selbsteingeschätzten prozeduralen und deklarativen Wissen von Thiel, Ophardt und Piwowar (2013) eingesetzt. Zur Überprüfung der Hypothesen wurden im ersten Schritt Varianzanalysen mit Messwiederholung durchgeführt mit den Faktoren Zeit (Prätest/Posttest) und Gruppenzugehörigkeit. Im zweiten Schritt wurden Kontrastanalysen der Differenzvariablen Prä/Post für den paarweisen Vergleich zwischen den Gruppen durchgeführt.

Ergebnisse

Die Varianzanalyse ergab für die Skalen zu den selbsteingeschätzten Kompetenzen signifikante Interaktionseffekte zwischen den drei Gruppen über die Zeit in die erwartete Richtung insbesondere für die Skalen Regeln und Störungsintervention. Die Interaktion für die Skala Zeitmanagement war nicht signifikant. Die Kontrastanalysen der Differenzvariablen ergaben signifikante Differenzen für alle Skalen zwischen der Trainings- und der Kontrollgruppe mit der Ausnahme der Skala Zeitmanagement. Signifikante Kontraste ergaben sich zwischen Trainings- und Literaturkurs-Gruppe für die Skalen Regeln, Klarheit des Handlungsprogramms und Störungsintervention. Zwischen Literaturkurs- und Kontrollgruppe ergaben sich signifikante Kontraste für die Skalen Störungsintervention und Monitoring sowie für die Skala Umgang mit Konflikten unter Schülerinnen und Schülern. Die Varianzanalyse zu den Skalen im selbsteingeschätzten Wissen zeigte signifikante Interaktionseffekte zwischen den drei Gruppen über die Zeit insbesondere für die Skalen Störungsintervention, Regeln sowie Prozeduren/Abläufe. Die Kontrastanalysen der Differenzvariablen ergaben signifikante Kontraste für alle Skalen zwischen Trainings- und Kontrollgruppe. Zwischen Literaturkurs- und Kontrollgruppe ergaben sich signifikante Kontraste für alle Skalen. Zwischen Trainings- und Literaturkurs-Gruppe waren die Kontraste zwischen den Skalen Störungsintervention und Regeln signifikant. Die Unterschiede zwischen Trainings-, Literatur- und Kontrollgruppe blieben bis zum dritten Messzeitpunkt stabil.

Zusammenfassend lässt sich schließen, dass durch das Training CM-Kompetenzen angebahnt werden, die Voraussetzung für gelingende Lehrer-Schüler-Interaktionen sind.

## A17

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H114

### Wirkungen von Unterrichtsqualität: Bisherige Forschungsbefunde und methodische Herausforderungen

Chair(s): **Carmen Koehler** (DIPF, Deutschland), **Anna-Katharina Praetorius** (Universität Zürich)

DiskutantIn(nen): **Marcus Pietsch** (Leuphana Universität Lüneburg)

Ein zentrales Interesse der Unterrichtsforschung besteht darin, Wirkungen von Unterrichtsqualität auf die Lern- und Leistungsentwicklung sowie die motivationale Entwicklung von Schülerinnen und Schülern zu überprüfen, um diejenigen Merkmale zu identifizieren, die besonders lern- und motivationsförderlich sind (z. B. Seidel & Shavelson, 2007). Eine solche Identifikation ist methodisch hoch anspruchsvoll, da verschiedene Komponenten hinreichend gegeben sein müssen. Unter anderem muss eine reliable und valide Messung (a) relevanter Unterrichtsmerkmale (z. B. Praetorius, Lenske & Helmke, 2012), (b) wesentlicher Outcomes auf Seiten der Schülerinnen und Schüler auf Individual- und Klassenebene (Naumann, Hartig & Hochweber, 2017) sowie (c) eine methodisch angemessene Verbindung und Berücksichtigung einer gegebenenfalls wechselseitige Beeinflussung zwischen beiden gewährleistet sein (z. B. Marsh, Lüdtke, Nagengast, Trautwein, Morin, Abduljabbar, & Köller, 2012). Diese Herausforderungen stehen im Fokus des Symposiums. Dabei sollen einerseits der aktuelle Stand der Forschung und die zurzeit gängigen methodische Ansätze vorgestellt werden, andererseits gilt es auch diese kritisch zu reflektieren und einen Ausblick auf notwendige Weiterentwicklungen zu geben.

Der erste Beitrag gibt einen Überblick über die bisherige inkonsistente Befundlage zu den Wirkungen der drei Basisdimensionen von Unterrichtsqualität (Klassenführung, konstruktive Unterstützung und kognitive Aktivierung, z. B. Klieme, Lipowsky, & Rakoczy, 2006) auf Schülerleistungen und motivationale Aspekte und diskutiert mögliche Ursachen für die Heterogenität der Befunde.

Im zweiten Beitrag liegt der Fokus auf der adäquaten Modellierung von Unterrichtsmerkmalen für internationale Studien. Die Vergleichbarkeit von Unterrichtsskalen über Länder hinweg ist für die Verwendung solcher Merkmale als Prädiktoren für Schüleroutcomes essentiell, da nur eine – für alle Länder – vergleichbare Messung der Prädiktoren ländervergleichende Analysen zur Wirkung von Unterricht erlauben. Die Befunde deuten auf metrische, nicht aber skalare Messinvarianz der Skalen und interessante Unterschiede innerhalb von Sprachgruppen hin.

Der dritte Beitrag stellt verschiedene Modelle vor, die für die Untersuchung der Wirkung von Unterricht genutzt werden. Hier werden vor allem methodische Aspekte hervorgehoben, die bei der Analyse des Einflusses von Unterrichtsmerkmalen auf Lernergebnisse eine Rolle spielen und bedacht werden müssen. Verschiedene aktuelle Mehrebenen-Modelle, welche die typische genestete Datenstruktur in der Unterrichtsforschung berücksichtigen, werden präsentiert und miteinander verglichen. Sie werden dahingehend diskutiert, ob sie der eigentliche Forschungsintention – nämlich die Frage nach der Wirkung von Unterricht zu beantworten – überhaupt nachkommen.

Marcus Pietsch diskutiert abschließend die Beiträge im Hinblick auf richtungsweisende Ansätze in der Forschung von Wirksamkeit von Merkmalen der Unterrichtsqualität.

*Beiträge des Symposiums*

### Einheitliche Befunde? Ein Review zu den Wirkungen der drei Basisdimensionen von Unterrichtsqualität

**Eckhard Klieme<sup>1</sup>, Anna-Katharina Praetorius<sup>2</sup>, Benjamin Herbert<sup>1</sup>**

<sup>1</sup>DIPF, Deutschland, <sup>2</sup>Universität Zürich

Theoretischer Hintergrund und Fragestellung

Das Modell der drei Basisdimensionen von Unterrichtsqualität – Klassenführung, konstruktive Unterstützung und kognitive Aktivierung – stellt im deutschsprachigen Raum den mittlerweile am weitesten verbreiteten Ansatz zur Konzeptualisierung von Unterrichtsqualität in der quantitativen empirischen Unterrichtsforschung dar. Es wurde im Rahmen der deutschen Erweiterung der TIMSS-Videostudie 1995 entwickelt und basiert auf allgemeinen Unterrichtstheorien sowie etablierten Forschungstraditionen der Unterrichtspsychologie (u.a. Selbstbestimmungstheorie, siehe Ryan & Deci, 2000; Konstruktivismus, siehe Staub & Stern, 2002). Den theoretischen Überlegungen zufolge werden Klassenführung und kognitive Aktivierung als prädiktiv für die Leistungsentwicklung von Schüler/-innen sowie Klassenführung und konstruktive Unterstützung als prädiktiv für die motivationale Entwicklung der Schüler/-innen angenommen. Erste unterstützende Befunde für diese Annahmen finden sich bei Klieme, Schümer und Knoll (2001). In den letzten Jahren wurde eine Reihe weiterer Studien zur prädiktiven Validität der drei Basisdimensionen in verschiedenen Fächern und mit unterschiedlichen Operationalisierungen und Messansätzen durchgeführt. Eine Übersicht zur Einheitlichkeit dieser Befunde existiert bislang nicht; dies war entsprechend Ziel des vorliegenden Reviews.

Methode

Basierend auf einer umfassenden Literaturrecherche mittels Education Resources Information Center (ERIC) sowie Fachinformationssystem Bildung (FIS Bildung) sowie ergänzenden Suchen per Hand sowie Schneeballsystem wurden basierend auf den folgenden Kriterien zunächst 39 Publikationen mit Bezug zu 19 empirischen Studien identifiziert. Es wurden diejenigen Studien ausgewählt, (a) bei denen es sich um quantitativ-empirische Studien handelt, (b) die sich explizit auf das Modell der drei Basisdimensionen beziehen, (c) die ihre jeweilige Operationalisierung der Basisdimensionen beschreiben oder auf eine solche verweisen und (d) die zur Datenerhebung hoch-inferente Skalen nutzen. Von diesen 39 Publikationen berichten 13 über die prädiktive Validität der drei Basisdimensionen. Um den mehrfachen Einbezug derselben Daten zu vermeiden, wurde für jedes in einer Studie berichtete Schüleroutcome nur eine Publikation einbezogen. Zudem wurden aufgrund der andernfalls eingeschränkten Interpretierbarkeit lediglich Befunde in das Review aufgenommen, die auf einem längsschnittlichen Mehrebenen-Design basieren. Die Anzahl einbezogener Beiträge bezieht sich entsprechend auf 7 Publikationen zu 6 Studien mit insgesamt 36 berichteten Effekten zu Leistungsauscomes sowie motivationalen Outcomes der Schüler/-innen.

Ergebnisse und Diskussion

Eine Zusammenschau der Befunde zeigt, dass Effekte der drei Basisdimensionen auf die Leistungs- und Motivationsentwicklung der Schüler/-innen identifiziert werden können. Allerdings lassen sich die theoretischen Annahmen dabei nur teilweise empirisch bestätigen. So konnten von den erwarteten 10 Effekten der Klassenführung auf die Leistungsentwicklung der Schüler/-innen lediglich 6 identifiziert werden, für die kognitive Aktivierung sogar lediglich einer von 6 Effekten. Von den sechs erwarteten Effekten der konstruktiven Unterstützung auf die motivationale Entwicklung der Schüler/-

innen konnten lediglich zwei statistisch abgesichert werden; für die Klassenführung war dies für drei der sechs Effekte der Fall, dabei fielen zwei erwartungswidrig jedoch negativ aus. Zusätzlich zu den theoretisch angenommenen Effekten fanden sich weitere statistisch signifikante Effekte, für die keine entsprechenden Annahmen getroffen wurden (1 von 9 Effekten für den Effekt konstruktiver Unterstützung auf die Leistungsentwicklung; 1 von 2 Effekten für den Effekt kognitiver Aktivierung auf die Motivationsentwicklung). Der Beitrag schließt mit einer Diskussion möglicher Ursachen für die inkonsistente Befundlage (u.a. Zentrierungen, Modellierung der Variablen zum ersten Messzeitpunkt, Operationalisierungen und Suppressionseffekte) sowie mit einem Ausblick auf zukünftige Forschung.

## **Internationale Vergleichbarkeit von Schülerratings zur Erfassung von Unterrichtsqualität**

**Jessica Fischer<sup>1</sup>, Anna-Katharina Praetorius<sup>2</sup>, Eckhard Klieme<sup>1</sup>**

<sup>1</sup>DIPF, Deutschland, <sup>2</sup>Universität Zürich

### Theoretischer Hintergrund

Fragebogendaten werden oftmals verwendet um zwischen Ländern variierende Schülerleistungen zu erklären (Schulz, 2003). Ein Faktor, der Schülerleistungen positiv beeinflusst und zwischen Ländern variiert ist die Unterrichtsqualität (Nielsen & Gustafsson, 2016), die oft über die drei Basisdimensionen Klassenführung, konstruktive Unterstützung und kognitive Aktivierung konzeptualisiert wird (Klieme, Lipowsky & Rakoczy, 2006). Ist die Datenvergleichbarkeit gewährleistet, kann der Einfluss dieser Faktoren auf Schülerleistungen zwischen Ländern analysiert werden (Klieme & Baumert, 2001). Nach van de Vijver und Leung (1997) sind Fragebogendaten vergleichbar, wenn dasselbe Konstrukt (konfigurale Äquivalenz) mit derselben Messeinheit (metrische Äquivalenz) und demselben Ursprung (skalare Äquivalenz) in allen Gruppen gemessen wird. Eine Äquivalenz interkultureller Daten ist jedoch nicht unbedingt gegeben, unter anderem durch Unterschiede im Item-Verständnis und -Beantwortung zwischen Ländern (Miller, Mont, Maitland, Altman & Madans, 2011). Ob Daten interkulturell vergleichbar sind, wird trotz der hohen Relevanz oftmals nicht überprüft (Schulz, 2005).

Unterrichts-Items sind begrenzt zwischen Ländern vergleichbar. In der Regel werden dieselben Konstrukte gemessen, skalare Äquivalenz wird jedoch oftmals nicht erfüllt (Çetin, 2010; He, Buchholz, & Klieme, 2017; Schulz, 2005; Schulz, 2003). Die Modellgüte variiert unabhängig von den Äquivalenzstufen zwischen Ländern, vor allem für asiatische Länder ist diese nicht zufriedenstellend. Restriktionen wirken sich in dieser Gruppe weiterhin besonders negativ aus (Schulz, 2005). Die Wahrscheinlichkeit der Erreichung der Äquivalenzstufen ist insgesamt höher, je ähnlicher die Kultur und die Sprache der Länder sind (Kankarash & Moors, 2014). Dementsprechend sollte eine Unterteilung nach Sprachgruppen bei der Invarianzprüfung berücksichtigt werden.

### Fragestellung

Ziel dieses Beitrages ist es den lückenhaften Forschungsstand der Äquivalenz von Unterrichts-skalen ein Stück weit zu schließen, in dem die Messinvarianz innerhalb und zwischen jeweils drei chinesisch-, deutsch- und spanischsprachigen Ländern untersucht wird. Dahingehend wird angenommen, dass Konstruktäquivalenz innerhalb und zwischen Sprachgruppen erreicht wird (H1). Innerhalb der Sprachgruppen wird für die chinesische das geringste Ausmaß an Invarianz vorhergesagt (H2), da diese kulturell und sprachlich die heterogenste Sprachgruppe darstellt (Schulz, 2005). Weiterhin wird erwartet, dass innerhalb der Sprachgruppen die Skalen ein größeres Ausmaß an Invarianz aufweisen als zwischen Sprachgruppen (H3).

### Methode

Die Datengrundlage bilden Schüler-Fragebogendaten der PISA-Erhebung 2012 für jeweils drei chinesischsprachige (Macao, Shanghai, Taipei), deutschsprachige sowie spanischsprachige Länder (Deutschland, Österreich, Schweiz sowie Chile, Kolumbien, Mexiko). Klassenführung und konstruktive Unterstützung werden anhand von fünf Items gemessen, kognitive Aktivierung anhand von neun Items. Die Skalenreliabilität aller Konstrukte ist in allen Ländern sehr gut (Range  $\alpha$ : Klassenführung 0.80-0.91; konstruktive Unterstützung 0.80-0.87; kognitive Aktivierung 0.79-0.85).

Zur Beantwortung der Hypothesen werden konfirmatorische Mehrgruppen-Faktorenanalysen durchgeführt. Die Modellgüte wird mit dem Comparative fit index (CFI) evaluiert. Das striktere Modell wird abgelehnt, wenn die Veränderung des CFI vom konfiguralen zum metrischen mehr als 0.02 und vom metrischen zum skalaren Modell mehr als 0.01 beträgt (He et al., 2017).

### Ergebnisse

Die Ergebnisse sprechen für jede Unterrichtsdimension für eine Ein-Faktorstruktur in allen Ländern als auch für eine Invarianz der Faktorladungen innerhalb sowie zwischen den Sprachgruppen (Veränderung CFI maximal 0.016). Konfigurale und metrische Äquivalenz sind demnach gegeben (spricht für H1). Skalare Äquivalenz konnte lediglich für Klassenführung in der deutschen (Veränderung CFI 0.007) sowie für kognitive Aktivierung in der spanischen Sprachgruppe (Veränderung CFI 0.009) nachgewiesen werden. Keine der Dimensionen erfüllt die Bedingung der skalaren Äquivalenz in der chinesischsprachigen Gruppe sowie über alle Sprachgruppen hinweg. Auffällig ist die besonders negative Auswirkung der Modellparameter-Restriktionen in der chinesischsprachigen Gruppe beim Modellvergleich eines metrischen versus skalaren Modells (Veränderung CFI zwischen 0.035 und 0.125 im Vergleich zu 0.011 und 0.025 in den anderen Sprachgruppen) (spricht für H2). Besonders hoch fällt diese für die kognitive Aktivierung aus (Veränderung CFI 0.125). Die Ergebnisse sprechen folglich dafür, dass Wirkungen von Unterrichtsqualität in und zwischen Sprachgruppen auf Basis von Effekten, nicht aber Mittelwerten analysiert werden können. Des Weiteren sollten chinesische Regionen getrennt analysiert werden.

## **Multilevel Modelle zur Evaluierung von Wirkung im Unterricht: Konzeptuelle und Methodische Herausforderungen**

**Carmen Köhler, Susanne Kuger, Alexander Naumann, Johannes Hartig**

DIPF, Deutschland

### Theoretischer Hintergrund

Ein Ziel der Unterrichtsforschung ist die Identifikation von Einflüssen des unterrichtlichen Handelns Lehrerverhaltens auf Schülerinnen und Schüler, insbesondere auf deren Lernerfolg oder deren Lernmotivation. Entsprechende Fragestellungen und die damit verbundene genestete Datenstruktur mit Schülern in Klassen, Schulen, etc. erfordern eine vergleichsweise komplexe Mehrebenenmodellierung, um Einflüsse auf der Gruppenebene zu identifizieren. Damit gehen einige methodische Herausforderungen in Bezug auf die Erstellung des Erhebungs- und Datendesigns, die Wahl des statistischen Modells und die adäquate Interpretation der Ergebnisse einher (Marsh, Lüdtke, Nagengast, Trautwein, Morin, Abduljabbar, & Köller, 2012; Miller & Murdock, 2007).

### Fragestellung

Der folgende Beitrag befasst sich mit der Frage, wie Wirkung von Unterricht empirisch erfasst werden kann. Es werden relevante Aspekte des Designs, z.B. auf welcher Ebene das Lehrerverhalten gemessen wird, aufgegriffen und in Bezug auf verschiedene inhaltliche Fragestellungen der Unterrichtsforschung erörtert. Außerdem werden verschiedene Mehrebenenmodelle, welche die Identifikation von Unterrichtseinflüssen erlauben sollen, vorgestellt und beispielhaft angewendet. Hierbei wird auf die verschiedenen Möglichkeiten der Zentrierung von Variablen eingegangen (Enders & Tofghi, 2007). Darüber hinaus werden Modelle mit latenten Veränderungsfaktoren mit Modellen, bei denen der Ausgangswert als Kontrolle mit einbezogen wird, verglichen. Es wird eruiert, welche Modelle welche inhaltlichen Interpretationen erlauben sowie welche Modellierung welchen Begrenzungen unterliegt.

#### Methode

Als Datengrundlage diente die 2. und 3. Welle der BiKS-8-14 (Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter) Studie, welche Untersuchungen zur sprachlichen und kognitiven Entwicklung bei Kindern und Jugendlichen durchführte. Die 2. und 3. Welle bezieht sich auf die Erhebungen im 1. und 2. Halbjahr der 4. Klassenstufe. Die Anzahl der Schüler betrug  $NS = 1747$ , die Anzahl der Lehrer – und somit der Cluster – betrug  $NT = 124$ . Im Beitrag wurde exemplarisch der Zusammenhang zwischen der Schüler-Lehrer-Beziehung und der Mathematikleistung der Schüler analysiert. Verschiedene latente und doppelt-latente Mehrebenenmodelle (s. Marsh, Lüdtke, Robitzsch, Trautwein, Asparouhov, Muthén, & Nagengast, 2009) mit und ohne Längsschnittberücksichtigung wurden mit dem Programm Mplus (Muthén & Muthén, 1998-2015) berechnet. Zur Berücksichtigung des Längsschnitts wurden Latent Growth Modelle (LGM) und kovarianzanalytische Modelle, bei denen der Messwert zum 1. Zeitpunkt als Kontrollvariable diente, analysiert und miteinander verglichen.

#### Ergebnisse

Ergebnisse aus den Analysen zeigten, dass sich je nach Modellierung die Aussagen über den Zusammenhang zwischen der Lehrer-Schüler-Beziehung und der Mathematikleistung ändert. Inwieweit welche Modellierungen Aussagen über die Wirkung von Unterricht zulassen und welche Vor- und Nachteile die einzelnen Modelle mit sich bringen wird evaluiert. Die Ergebnisse lenken damit insgesamt die Aufmerksamkeit auf Details der Modellierung, denen nicht immer ausreichend Beachtung geschenkt wird. Die Diskussion richtet sich u.a. auf die Relevanz der Fragestellung bei der Entscheidung, wie mit diesen Details umgegangen werden sollte.



## A18

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H102

### **Measurement and effects of the cultural diversity context at school and beyond: Implications for student and teacher development, well-being and performance**

Chair(s): **Maja Schachner** (Universität Potsdam, Deutschland), **Linda Juang** (Universität Potsdam, Deutschland)

DiskutantIn(nen): **Petra Stanat** (Humboldt-Universität zu Berlin)

Cultural diversity and how this diversity is handled in schools have important implications for student and teacher development, well-being and performance (for recent reviews, see Rjosk, Richter, Lüdtke, & Eccles, 2016; Schachner, 2017). To better understand these implications, there is an urgent need for valid and reliable measures of different levels of the diversity context in schools. Starting out with a conceptual reflection of the very nature of group representation and diversity (study 1), this symposium brings together a collection of studies employing different types of measures of how this diversity is approached, from student and teacher self-reports of beliefs, climate and practices (studies 2, 3 and 4), to video observations of teacher behaviour (study 4). The studies also tap into different facets of the diversity context, such as structural and perceived group representations and diversity (study 1), individual diversity beliefs (study 4), different facets of the intergroup and multicultural climate (study 2), differentiation and individualisation in teaching more generally (study 3), and finally culturally responsive teaching as one specific facet of differentiation (study 4). Together, the studies highlight the multidimensional nature of cultural diversity contexts, the need for more precision in how we define and study such contexts, and how an understanding of the diversity context matters for both teachers and students.

*Beiträge des Symposiums*

### **Understanding Ethnic Density in School and Community Contexts**

Moin Syed<sup>1</sup>, Linda Juang<sup>2</sup>, Ylva Svensson<sup>3</sup>

<sup>1</sup>University of Minnesota, USA, <sup>2</sup>Universität Potsdam, Deutschland, <sup>3</sup>University of Gothenburg, Schweden

In increasingly culturally diverse societies there is a growing recognition of the importance of the surrounding context to understanding children's school experiences. One key aspect of the context is ethnic density, or the proportion and nature of co-ethnics within an individual's environment. Density is a contextual factor with great relevance to many aspects of child development, such as ethnic identity, discrimination, schooling, inter-ethnic peer relations, and neighborhood influences (see Graham, 2016; van der Meer & Tolsma, 2014). For instance, increased ethnic heterogeneity at both classroom and school levels was related to perceptions of greater safety and social satisfaction for ethnic minority youth (Juvonen, Nishina, & Graham, 2006), and a greater proportion of students of immigrant background in the classroom was positively associated with outgroup orientation and contact in some studies (Schachner et al., 2015; Schwarzenthal et al., 2017), but not others (Baysu, Phalet, & Brown, 2014; Brenick, et al., 2012). Because ethnic density has been measured and studied in multiple and inconsistent ways, it has been difficult to come to clear conclusions about the role of ethnic density across the various studies. Thus, in this paper we present a conceptual framework for the study of ethnic density that consists of four dimensions: 1) perspective (that density can be understood in both objective and subjective terms); 2) heterogeneity (the range of ethnic groups in a setting); 3) differentiation (how groups are defined in a setting); and 4) proximity (the distance between the individual and the setting). Our focus is on how to conceptualize ethnic density itself.

Our review shows that the dimension of perspective that most studies adopt is using objective measures of density. Most studies also have focused much more on the dimension of differentiation than on heterogeneity. Heterogeneity has been measured more consistently, with many studies using Simpson's Index. Differentiation, in contrast, has been calculated based on group boundaries defined in a variety of ways: by specific ethnic group, migration status, racial group, minority versus majority group, and People of Color versus not. Finally, most studies consider only one level of proximity, ranging from the people around you, classroom, school, university, neighborhood, city, city type (urban versus suburban), city district, county, region (e.g., coast versus highland), state, and country. When considering the different ways in which each of the four dimensions can be defined and combined, these variations in focus and measurement of each dimension render it difficult to develop a cumulative understanding of how, when, and why density matters, because the definitions are highly inconsistent across studies. Our definition of density—which incorporates all four dimensions—is meant to highlight how the dimensions are always inter-related; that conceptualizing and measuring one dimension also requires a consideration of the others, and when measures are selected, the other dimensions are always at least implicitly defined.

### **Classroom Cultural Diversity Climate – Conceptualisation, Measurement and Associations amongst Secondary School Students**

**Maja Schachner, Miriam Schwarzenthal, Ursula Moffitt, Sauro Civitillo, Linda Juang**  
Universität Potsdam, Deutschland

One of the challenges of the 21st century is to create settings in schools where all students can do well, regardless of ethnic background. Such settings should promote understanding and acceptance of people of diverse cultural backgrounds and prepare students to become citizens of a culturally diverse society. In the current study, we developed a measure for different facets of the cultural diversity climate in schools, the Classroom Cultural Diversity Climate Scale (CCDCS). The first version of the scale (CCDCS-1) was developed for early adolescent students. It included eight subscales, namely unequal treatment by teachers, support for contact and cooperation by teachers, unequal treatment by students, support for contact by students, support for cooperation by students, interest in ethnic background of students, multicultural curriculum – relationships, and multicultural curriculum – topics (38 items in total). Factorial invariance was demonstrated across students of immigrant and non-immigrant background, using data from 759 students (Mage = 11.2 years; 51% male; 63% immigrant background). Test-retest reliability across the first year at secondary school was also established. A second version of the scale (CCDCS-2) was developed for mid-adolescents, including the six subscales contact and cooperation, equal treatment, critical consciousness, heritage and intercultural learning, colorblindness, and polyculturalism (31 items in total). Factorial invariance across students of immigrant and non-immigrant background was demonstrated using data from 1335 students (Mage = 14.7; 51% male; 51% immigrant background). Subscales showed meaningful correlations with students' individual diversity beliefs as well as psychological and school adjustment outcomes (life satisfaction, subjective school values, grade average and behavioral disengagement). Although some differences in associations were observed between subscales and students of immigrant and

non-immigrant background, the overall pattern suggests that a positive diversity climate, characterized by positive interethnic relations and opportunities to learn about cultural diversity, is associated with better outcomes amongst all students.

## **Teachers' Willingness to Differentiate Impacts Student-perceived Instructional Quality**

**Axinja Hachfeld<sup>1</sup>, Rebecca Lazarides<sup>2</sup>**

<sup>1</sup>Universität Konstanz, Deutschland, <sup>2</sup>UniPotsdam, Deutschland

### Objectives

In the context of growing cultural diversity in the classroom, teachers' willingness and their ability to adapt and differentiate their teaching practices have been discussed as important aspects of professional competence. Studies show that, for example, teachers' cultural beliefs and their notion of costs and benefits of diversity affect their motivational orientations and their self-reported intention for adaptive teaching (Hachfeld, Hahn, Schroeder, Anders, & Kunter, 2015; Gebauer & McElvany, 2017). From a social equity perspective, adaptive teaching should be especially important for minority students. Yet, it is unclear how teachers' willingness to differentiate is perceived by students and if it results in higher student-perceived instructional quality. This study investigates the relationship between teachers' willingness to address students' individual needs in their instructional practices in a longitudinal design taking into account teacher and student ratings.

### Methods and Data

Multilevel regression analyses were applied using data of three waves of the German longitudinal Motivation in Learning in Mathematics study (MOVE). For the present analyses, data was used from 1,115 ninth to tenth graders (45.3% boys; age:  $M=14.56$  years,  $SD=0.88$ ) from 58 classrooms in Berlin and their mathematics teachers. Teachers' willingness to address students' individual needs in their instructional practices was measured with two scales: differentiation and individualization in mathematics classrooms ("In group work I give different tasks to more and less advanced students."). Teachers beliefs about classroom heterogeneity was assessed with one scale ("The heterogeneity in my class inspires me to try new methods."). Instructional quality was measured with a scale assessing students' perceived monitoring (e.g., "Our teachers know what is going on in the classroom.") and cognitive activation ("Our teachers ask us to show different solutions to a problem.").

### Results and Significance

Teacher-reported differentiation ( $\beta=.27$ ,  $SE=.12$ ,  $p=.03$ ) at the beginning of the school year was significantly and positively related to student-perceived monitoring at the classroom level at mid-year when controlling for student-perceived monitoring at the beginning of the school year. Teacher-reported beliefs regarding the heterogeneity in class was marginally significantly and negatively associated with student-perceived class-level monitoring at mid-year ( $\beta=-.18$ ,  $SE=.10$ ,  $p=.056$ ) when controlling for classroom composition. Classroom composition regarding students' migration background was negatively associated with student-perceived monitoring at the beginning of the school year ( $\beta=-.30$ ,  $SE=.10$ ,  $p=.003$ ). No significant effects were found for student-perceived cognitive activation (differentiation:  $\beta=-.05$ ,  $SE=.11$ ,  $p=.63$ ; individualization:  $\beta=.09$ ,  $SE=.12$ ,  $p=.46$ ).

The results stimulate the discussion about differentiation in the classroom as a means of providing equal opportunities for students with different (cultural) backgrounds. In multicultural classrooms, teacher-reported differentiation, but not their beliefs about heterogeneity, improved perceived student-reported monitoring. This relationship, however, was not specific for minority students. Questions remain how teachers apply adaptive teaching strategies in multicultural settings and how they are linked to educational outcomes of students, in particular minority students.

## **Investigating cultural diversity beliefs, culturally responsive teaching, and self-reflections skills with in-service teachers: A case study**

**Sauro Civitillo, Linda Juang, Maja Schachner, Marcel Badra**

Universität Potsdam, Deutschland

Theoretical framework. Different teacher beliefs about cultural diversity (e.g., multiculturalism, colorblindness) are assumed to be linked with different teaching behaviors (e.g., acknowledging cultural differences versus denying the cultural background of students). Yet few studies have empirically tested these associations. In addition, self-reflection is essential to align beliefs and practices for novice as well as experienced teachers and provides an opportunity to discuss the tensions between beliefs and practices.

The inclusive framework of Culturally Responsive Teaching (CRT, Gay, 2010) provides insights into addressing the needs of culturally diverse student populations. The CRT is a multidimensional construct, and encompasses curriculum content, learning context, classroom climate, instructional techniques, and performance assessment. CRT is not only a set of teaching strategies but it requires teachers holding beliefs that challenge a deficit view of diverse students and their families and communities. Thus, the aims of this case study was to assess how the relation between teacher beliefs about cultural diversity relate to culturally responsive teaching and examine how teachers critically reflect on their own teaching behaviors.

Methods. This qualitative case study employed classroom observations and post-observation interviews. A modified version of the Culturally Responsive Instruction Observation Protocol (CRIOP; Powell, et al., 2014) was used for measuring teachers' use of culturally responsive teaching. The CRIOP contains five dimensions: classroom relationships, assessment practices, instructional practices, thoughtful discourse, and sociopolitical consciousness, rated on a 4-point scale (1 = Not at all, 2 = Occasionally, 3 = Often, 4 = To a great extent). A post-observation interview was conducted using a think-aloud procedure in which participants were asked to verbalise their thoughts while they viewed five selected sequences of the video observations. Additionally, open-ended questions aiming at exploring participants' beliefs about cultural diversity in education constituted a further source of data.

Data. The data are drawn from four teachers who worked in an ethnically diverse school located in Berlin, Germany. Three classroom observations for each were analyzed by rating each dimension of CRIOP. For analysing participants' self-reflection responses, we coded answers to one of three self-reflection categories (description, explanation, and integration) identified previously in the literature (e.g., Blomberg et al., 2014). Answers to open-ended questions related to cultural diversity beliefs were analyzed using a thematic analysis with a theoretical top-down approach (Braun & Clarke, 2006). Interrater agreement was established for both classroom observations and post-observation interviews.

Results. Our findings suggest that within the same school, teachers reported different degrees of culturally responsive practices as well as different beliefs towards cultural diversity. Importantly, these practices and beliefs are linked. Teachers with high scores in the CRIOP reported higher self-reflection, and held beliefs more multicultural diversity oriented. In contrast, teachers

with low scores in the CRIOP reflected less critically about their own strategies, and endorsed more egalitarian and colorblind beliefs about diversity.

## A19

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H001

### Strategien zum Verstehen von Fachtexten

Chair(s): **Anke Schmitz** (Universität zu Köln, Deutschland), **Fabiana Karstens** (Universität zu Köln), **Jörg Jost** (Universität zu Köln), **Anselm Strohmaier** (Technische Universität München), **Matthias C. Lehner** (Technische Universität München), **Kristina Reiss** (Technische Universität München)

DiskutantIn(nen): **Michael Becker-Mrotzek** (Universität zu Köln, Deutschland)

Sowohl im schulischen Bereich als auch in der Hochschule erfolgt die Wissensvermittlung bzw. der Wissenserwerb weitgehend textbasiert, weshalb das Verstehen und die zielgerichtete Nutzung von Texten als Schlüsselqualifikationen bezeichnet werden (Becker-Mrotzek et al., 2013; Schnotz, 1994; Weis et al., 2016). Im engeren Sinn beinhaltet das etwa das Verstehen von Sachtexten im Fach Deutsch und in den Naturwissenschaften oder die Verarbeitung von Textaufgaben in der Mathematik. Solche Fachtexte fordern aus psychologischer Perspektive eine „strategische, ziel-orientierte Textverarbeitung“ (Golke, Matthäi & Artelt, 2013) auf Seiten der Lernenden. Selbstreguliertes, strategisches Lesen gewinnt daher über die Sprachdidaktik hinaus an Bedeutung für die Forschung zu qualitativollen Lehr-Lernprozessen.

Die Vermittelbarkeit von Lesestrategien und Selbstregulationstechniken und die Bedeutung derer für die Informationsaufnahme und für das Textverständnis wurde vielfach nachgewiesen (u.a. Dignath & Büttner, 2008; Labuhn et al., 2008; Schüneman et al., 2013). Fragen nach der Spezifik des strategischen, selbstregulierten Lesens von fachlichen Texten im Hinblick auf ihre prototypischen Leseanforderungen wurden in den Studien der bisherigen pädagogisch-psychologischen Forschung allerdings nur in geringem Umfang betrachtet, welches sowohl für den schulischen als auch für den universitären Bereich gilt. Dabei erscheint eine differenzierte Betrachtung von strategieorientierten Lern- bzw. Leseprozessen unter Berücksichtigung der Fächerkontexte für die Unterrichtsentwicklung und Lehrerbildung sowie für die universitäre Lehre wesentlich. Es gibt folglich zahlreiche Ansatzpunkte für interdisziplinäre Forschung, was die Fachdidaktiken untereinander und einen Austausch mit der pädagogischen Psychologie und der Lehr-Lernforschung betrifft.

In diesem Symposium mit Vertretern aus der Bildungsforschung, der Mathematik- und Sprachdidaktik werden unterschiedliche Perspektiven auf den strategischen Zugang für das Verarbeiten von Informationen aus fachlichen Texten am Beispiel der Fächer Deutsch, Mathematik und Biologie präsentiert. Ausgehend von einer Analyse der Bedeutung von unterschiedlichen Strategien für das Lernen mit Text-Bild-Kombinationen biologischer Fachtexte und einem spezifischen Fokus auf die Informationsaufnahme und -verarbeitung von mathematischen Aufgaben bei Schüler/-innen und Studierenden werden fächervergleichend Perspektiven dargestellt, wie sich strategieorientierter Unterricht in Deutsch, Mathematik und Biologie unterscheidet. Ziel ist es, fächerübergreifende Gemeinsamkeiten und Besonderheiten des strategieorientierten Zugangs zu Fachtexten zu ermitteln, um das Konstrukt selbstreguliertes Lesen zu spezifizieren und Erkenntnisse für die Professionalisierung von Lehrkräften und die Entwicklung eines sprachsensiblen Fachunterrichtes zu generieren.

Der erste Beitrag befasst sich mit Strategien für das Lernen mit illustrierten Sachtexten des Faches Biologie. Es werden Ergebnisse einer quasi-experimentellen Studie vorgestellt, welche die Effekte einzelner Lerntechniken auf den Lernerfolg der Schüler/-innen untersucht. Der zweite Beitrag stellt Studien mit Schüler/-innen vor, die das Lesen von Aufgaben in der Mathematik und die Bedeutung des Anfertigns von Notizen im Rahmen der Informationsaufnahme und -verarbeitung thematisieren. Im dritten Beitrag wird diskutiert, wie sich Veränderungen in der Darstellung mathematischer Aufgaben auf die Informationsaufnahme und -verarbeitung von Items bei Studierenden auswirken. Der abschließende Beitrag nimmt einer fächervergleichende Perspektive ein und stellt Ergebnisse einer Lehrerbefragung vor, in welcher das selbstregulierte Lesen, hinsichtlich seines Stellenwerts, der Strategienutzung sowie der Instruktionsqualität in den Fächern Deutsch, Mathematik und Biologie erhoben wurde.

In einer abschließenden Diskussion werden die Befunde zum strategieorientierten Verstehen von fachlichen Texten zusammengefasst und Konvergenzen sowie Kontroversen für die weitere Beforschung des Forschungsgegenstandes aufgezeigt.

*Beiträge des Symposiums*

### Strategisches Lernen mit Text-Bild-Kombinationen im naturwissenschaftlichen Unterricht

**Sabine Schlag**

Bergische Universität Wuppertal

Zahlreiche Forschungsarbeiten zeigen die Bedeutsamkeit von Strategien beim Lernen mit Texten (Mandl & Friedrich, 2006). Lernstrategien werden dabei als eine systematische Verknüpfung von Lerntechniken verstanden, die zielgerichtet ausgeführt werden, sich zunehmend automatisieren, jedoch bewussteinfähig bleiben (Streblo & Schiefele, 2006). Im Gegensatz zum Lernen mit Texten liegen für das Lernen mit Text-Bild-Kombinationen noch wenige Lernstrategien vor (Plötzner et al., 2009; Scheiter et al., 2015). Meist wird angenommen, dass sich beim Lernen mit illustrierten Texten im Vergleich zum Lernen mit Texten ohne Bilder der Lernerfolg verbessert (Levie & Lentz, 1982; Mayer, 2001). Es zeigt sich jedoch, dass Lernende nicht immer in der Lage sind, Text-Bild-Kombinationen angemessen zu verarbeiten (Brünken et al., 2005; Weidenmann, 1991). Dies ist insbesondere dann der Fall, wenn auch die Bilder eine gezielte Informationsverarbeitung erfordern. Lernende benötigen folglich Unterstützung, um von Bildern in Texten profitieren zu können (Ainsworth et al., 2002; Seufert, 2003). Im Schulkontext sind dabei insbesondere im naturwissenschaftlichen Unterricht, wie z.B. in Schulbüchern für den Biologieunterricht, illustrierte Sachtexte zu finden, bei denen die Kombination aus Text und Bild zu einem tieferen Verständnis der Inhalte beitragen soll (Ainsworth, 2006).

Auf der Grundlage von Theorien zum multimedialen Lernen (Mayer, 2001) wurde eine Strategie für das Lernen mit illustrierten Sachtexten entwickelt (Schlag, 2011). Dabei wurden Lerntechniken kombiniert, welche die vier relevanten Informationsverarbeitungsprozesse Selektion, Organisation, Integration und Transformation anregen (ebd.). In mehreren empirischen Untersuchungen konnte die Lernwirksamkeit der Strategie nachgewiesen werden (Schlag, 2011). Klauer (2010) merkt jedoch an, dass sich bei komplexen Lernstrategien die Frage nach Inferenzen stellt. Dabei können die kombinierten Lerntechniken sowohl zu einer Steigerung als auch zu einer Abschwächung des Lernerfolgs im Vergleich zur Anwendung einer einzelnen Technik führen. Für einzelne Techniken konnten bereits lernförderliche Effekte gezeigt werden (z.B. Skizze erstellen: Leutner & Schmeck, 2014).

In diesem Beitrag wird der Frage nachgegangen, ob die einzelnen Lerntechniken der Strategie zu höherem, gleichem oder geringerem Lernerfolg beim Lernen mit illustrierten Sachtexten führen, als die Anwendung der gesamten Lernstrategie. Dies ist insbesondere von Interesse, weil sich das Training der komplexen Lernstrategie meist als recht zeitaufwändig erweist und die gefunden Effekte nicht immer die hohen Erwartungen bestätigen können (Klauer, 2010; Scheiter et al., 2015).

In einer quasi-experimentellen Studie wurden Schüler/-innen aus sechsten Klassen von Gymnasien und Gesamtschulen in vier Strategie-Gruppen eingeteilt. Dabei wurde eine Gruppe aufgefordert, die gesamte Lernstrategie (Gruppe (a) Überblick verschaffen, wichtige Text- und Bildelemente markieren und verknüpfen, zusammenfassen und skizzieren) anzuwenden. Die anderen drei Gruppen erhielten jeweils einzelne Techniken der Strategie: Gruppe (b) wichtige Text- und Bildelemente markieren und verknüpfen, Gruppe (c) zusammenfassen und Gruppe (d) skizzieren. Nach einem Vortest (Fragen zum Behaltenswissen des Nachtests), wurden die Schüler/-innen in die Strategie bzw. Technik eingeführt. Anschließend hatten diese 40 Minuten Zeit, das Lernmaterial, bestehend aus vier Text-Bild-Kombinationen zu einem biologischen Thema (Tanz der Honigbiene), zu lernen und die Strategie (Gruppe a) bzw. Technik (Gruppen b bis d) anzuwenden. Anschließend folgte der Nachtest mit Fragen zum Behaltens-, Konzept- und Transferwissen (Anderson & Krathwohl, 2001).

Die Ergebnisse der Untersuchung werden derzeit ausgewertet. Bezogen auf die Lerntechnik des Zusammenfassens hat die Studie von Schlag (2011) bereits gezeigt, dass die Lernenden, welche die Strategie anwendeten, der Zusammenfassungsgruppe signifikant überlegen waren (Vorstudie:  $F(3,56) = 10.19, p < .01, \eta^2 = .35, d=1.2$ ; Hauptstudie:  $F(1,130) = 24.55, p < .01, \eta^2 = .16, d=.98$ ). Ob sich diese Ergebnisse erneut replizieren lassen und welchen Einfluss auf den Lernerfolg die anderen beiden Lerntechniken haben, soll auf der Konferenz vorgestellt werden. Diese Ergebnisse können dazu beitragen, zu verstehen, welche Strategien bzw. Techniken beim Lernen mit illustrierten Sachtexten am besten im Unterricht vermittelt werden sollten.

### **Fachspezifische Leseprozesse und Notizen - Empirische Studien zu zwei irritierenden Elementen im kompetenzorientierten (Mathematik-)Unterricht**

**Dominik Leiss, Jennifer Plath**  
Leuphana Universität Lüneburg

Spätestens seit den empirischen Zusammenhängen zwischen sprachlichen und mathematischen Performanzen bei PISA 2000 und der Einführung kompetenzbasierter Bildungsstandards müssen sich die Lehrkräfte aller Unterrichtsfächer mit sprachlichen Aspekten im Rahmen von Lehr-Lernprozessen beschäftigen (Becker-Mrotzek et al., 2013).

So verlangen fachspezifische Kompetenzen wie z.B. mathematisches Modellieren, dass Schüler/-innen sich nun auch im Mathematikunterricht mit längeren (>2 Zeilen) lebensweltlichen Texten beschäftigen und in einen mündlichen und/oder schriftlichen Diskurs treten. Dabei scheint dieser fachliche Diskurs in besonderem Maße durch die Dualität von Sprache, also dem Zusammenspiel von kommunikativen und kognitiven Aspekten, geprägt zu sein. Detailliertere Erkenntnisse stellen jedoch noch ein Forschungsdesiderat dar (Paetsch et al., 2016). Für den Mathematikunterricht, der im Wesentlichen durch aufgabenbasierte Lehr-Lernumgebungen bestimmt wird, gilt es entsprechend die verwendeten Aufgabebearbeitungen nicht nur hinsichtlich der mathematischen Verfahren, sondern insbesondere hinsichtlich der sprachbasierten Verständnisprozesse zu analysieren.

Entsprechend werden im Vortrag zwei empirische Studien vorgestellt, die den mixed-method-Ansatz verfolgen und inhaltlich auf den strategischen Verstehensprozess mathematischer Problemstellungen und darauf aufbauend auf den bisher eher unsystematisch gelehnten epistemischen Schreibprozess des Anfertigen von Notizen fokussieren. In der ersten Laborstudie wurde anhand von 165 individuellen Aufgabenlösungsprozessen mit Lautem Denken der Verstehensprozess von Schüler/-innen der 7. Jahrgangsstufe qualitativ empirisch rekonstruiert. Hierzu wurden die videographierten und transkribierten Laborsitzungen mittels qualitativer Inhaltsanalyse zunächst mit einem hoch-inferenten Rating kodiert und die in diesem Rahmen erhaltenen quantifizierten Prozessdaten mit den durch Fragebögen und Tests erhobenen Hintergrunds- und Leistungsdaten (u.a. Lesegewohnheiten, Migrationshintergrund, allgemeine mathematische Fähigkeit, Leseverständnis, Lesegeschwindigkeit) ausgewertet. Anhand von Regressionsanalysen zeigte sich, dass die mathematische Fähigkeit, das Leseverständnis und die Komplexität der jeweiligen Aufgabe einen signifikanten Einfluss auf das Bilden eines adäquaten Situationsmodells (Leiss 2010) einer mathematischen Aufgabe haben ( $F(3,98) = 7.472; p < .01; \text{corr. } R^2 = .161$ ). Es zeigte sich allerdings auch, dass mit einer Ausnahme sich keine der von den Schüler/-innen angewendeten Lesestrategien positiv auf das Situationsmodell oder die Aufgabenlösung auswirkt. Lediglich das Anfertigen von Notizen hat einen geringen aber signifikant positiven Einfluss sowohl auf die Konstruktion eines adäquaten Situationsmodells als auch auf die erfolgreiche Bearbeitung der Aufgabe. Letztgenannter Effekt wird dabei durch die Komplexität der Aufgaben moderiert ( $F(23,135) = 2.601; p < .01; \text{corr. } R^2 = .189$ ).

Um das Anfertigen von Notizen, also die Fähigkeit eine fachspezifische Testsorte zu erstellen, näher zu untersuchen, wurden in einer quantitativen Leistungsteststudie offene Mathematikaufgaben auf verschiedenen Sprachniveaus konstruiert und in Klassentestungen mit insgesamt 1346 Schüler/-innen der Jahrgangsstufen 7 und 8 eingesetzt. Neben verschiedenen mit Fragebögen erhobenen Hintergrundvariablen und spezifischen Leistungsdaten wurde nicht nur die mathematische Korrektheit des Bearbeitungsprozesses, sondern insbesondere die selbst erstellten Notizen erfasst, kodiert und bewertet. Hierbei wurde nicht nur der Einfluss der sprachlichen Komplexität einer Aufgabe auf die Lösungsrate erneut bestätigt (Plath & Leiss, accepted), sondern unterschiedliche Notizentypen konnten reliabel identifiziert und bezüglich ihres Zusammenhangs mit der erfolgreichen Aufgabebearbeitung in Beziehung gesetzt werden.

### **Schwierigkeitsbezogene Unterschiede beim Lesen von Mathematikaufgaben mit offenem und klar definiertem Anfangs- und Zielzustand**

**Matthias C. Lehner, Kristina Reiss**  
Technische Universität München

In der aktuellen mathematikdidaktischen Forschung zur Schwierigkeit von Mathematikaufgaben wird häufig die Komplexität einer Aufgabe mithilfe von vordefinierten Kategoriensystemen eingeschätzt. Meist werden dabei die Prozesse bei der Aufnahme und Verarbeitung relevanter Informationen – dem Lesen – einer Aufgabe und die dabei verwendeten Strategien mit äußeren Merkmalen der Aufgabe in Verbindung gebracht. Die von unterschiedlichen Forschungsgruppen verwendeten Kategorien überschneiden sich teilweise (z.B. Cohors-Fresenborg et al., 2004; Kleinknecht et al., 2014; Neubrand et al., 2002; Turner et al., 2015). Jedoch konzentrieren sich diese Forschungsarbeiten vorwiegend auf Aufgaben in der Sekundarstufe, während nur wenige Studien die Schwierigkeit von Mathematikaufgaben auf Universitätsniveau thematisieren oder den Umgang von Studierenden mit unterschiedlich schwierigen Aufgaben fokussieren.

Da sich die Arbeitsweisen in der Mathematik an Schule und Hochschule unterscheiden, ist der Übergang zwischen Schulmathematik und Hochschulmathematik eine Herausforderung für viele angehende Studierende (Reichersdorfer, 2013). Mathematikaufgaben können oft nicht mehr mit Hilfe gewohnter Routinen gelöst werden und werden daher für Studierende zu Problemen. Ein Problem ist nach Newell und Simon (1972) eine Situation, in der eine Person ein Ziel erreichen will, aber zunächst nicht weiß, welche Mittel zu diesem Ziel führen. Oftmals werden Probleme mit einem Anfangszustand beschrieben, der zu Beginn eines Problemlöseprozesses vorgegeben ist, sowie einem Zielzustand, der im Problemlöseprozess aus dem Anfangszustand abgeleitet werden soll (z.B. Mayer, 1992). Für die Erklärung von Aufgabenschwierigkeiten ist zu berücksichtigen, dass Anfangs- und Zielzustand einer Aufgabe jeweils klar definiert oder offener formuliert sein können. Beispielsweise lässt sich ein offenerer Anfangszustand durch die Präsentation zusätzlicher Informationen operationalisieren, die alternative und komplexere Lösungswege ermöglichen. Ein offener Zielzustand lässt sich durch das Weglassen einer expliziten Formulierung des Ziels einer Aufgabe umsetzen.

Entsprechend der cognitive load theory (Sweller et al., 2011) können sich zusätzliche Informationen in der Angabe beziehungsweise ein offen gefasstes Ziel negativ auf den strategischen Lösungsprozess der Probleme auswirken. Durch sie entsteht zusätzlicher cognitive load beim Lösen einer Aufgabe. Wegen der begrenzten Verarbeitungskapazität des Arbeitsgedächtnisses können zusätzliche Informationen dazu führen, dass weniger Arbeitsgedächtniskapazität für die Lösung der Aufgabe genutzt werden kann. In diesem Zusammenhang wird in der vorliegenden Studie der Frage nachgegangen, inwieweit zusätzliche Informationen und eine offene Zielformulierung das Aufnehmen von Informationen in einer Problemstellung – das Lesen – und das strategische Lösen einer Aufgabe bei Studierenden beeinflussen.

Es wurden vier Variationen von neun verschiedenen Items auf Universitätsniveau entwickelt. Die Variationen beinhalten zwar dasselbe mathematische Problem, unterscheiden sich aber durch einen klar definierten beziehungsweise offenen Anfangs- und Zielzustand. Die Items wurden in zwölf Testheften angeordnet, die von jedem der neun Probleme eine Variation enthalten. Die Testhefte wurden 450 Studierenden der Life Sciences zu Beginn des ersten Semesters vorgelegt. Die Items wurden von einem geschulten Rater dichotom codiert und mithilfe der probabilistischen Testtheorie ausgewertet. Dies führt zu 36 geschätzten Aufgabenschwierigkeitsparametern, die anhand von neun Varianzanalysen mit dichotomen Faktoren „Klarheit des Anfangszustandes“ und „Klarheit des Zielzustands“ verglichen wurden.

In den Ergebnissen zeigt sich in sieben der neun Items, dass die Aufgabenschwierigkeit bei Aufgaben mit einem klar definierten Zielzustand signifikant geringer ist als bei Aufgaben mit einem offen formulierten Ziel ( $p < .02$ ). Hingegen ist kein klarer Zusammenhang zwischen der Aufgabenschwierigkeit und der Präsentation zusätzlicher Informationen in der Aufgabenstellung zu erkennen. Diese Befunde sind möglicherweise darauf zurückzuführen, dass neben der Charakterisierung der vorliegenden Items anhand objektiv beobachtbarer Merkmale individuelle kognitive Voraussetzungen der Studierenden eine Rolle für die individuellen Anforderungen von Mathematikaufgaben spielen.

Insgesamt lässt sich feststellen, dass die Schwierigkeit von Aufgaben auf Universitätsniveau objektiv nur schwer eingeschätzt werden kann. Dies kann daran liegen, dass sie vor allem von dem Vorwissen der angehenden Studierenden beeinflusst wird. Die Vorgabe eines klaren Zielzustandes hingegen erhöht die Lösungshäufigkeit einer Aufgabe und reduziert deren Schwierigkeit für Studierende signifikant.

## **Selbstreguliertes Lesen in den Fächern Deutsch, Biologie und Mathematik. Fachspezifische Unterschiede und überfachliche Gemeinsamkeiten**

**Anke Schmitz, Fabiana Karstens, Jörg Jost**  
Universität zu Köln

Eine Möglichkeit Schüler/-innen mit Blick auf hierarchiehöhere Fähigkeiten des Textverstehens zu fördern, besteht in der Befähigung zum selbstregulierten Lesen, d.h. der Nutzung von Lesestrategien und der eigenständigen Regulation des Leseprozesses. Die Bedeutung von kognitiven und metakognitiven Strategien für das Leseverständnis wurde durch Bedingungsmodelle nachgewiesen (u.a. Cromley & Azevedo, 2007). Darüber hinaus besteht Evidenz darüber, dass Trainings zur Förderung des selbstregulierten Lesens leistungssteigernd auf das Leseverständnis wirken (Duke & Pearson, 2002; Mokhesgerami et al., 2007; Schünemann et al., 2013), insbesondere wenn Selbstregulationstechniken und Lesestrategien mit fachlichen Inhalten verknüpft werden (Perels et al., 2005). Die Mehrheit der benannten Studien fokussiert jedoch auf fächerübergreifende Selbstregulationstechniken (Kistner et al., 2015; Labuhn et al., 2008), sodass fachspezifische Leseanforderungen und -strategien unberücksichtigt bleiben. Wie eine passgenaue, fachbezogene Förderung von strategischem Lesen zur Verbesserung des Leseverständnisses aussehen könnte, gilt als fraglich, da detaillierte Erkenntnisse zum strategieorientierten Lesen in der fachlichen Unterrichtspraxis ein Forschungsdesiderat darstellen.

In diesem Beitrag wird eine Studie vorgestellt, in der Lehrkräfte dazu befragt wurden, wie sie mit ihrem Unterrichtsangebot (Helmeke, 2004) Formen des selbstgesteuerten Lesens im Fachunterricht ermöglichen. Ziel ist es, fachspezifische Differenzen sowie fächerübergreifenden Konvergenzen beim strategieorientierten Lesen von prototypischen Fachtexten in den Fächern Deutsch, Mathematik und Biologie zu analysieren. Die Operationalisierung des selbstregulierten Lesens basiert auf der Lernstrategietaxonomie von Schiefele und Pekrun (1996), dem Lesestrategieinventar von Mandl und Friedrich (2006) sowie empirischer Evidenz zur Instruktionsqualität des selbstregulierten Lesens mittels kognitiver und metakognitiver Strategien (Dole et al., 2009; Duke & Pearson, 2002; Pressley et al., 1998). An der Befragung nahmen 227 Lehrkräfte aus 23 weiterführenden Schulen teil. Sie wurden zum selbstregulierten Lesen von Sachtexten (Deutsch und Biologie) bzw. von Textaufgaben (Mathematik) befragt. Neben der Selbsteinschätzung der Häufigkeit des Einsatzes ausgewählter kognitiver Lesestrategien und der Anregung metakognitiver Regulationsprozesse (Zimmerman, 2002), wurden die Lehrkräfte zur Instruktionsqualität bei der Vermittlung von Lesestrategien (Mandl & Friedrich, 2006; Duke & Pearson, 2002) sowie zur Relevanz von Lesen und Lesestrategien in ihrem Fach befragt (in Anlehnung an PISA, 2006).

Die Ergebnisse zeigen, dass sich die Deutschlehrkräfte in der Beurteilung der Relevanz von Lesestrategien für die Textarbeit und der Betonung der Relevanz von Lesestrategien und Metakognition im Unterricht signifikant von den Mathematik- und Biologielehrkräften unterscheiden ( $F(2, 225) = 21.06, p < .001, \eta^2 = .16$ ). Auch die Anregung zu metakognitiven Reflexionsprozessen scheint für die Mathematik- und Biologielehrkräfte eine weniger wichtige Rolle zu spielen als für Deutschlehrer/-innen. Beim Einsatz von kognitiven Lesestrategien im Unterricht liegt ein signifikanter wechselseitiger Unterschied vor ( $F(2, 226) = 36.94, p < .001, \eta^2 = .25$ ). Die Deutschlehrkräfte bekunden, kognitive Lesestrategien in den meisten Unterrichtsstunden zu nutzen, gefolgt vom Einsatz in einigen Stunden im Fach Mathematik und wenigen Stunden in Biologie. Dieses fachspezifische Bild zeigt sich auch bei der spezifischen Betrachtung der Nutzung von Organisationsstrategien. Elaborations- und Wiederholungsstrategien hingegen werden von Deutsch- und Mathematiklehrer/-innen gleich oft und häufiger als von Biologielehrkräften genutzt. Konvergenzen von Deutsch- und Mathematiklehrkräften betreffen zudem die

Instruktionsqualität bei der Vermittlung von Lesestrategien. Die Lehrkräfte bekunden, den Nutzen und die Anwendung eher explizit zu erklären im Gegensatz zu den Biologielehrkräften  $F(2, 213) = 19.04$ ,  $p < .001$ ,  $\eta^2 = .15$ ).

Die Ergebnisse verdeutlichen, dass die Thematik des selbstregulierten Lesens von Deutschlehrkräften zwar als wichtiger angesehen wird, jedoch auch deutliche überfachliche Gemeinsamkeiten, insbesondere den Mathematik- und Deutschunterricht betreffend, vorliegen. Zum aktuellen Zeitpunkt werden Unterrichtsbeobachtungen ausgewertet, die in 16 Schulen durchgeführt wurden ( $N = 84$  beobachtete Unterrichtsstunden bzw. 3.780 Minuten, davon 42 Deutschstunden und jeweils 21 Mathematik- und Biologiestunden), wodurch ein Perspektivabgleich mit der Lehrerbefragung erfolgen kann (Clausen, 2003). Entsprechende Erkenntnisse können für die Entwicklung fachspezifischer Lesestrategietrainings, die Lehrerprofessionalisierung und eine sprachensible Unterrichtsentwicklung genutzt werden.

## A20

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Kollegienhaus H120

### Kompetenz von Lehr- und Fachkräften in verschiedenen Bildungsbereichen: Entwicklung von Instrumenten und Forschungsergebnisse

Chair(s): **Andrea Wullschleger** (Universität Zürich, Schweiz), **Simone Dunekacke** (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel)

DiskutantIn(nen): **Lars Jenßen** (Humboldt-Universität zu Berlin)

Sowohl im Elementar- als auch im Primarbereich sind Fragen nach dem komplexen Zusammenwirken verschiedener Elemente die zum Gelingen von Lehr-Lernprozessen führen von großer Bedeutung und werden in sog. Angebots-Nutzungs-Modellen (Helmke, 2015) bzw. Struktur-Prozess-Modellen (Tietze, 2013) modellhaft dargestellt. Erst hierdurch wird es möglich, einzelne Elemente und ihr Zusammenwirken detailliert zu untersuchen. Insbesondere das Element „professionelle Kompetenz der Lehr- und Fachkräfte“ und dessen Auswirkungen auf die Qualität von Lernangeboten sowie die Kompetenzentwicklung der Lernenden sind zu einem zentralen Gegenstand der empirischen Forschung geworden.

Im Primar- und Sekundarbereich werden die Kompetenzen der Lehrkräfte sowie deren Entwicklung seit längerem eingehend untersucht (Blömeke, Kaiser & Lehmann, 2010; Kunter, Baumert, Blum & Neubrand, 2011), allerdings nicht für alle Fächer in gleichem Umfang. Für das Fach Deutsch liegen deutlich weniger Erkenntnisse vor als für die Mathematik bzw. das naturwissenschaftliche Lernen (Blömeke et al., 2011). Die empirische Untersuchung professioneller Kompetenzen frühpädagogischer Fachkräfte hat erst deutlich später begonnen (Gasteiger & Benz, 2016) und ist nach wie vor wenig erforscht (Anders & Rossbach, 2015).

Zielsetzung des Symposiums:

Zielsetzung des Symposiums ist zum einen, für Fächer und für Bildungsbereiche, in denen noch wenige Erkenntnisse vorliegen, namentlich das Fach Deutsch und die professionellen Kompetenzen von Fachkräften im Kindergarten, die Entwicklung von Instrumenten in den Blick zu nehmen. Zum anderen soll ein Beitrag dazu geleistet werden, das Zusammenwirken der professionellen Kompetenz von Lehr- und Fachkräften mit anderen Elementen des Angebots-Nutzungs-Modells am Beispiel von Studien aus verschiedenen Bildungsbereichen und Fächern besser zu verstehen.

Beiträge:

Der erste Beitrag befasst sich mit der Erfassung professioneller Kompetenzen von frühpädagogischen Fachkräften im mathematischen Bereich. Vorgestellt wird ein Instrument, in dem durch den Einsatz von Video-, Bild- und Textvignetten handlungsnahe Kompetenzen von frühpädagogischen Fachkräften untersucht werden.

Der zweite Beitrag ist auf der Primarstufe und im Bereich der Lesedidaktik und Lesediagnostik angesiedelt. Es wird ein Instrument vorgestellt, mit dem das Wissen von Lehrkräften mit unterschiedlichen Ausbildungen in den Grundlagen der Lesedidaktik und Lesediagnostik unter besonderer Berücksichtigung von Leseschwächen erfasst werden kann.

Im dritten Beitrag werden die Zusammenhänge zwischen naturwissenschaftlichen Fortbildungen, professionellem Austausch im Team und der Häufigkeit naturwissenschaftlicher Lernangebote untersucht und diskutiert, welchen Beitrag diese Professionalisierungsmaßnahmen für die Implementation des naturwissenschaftlichen Bildungsbereichs leisten können

Im vierten Beitrag wird die Entwicklung des mathematischen und mathematikdidaktischen Wissens von Primar- und Sekundarlehrkräften während des Studiums und während den ersten beiden Berufsjahren untersucht. Neben dem Effekt der Ausbildung und der Berufserfahrung analysieren die Forschenden auch den Einfluss von Berufswahlmotiven.

Abschließend erfolgt eine Diskussion der vorgestellten Studien, die einerseits die gewählten Ansätze einordnet und andererseits deren Potenzial für die Weiterentwicklung des Forschungsstands diskutiert.

*Beiträge des Symposiums*

### Struktur professioneller Kompetenz und deren Erfassung bei frühpädagogischen Fachkräften im Bereich Mathematik

**Selma Seemann<sup>1</sup>, Anke Lindmeier<sup>1</sup>, Miriam Leuchter<sup>2</sup>, Elisabeth Moser Opitz<sup>3</sup>, Franziska Vogt<sup>4</sup>, Aiso Heinze<sup>1</sup>**

<sup>1</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel, <sup>2</sup>Universität Landau, <sup>3</sup>Universität Zürich, <sup>4</sup>Pädagogische Hochschule St.Gallen

Theoretischer Hintergrund und Fragestellung:

Aufgrund ihrer prädiktiven Bedeutung für die spätere Kompetenzentwicklung gilt die Förderung mathematischer Fähigkeiten als eine Kernaufgabe frühpädagogischer Einrichtungen. Dabei wird angenommen, dass die professionellen Kompetenzen der Fachkräfte die Qualität der Lernumgebung und darüber den Fähigkeitserwerb bei den Kindern beeinflussen (Gasteiger, 2012; Mashburn & Pianta, 2010). Bisher fehlen differenzierte Erkenntnisse über die Struktur von insbesondere fachspezifischen professionellen Kompetenzen von pädagogischen Fachkräften. Diese sind für die Erforschung von mathematikbezogenen Bildungsprozessen sowie der Gestaltung und Evaluation von Aus- und Fortbildungen der Fachkräfte von Bedeutung.

Insbesondere in Deutschland stellt die Erfassung von mathematikspezifischer professioneller Kompetenz eine Herausforderung dar. Viele Fachkräfte haben aufgrund des nicht-akademischen Ausbildungssystems kaum systematische Lerngelegenheiten zum Erwerb mathematikdidaktischen Wissens. Entsprechend dürfte für die Performanz der Fachkräfte implizites Wissen eine wichtige Rolle spielen (Dunekacke et al., 2016). Zur Erfassung bietet sich deshalb das Strukturmodell professioneller Kompetenz von Lindmeier (2011) an, das ausgehend von antizipierten fachspezifischen Anforderungen einer pädagogischen Fachkraft zwei Kompetenzdimensionen konzeptualisiert: die Reflexive Kompetenz (RC), welche Fähigkeiten zur adaptiven Vorbereitung von Lerngelegenheiten und zur Analyse von Lernprozessen ohne Zeitdruck umfasst, und die Aktionsbezogenen Kompetenz (AC), die Fähigkeiten der spontanen, unmittelbaren und kognitiv aktivierenden Lernbegleitung in pädagogischen Situationen unter Zeitdruck beinhaltet (Lindmeier et al., 2016). Ergebnisse von Hepberger et al. (2017) deuteten bereits an, dass eine reliable Messung von RC und AC möglich ist und diese empirisch trennbar sind. Diese Studie unterlag jedoch Einschränkungen, da u.a. die komplexe Erhebung von AC (s.u.) als Gruppentestung und nicht als computergestützte Einzeltestung erfolgte. Ziel dieses Beitrags ist einerseits die Replikation der Ergebnisse von Hepberger et al. (2017) (reliable Erfassung, Trennbarkeit von RC und AC) unter optimierten Erhebungsbedingungen. Andererseits sollen Indikatoren für die Validität der Messungen untersucht werden.



#### Methode:

Die Stichprobe umfasst 161 pädagogische Fachkräfte (80 aus Deutschland und 81 aus der Schweiz). Aufgrund der unterschiedlichen Ausbildungssysteme in den beiden Ländern (beruflich vs. akademisch) werden u.a. Unterschiede in der mathematikspezifischen professionellen Kompetenz erwartet.

Zur Erfassung der Kompetenzkomponenten RC und AC wurden Items von Hepberger et al. (2017) adaptiert und neue Items entwickelt. Der Test umfasst 9 AC-Items sowie 12 RC-Items und ist inhaltlich am Bereich Zahlen, Mengen und Operationen ausgerichtet. Um eine für den Kindergartenalltag typische Anforderungsspezifik abzubilden, enthält jedes Item eine kurze Videovignette (AC) bzw. eine Bild- oder Textvignette (RC). Die Beantwortung erfolgt an einem Laptop für AC-Items jeweils unter Speedbedingung mündlich in ein Head-Set und für RC-Items unter Powerbedingung schriftlich. Die Antworten wurden gemäß eines Kodierschemas kodiert. Pro Land wurden Daten von 23 randomisiert gezogenen Fachkräften doppelt kodiert, wobei sich eine gute Interraterreliabilität zeigte (Cohens Kappa RC/AC=.92/.90).

#### Ergebnisse:

Die Kodierung der RC-Items ist für die Gesamtstichprobe abgeschlossen, während für AC bisher nur Daten der 46 doppelt kodierten Fälle vorliegen. Die Auswertungen mit AC-Items beschränken sich auf diese Teilstichprobe. Erste Analysen zeigen eine akzeptable Reliabilität (Cronbachs  $\alpha$  für RC/AC=.65/.66) und es ergaben sich keine Boden- und Deckeneffekte. Die Skalen korrelieren in mittlerer Höhe ( $r=.37^{***}$ ), sodass RC und AC erwartungsgemäß trennbar sind (vgl. auch Knievel et al., 2015 für Grundschullehrkräfte). Des Weiteren bilden die Skalen ab, dass die akademisch ausgebildeten Fachkräfte aus der Schweiz deutlich bessere Kompetenzen aufweisen als die beruflich ausgebildeten Fachkräfte aus Deutschland (RC:  $M=12.75$ ,  $SD=3.60$  vs.  $M=9.11$   $SD=3.75$ ,  $t(159)=6.28^{***}$ ,  $d=0.99$ ; AC:  $M=10.00$ ,  $SD=2.91$  vs.  $M=6.36$ ,  $SD=3.05$ ,  $t(44)=3.93^{***}$ ,  $d=1.22$ ).

Bis zum Vortrag werden die Analysen (u.a. CFA, IRT) des vollständigen Datensatzes zur Skalengüte vorliegen und wir erwarten eine Bestätigung der vorläufigen Ergebnisse. Die Validität des zweidimensionalen Modells soll anschließend durch Überprüfung von differenziellen Effekten verschiedener Fortbildungen auf RC bzw. AC weiter gestützt werden. Schließlich werden Zusammenhänge von RC und AC zur mathematikspezifischen Prozessqualität (als Performanzindikator für Fachkräfte) analysiert und der Erklärungsgehalt des zweidimensionalen Modells untersucht.

## **Erheben von Wissen zu Grundlagen der Lesedidaktik und zur Diagnostik bei Leseschwierigkeiten: Entwicklung eines Instruments**

**Meike Lietz**

Pädagogische Hochschule Bern

#### Theoretischer Hintergrund:

Professionelle Kompetenzen von Lehrkräften wurden bislang vor allem im Bereich Mathematik und Naturwissenschaften (Kunter et al., 2011) untersucht. Auch wenn inzwischen Studien zu anderen Fächern vorliegen (z.B. Krauss et al., 2017), muss festgestellt werden, dass es nur wenige Studien gibt, die sich mit der Kompetenz von Lehrkräften im Deutschunterricht befassen (Blömeke et al., 2011). Ausnahmen stellen hier die Studien PlanvoLL-D (Bremerich-Vos et al., 2016) sowie FALKO-D (Pissarek & Schilcher, 2017) und für die Domäne Schreiben die Studien von Corvacho del Toro (2013) und Sturm et al. (2016) dar. Studien, die explizit auf die Kompetenzmessung bei Lehrpersonen im Bereich Lesen abzielen, gibt es - abgesehen von der Studie von Rutsch (2016) vor allem im US-Amerikanischen Raum (u.a. Carlisle et al., 2011; Moats & Foorman, 2003; Piasta et al., 2009). Dies ist insbesondere deshalb erstaunlich, weil Schwierigkeiten beim Lesen (und auch beim Schreiben) gemäß Hartmann (2013, 100) zu den häufigsten Lernproblemen gehören. Zur Schließung dieser Forschungslücke will die vorliegende Studie einen Beitrag leisten.

#### Zielsetzung:

Ziel der vorgestellten Studie ist die Entwicklung eines Instruments, mit dem das professionelle Wissen von Lehrkräften mit unterschiedlichem Ausbildungshintergrund in den Bereichen Grundlagen der Lesedidaktik und Lesediagnostik unter besonderer Berücksichtigung von Leseschwächen erfasst werden kann. Das Instrument wird im Hinblick auf psychometrische Gütekriterien und die Dimensionalität beurteilt. Zudem wurde untersucht, ob sich die Leistungen von Personen mit unterschiedlichen Ausbildungen unterscheiden.

#### Methode:

Die theoretische Konzipierung des Instrumentes basiert auf dem Kompetenzmodell zum Professionswissen von Baumert & Kunter (2006). Das Modell berücksichtigt die auf Shulman (1987) zurückgehenden Bereiche 'Fachwissen' und 'fachdidaktisches Wissen'. Das fachliche und fachdidaktische Wissen wurde zu folgenden zwei Inhaltsbereichen erfasst: 'Grundlagen zum Leseerwerb und Lesekompetenz' und 'Diagnostik bei Leseschwierigkeiten'. Hierzu wurden verschiedene Facetten erfasst wie z.B. Wissen über Phasen der Leseentwicklung, über Symptome von Leseschwierigkeiten sowie zu diagnostischen Verfahren.

Aus einem Pool von anfänglich 56 Items wurden auf der Basis der Ergebnisse qualitativer Pretests 24 Items ausgewählt. Daraus wurde ein Paper-Pencil Instrument (u.a. MC-Fragen, Audioitems) entwickelt und 181 Lehrpersonen vorgelegt.

Um ein möglichst großes Kompetenzspektrum zu erfassen, wurden drei Personengruppen befragt, von denen angenommen wird, dass sie sich hinsichtlich ihres Wissens unterscheiden: Lehramtstudierende bei Studienabschluss (STUD\_Abschluss  $n=66$ ); Grundschullehrkräfte mit mindestens zwei Jahren Berufserfahrung (GLK\_Erfahr  $n=61$ ); Grundschullehrkräfte mit mindestens zwei Jahren Berufserfahrung, die ein Aufbaustudium für das Lehramt Sonderpädagogik absolvieren am Ende des Studiums (GLK\_Aufbau  $n=54$ ).

Zur Abgrenzung des Professionswissens wurde das Instrument zudem einer Kontrollgruppe von Laien ( $n=22$ ) vorgelegt (Daten werden hier berichtet).

#### Ergebnisse:

Die Itemanalysen erfolgten mit dem eindimensionalen Raschmodell. Von den 24 Items erfüllten 15 die statistischen Anforderungen (MNSQ: .88 bis 1.11, rit .26 bis .46). Der Wert für die EAP Reliabilität ist mit .70 knapp befriedigend. Als besonders herausfordernd erweist sich die Konstruktion von Audioitems.

Die dimensionale Unterscheidung zwischen 'Fachwissen' und 'Fachdidaktischem Wissen' konnte faktorenanalytisch nicht gefunden werden (für ein vergleichbares Ergebnis bezüglich mathematischem Lernen vgl. Jandl & Moser Opitz, 2017). Jedoch konnten die Dimensionen entlang der zwei Inhalte 'Grundlagen Leseerwerb' und 'Diagnostik' nachgewiesen werden.

Die drei Gruppen unterscheiden sich bezüglich ihres Wissens erwartungsgemäß signifikant (STUD\_Abschluss M = 9.15, SD = 2.64; GLK\_Erfahr M = 10.98, SD = 2.78; GLK\_Aufbau M = 14.65, SD = 3.48; F [3] = 52.85, p < .001,  $\eta^2$  = .39).

Fazit:

Mit den bisherigen Forschungsarbeiten konnte ein erster Schritt zum Erreichen der Zielsetzung gemacht werden. Die Ergebnisse der Faktorenanalyse werfen die Frage auf, was fachliches Wissen bezüglich von Inhalten wie Leseerwerb oder Lesediagnostik ist bzw. ob es sich hier nicht immer auch schon um fachdidaktisches Wissen handelt. Ausgehend von den bisher vorliegenden Ergebnissen wird eine Weiterentwicklung des Instruments geplant.

### **Naturwissenschaftliche Bildung in der Kita – welchen Beitrag können Fortbildungen sowie professioneller Austausch im Team leisten?**

**Julia Barenthien<sup>1</sup>, Elisa Oppermann<sup>2</sup>, Yvonne Anders<sup>2</sup>, Mirjam Steffensky<sup>1</sup>**

<sup>1</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel, <sup>2</sup>Freie Universität Berlin

Theoretischer Hintergrund:

Qualitativ hochwertige Lerngelegenheiten in den Naturwissenschaften gelten gemäß dem Angebot-Nutzungs-Modell (Helmke, 2015) als Voraussetzung für die naturwissenschaftliche Kompetenzentwicklung von Kindern. Damit Kinder überhaupt von diesen Lerngelegenheiten profitieren können, müssen sie in einer ausreichenden Häufigkeit angeboten werden (Piasta et al., 2015). Internationale Befunde weisen jedoch auf eine geringe Häufigkeit naturwissenschaftlicher Lerngelegenheiten in der Kita hin (Tu, 2006). Die regelmäßige Teilnahme an Fortbildungsmaßnahmen gilt als hilfreich, um Fachkräfte dabei zu unterstützen, naturwissenschaftliche Lernangebote umzusetzen. So deuten aktuelle Befunde auf einen Zusammenhang zwischen Fortbildungsteilnahme und Häufigkeit naturwissenschaftlicher Lerngelegenheiten in der Kita hin (Gropen et al., 2017). Häufig fehlen in Kitas jedoch die Ressourcen für eine langfristige Fortbildung aller Fachkräfte einer Institution. Jedoch wäre denkbar, dass Fortbildungsinhalte ins Team getragen werden und auch nicht fortgebildete Fachkräfte durch diesen professionellen Austausch anregt werden, naturwissenschaftliche Lerngelegenheiten häufiger umzusetzen. Erste Befunde zeigten bereits positive Zusammenhänge zwischen dem professionellem Austausch und der Prozessqualität im Kontext der frühen sprachlichen Bildung (Resa et al., 2017). Für die Häufigkeit der Lerngelegenheiten sind diese Mechanismen bisher aber kaum untersucht.

Fragestellung:

Ziel des Beitrags ist es, zu prüfen,

(1) inwieweit naturwissenschaftliche Fortbildungen das Aufgreifen naturwissenschaftlicher Inhalte in Dienstbesprechungen beeinflussen?

(2) inwieweit der professionelle Austausch über naturwissenschaftliche Inhalte in Dienstbesprechungen mit der Häufigkeit der naturwissenschaftlichen Lerngelegenheiten zusammenhängt?

(3) inwieweit das Aufgreifen naturwissenschaftlicher Inhalte in Dienstbesprechungen einen Einfluss auf die Häufigkeit der Lerngelegenheiten bei Fachkräften ohne Fortbildungsteilnahme hat?

Methode:

Die vorliegenden Analysen basieren auf einer Fragebogenerhebung mit 298 Fachkräften (M\_Alter= 37.61 Jahre; SD\_Alter= 11.15; M\_Berufserfahrung= 13.02 Jahre; SD\_Berufserfahrung= 10.87; 91% Frauen). Angaben zur Anzahl der naturwissenschaftlichen Fortbildungsteilnahmen im letzten Jahr wurde mit einer offenen Frage erhoben. Die Weitergabe von Fortbildungsinhalten in den Dienstbesprechungen der letzten drei Monate schätzen die Fachkräfte auf einer vierstufigen Skala von „kein Mal“ bis „auf allen Teamsitzungen“ ein. Diese Angaben wurden auf Kitaebene aggregiert. Die Häufigkeit naturwissenschaftlicher Lerngelegenheiten der vergangenen drei Monaten schätzten die Fachkräfte auf einer sechsstufigen Skala von „nie“ bis „täglich“ ein. Folgende Kontrollvariablen wurden bei den Analysen berücksichtigt: Häufigkeit der Dienstbesprechungen, Einrichtungsgröße und Anteil der U3-Kinder.

Für die Untersuchung der ersten beiden Fragestellungen wurde ein Pfadmodell in MPlus gerechnet. Die dritte Fragestellung wurde anhand einer Regressionsanalyse untersucht. Zur angemessenen Berücksichtigung der Mehrebenenstruktur der Daten wurde mit Type Complex eine Standardfehlerkorrektur durchgeführt.

Ergebnisse:

Die Ergebnisse des Pfadmodells zeigen einen Zusammenhang zwischen der Fortbildungsteilnahme und der Häufigkeit naturwissenschaftlicher Inhalte in den Dienstbesprechungen ( $\beta$  = .30; SE = .06; p < .01), welche wiederum mit einer höheren Anzahl an naturwissenschaftlichen Lernangeboten einhergeht ( $\beta$  = .15; SE = .07; p < .05). Fachkräfte ohne Fortbildung, in deren Dienstbesprechungen naturwissenschaftliche Inhalte angesprochen worden, boten signifikant häufiger naturwissenschaftliche Lerngelegenheiten an als nicht fortgebildete Fachkräfte mit Dienstbesprechungen ohne naturwissenschaftlichen Inhalte ( $\beta$  = .21; SE = .08; p < .05).

Diskussion:

Die vorliegenden Ergebnisse zeigen einen Zusammenhang zwischen naturwissenschaftsspezifischen Fortbildungsaktivitäten und der Häufigkeit naturwissenschaftlicher Inhalte in Dienstbesprechungen, die wiederum mit der Häufigkeit naturwissenschaftlicher Bildungsangebote zusammenhängt. Möglicherweise deuten sie darauf hin, dass die Wirkung bereichsspezifischer Fortbildungen sich erst dann entfalten kann, wenn in der Kita der Bildungsbereich mit hoher Qualität implementiert ist, wofür die Auseinandersetzung mit naturwissenschaftlichen Inhalten in Dienstbesprechungen ein Indikator ist. Dieser förderliche Effekt der Implementierungsqualität deutet sich auch für Fachkräfte ohne Fortbildung an. Einschränkend muss gesagt werden, dass die Daten aufgrund des querschnittlichen Designs keine Aussage über die Richtung der Zusammenhänge sowie die Qualität der Umsetzung zulassen. So wäre es auch denkbar, dass in Kitas, die Naturwissenschaften bereits in hoher Qualität implementiert haben, nicht nur häufiger Bildungsangebote umgesetzt als auch Fortbildungsangebote wahrgenommen werden. Gleichwohl machen die Ergebnisse auf die Bedeutung der Implementationsqualität von Bildungsbereichen in der Institution Kita aufmerksam, die zusätzlich zu Strukturmerkmalen und Indikatoren der Quantität und Qualität der Bildungsangebote in kommenden Studien berücksichtigt werden sollten.

### **Genese mathematischen und mathematikdidaktischen Wissens von Primar- und Sekundarlehrpersonen unter Berücksichtigung von Berufswahlmotiven**

Theoretischer Hintergrund:

Dem Lehrberuf als Profession, und damit verbunden dem Wissen und Können von Lehrpersonen, wird eine professionelle Kompetenz im Rahmen der Schule und ihrem spezifischen Berufsauftrag zugesprochen. Diese in Anlehnung an Weinert (2001) multidimensional konzipierte, als erwerb- und veränderbar zu betrachtende (Bromme, 2001; Darling-Hammond, 2006) und auf die Bewältigung berufsspezifischer Anforderungen ausgerichtete professionelle Kompetenz enthält kognitive sowie auch motivationale und handlungsbezogene Aspekte, welche sich gegenseitig beeinflussen (z.B. Baumert & Kunter, 2006; Blömeke, Kaiser, & Lehmann, 2008).

Der kognitive Aspekt des Professionswissens, bestehend aus fachwissenschaftlichem und fachdidaktischem Wissen in einzelnen Fächern sowie dem pädagogisch-psychologischen Wissen, gilt als zentraler Teil der professionellen Kompetenz, um Unterricht erfolgreich und lernwirksam planen und durchführen zu können (Baumert & Kunter, 2011a). Dies obwohl erst wenige empirische Erkenntnisse zu deren Wirkung auf die Schülerleistungen (Baumert & Kunter, 2011b; Hill, Rowan, & Loewenberg Ball, 2005) sowie die Unterrichtsqualität (z.B. Hattie, 2010; Kyriakides, Creemers, & Antoniou, 2009; Seidel & Shavelson, 2007) vorliegen. Zu den motivationalen Aspekten zählen beispielsweise Selbstwirksamkeitserwartungen und berufsbezogene Motive (Kunter, 2011). Auch sie stehen im Zusammenhang mit der erfolgreichen Bewältigung beruflicher Anforderungen (Rothland, 2014; Schwarzer & Warner, 2014).

Aufgrund der Annahme der Erwerb- und Veränderbarkeit professioneller Kompetenz wird der Lehrerausbildung eine dreifache Wirksamkeitserwartung zugeschrieben: (1) Die Lehrerausbildung soll zum Aufbau professioneller Kompetenzen und deren Weiterentwicklung im Beruf beitragen. (2) Die erworbenen Kompetenzen sollen in der Berufspraxis zu einer hohen Unterrichtsqualität führen, welche schliesslich (3) den Lernerfolg der Schülerinnen und Schüler über gelingende Lernprozesse positiv beeinflussen.

Fragestellung:

Die vorliegende Studie untersucht die erste Wirksamkeitserwartung der Lehrerausbildung, die Veränderung der professionellen Kompetenz im Sinne eines Angebots-Nutzungs-Modells (Helmke, 2012) unter Berücksichtigung individueller Merkmale. Konkret wird die Genese des fachlichen und fachdidaktischen Wissens in Mathematik unter Berücksichtigung von Berufswahlmotive von Lehrpersonen während der Ausbildung bis ins dritte Berufsjahr anhand folgender Forschungsfragen analysiert:

1. Wie entwickelt sich das mathematische und das mathematikdidaktische Wissen von Lehrpersonen der Primarstufe und der Sekundarstufe I
  - a. während des Studiums?
  - b. während den ersten beiden Berufsjahren?
2. Erweisen sich Berufswahlmotive als günstige Prädiktoren für Leistungsentwicklungen der Lehrpersonen in Mathematik und Mathematikdidaktik?

Methode:

Die längsschnittlichen Analysen beruhen auf 102 Primar- und 47 Sekundarlehrpersonen, von denen Daten zu Beginn (t1) und am Ende der Lehrerausbildung (t2) vorliegen. Zu einer Teilstichprobe von 44 Primar- und 30 Sekundarlehrpersonen liegen zudem Daten zu Beginn des dritten Jahres der Berufstätigkeit (t3) vor. Für die Analyse der Mittelwertvergleiche wurden die Testergebnisse zu den verschiedenen Messzeitpunkten als latente Variablen modelliert. Dabei wurden die Aufgabenparameter gemäss dem bei TEDS-M (Tatto et al., 2012) angewandten Skalierungsverfahren verankert. Die Berufswahlmotive sind durch insgesamt 16 Items erfasst, die in die vier Skalen fachbezogene Motivation (Cronbachs  $\alpha = .72$ ), extrinsische Berufsmotivation ( $\alpha = .58$ ), Interesse an den Schülerinnen und Schülern ( $\alpha = .64$ ) und Interesse an Vermittlung ( $\alpha = .61$ ) zusammengefasst wurden. Die Effekte der Berufswahlmotive auf die Entwicklung mathematischen und mathematikdidaktischen Wissens wurde, unter Kontrolle der Ausgangswerte bei den Wissenstests, regressionsanalytisch bestimmt.

Ergebnisse:

Erste Ergebnisse zeigen während des Studiums bei den Lehrpersonen der Primarstufe keine statistisch bedeutsame Veränderung im Mathematikwissen, jedoch im mathematikdidaktischen Wissen. Während den ersten beiden Berufsjahren zeigt sich weder eine signifikante Veränderung des Mathematikwissens noch des mathematikdidaktischen Wissens. Für die Sekundarstufe I findet sich hingegen sowohl in Mathematik als auch in Mathematikdidaktik ein statistisch bedeutsamer Anstieg des Wissens während der Lehramtsausbildung. Während den beiden ersten Berufsjahren bleiben auch hier beide Wissensformen unverändert. Die fachbezogene Motivation beeinflusst die Entwicklung des mathematischen Fachwissens bei den Primarlehrpersonen positiv, bei den Sekundarlehrpersonen hat das Interesse an der Vermittlung einen positiven Effekt auf das mathematische Fachwissen. Eine extrinsische Berufsmotivation ist für die Entwicklung des Mathematikwissens als kritisch zu betrachten.

## A21

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Bernoullianum 30, H148

### **Berufliche und fachspezifische Interessen im Schulkontext: Entwicklung, Zusammenhänge, Einflussfaktoren und Auswirkungen**

Chair(s): **Gundula Stoll** (Hector-Institut für Empirische Bildungsforschung, Universität Tübingen)

DiskutantIn(nen): **Gundula Stoll** (Universität Tübingen)

„There is no human competence which can be achieved in the absence of a sustaining interest.“

Dieses frühe Zitat von Tomkins (1962, p.343) spiegelt die Überzeugung vieler Wissenschaftler, dass Lernen und Kompetenzerwerb ohne anhaltendes Interesse nicht möglich ist. Das Zitat verdeutlicht, welche zentrale Funktion Interessen im Allgemeinen für den Erwerb von Wissen- und Fertigkeiten zugeschrieben wird. Als motivationale Faktoren lenken individuelle Interessen Aufmerksamkeitsprozesse, steuern Verhalten und beeinflussen somit wofür Personen ihre Kraft, ihre Fähigkeiten und ihre Zeit einsetzen. Damit sind Interessen für alle Arten von Lernprozessen sowie für bildungs-, ausbildungs- oder berufsbezogene Entscheidungen von zentraler Bedeutung.

In diesem Symposium sollen zwei unterschiedliche Konzeptionen von Interesse miteinander in Bezug gesetzt werden: Fachspezifische Interessen und berufliche Interessen. Fachspezifische Interessen spiegeln das Interesse an Schulfächern wieder und werden daher hauptsächlich im Schulkontext und mit Bezug auf schulische Lehr- und Lernprozesse untersucht. Berufliche Interessen spiegeln Präferenzen für unterschiedliche Arten von Tätigkeiten. Das prominenteste Rahmenmodell für die Klassifikation beruflicher Interessen stellt das RIASEC-Modell von Holland (Holland, 1959, 1997) dar. Berufliche Interessen werden hauptsächlich im Kontext der Berufswahl untersucht, können aber ebenfalls auf den Bildungs- und Ausbildungsbereich angewendet werden (Bergmann, 1992; Volodina, Nagy, & Retelsdorf, 2015).

Die vier Beiträge des Symposiums umspannen die gesamte Schulzeit von der Grundschule bis in die gymnasiale Oberstufe. Dabei stehen in den einzelnen Beiträgen mal berufliche Interessen; mal fachspezifische Interessen im Mittelpunkt. Zusätzlich werden Bezüge zwischen den beiden Interessenskonzeptionen, sowie Bezüge zu anderen zentralen Eigenschaften wie Persönlichkeit, Selbstkonzept und Anstrengungsbereitschaft hergestellt. Ziel dieser erweiterten Perspektive ist die Betrachtung von Gemeinsamkeiten und Unterschieden zwischen fachlichen und beruflichen Interessen.

Der Beitrag von Päßler betrachtet die Entwicklung beruflicher Interessen im Kindesalter. Vorgestellt werden Daten aus einer aktuellen Längsschnittstudie, in der Schweizer Grundschülerinnen und Grundschüler im Alter von 8 bis 13 Jahren zu drei Messzeitpunkten befragt wurden. Der Beitrag von Kneißler und Kollegen betrachtet das Zusammenspiel zwischen beruflichen Interessen und fachspezifischem Interesse für das Fach Mathematik in einer Stichprobe von Haupt- und Realschülern der 5. bis 8. Klassenstufe.

Der Beitrag von Rieger und Kollegen untersucht die Veränderbarkeit von Interesse durch äußere Einflüsse—am Beispiel eines Lehrerwechsels. Die Studie nimmt eine breitere Perspektive ein und vergleicht die Stabilität von Interesse und anderen sozialkognitiven Konstrukten (Selbstkonzept und Anstrengungsbereitschaft) mit der von Big Five Persönlichkeitseigenschaften.

Der Beitrag von Bergmann und Nagy überträgt ein zentrales Konstrukt der Forschung zu beruflichen Interessen—die Interessenkongruenz—in den Schulkontext und untersucht Zusammenhänge zwischen Kongruenz und Schulzufriedenheit in unterschiedlichen Gymnasialzweigen. Durch die zusätzliche Berücksichtigung von Leistungsmaßen und Persönlichkeitseigenschaften leistet diese Studie zusätzlich einen Beitrag zum konzeptionellen Verständnis beruflicher Interessen.

Im Anschluss an die einzelnen Beiträge erfolgt eine gemeinsame Diskussion, in der die unterschiedlichen Perspektiven und Befunde miteinander in Verbindung gebracht werden sollen. Die Diskussion wird vom Chair der Sitzung (Dr. Gundula Stoll) moderiert.

#### *Beiträge des Symposiums*

### **Wie entwickeln sich Interessen bei Kindern? Ergebnisse einer Längsschnittstudie**

**Katja Päßler**

Hochschule für Angewandte Psychologie, Fachhochschule Nordwestschweiz

Berufliche Interessen sind wichtige Determinanten der Ausbildungs- und Studienwahl (Humphreys, Lubinski, & Yao, 1993) sowie bedeutsame Prädiktoren akademischer und beruflicher Leistung (Nye, Su, Rounds, & Drasgow, 2012) und Zufriedenheit (Nye, Su, Rounds, & Drasgow, 2017; Tsabari, Tziner, & Meir, 2005). Obwohl sich individuelle Unterschiede in Aktivitäts- und Tätigkeitspräferenzen bereits sehr früh zeigen, weiß man bisher wenig darüber, wie sich Interessen bei Kindern ausbilden und entwickeln (Hartung, Porfeli, & Vondracek, 2005).

In einschlägigen Metaanalysen zeigen sich erhebliche Geschlechtsdifferenzen in den beruflichen Interessen von Jugendlichen und Erwachsenen: Während Männer die Arbeit mit unbelebten Objekten (z.B. Maschinen, Werkzeugen) bevorzugen, präferieren Frauen den Umgang mit anderen Menschen (Su, Rounds, & Armstrong, 2009). In Bezug auf das Interessenmodell von Holland (1997) sind das praktisch-technische und das intellektuell-forschende Interessen bei Männern und das künstlerisch-sprachliche sowie das soziale Interesse bei Frauen stärker ausgeprägt. Bis-her ist jedoch wenig darüber bekannt, wie sich geschlechtsspezifische Unterschiede in beruflichen Interessen über die Lebensspanne entwickeln und festigen. Studien zeigen, dass sich bereits am Ende der Grundschule Unterschiede in den Interessen von Mädchen und Jungen nachweisen lassen (Maurice & Bäumer, 2015), sie sich beim Übergang in die weiterführende Schule verstärken (Tracey & Ward, 1998). Zudem nimmt das allgemeine Interessenniveau von Mädchen und Jungen beim Übergang von der Grundschule in die weiterführende Schule ab (Tracey, 2002), wobei die Abnahme bei Mädchen stärker ausfällt als bei Jungen (Tracey, Robbins, & Hofsess, 2005). Als ursächlich für diese Interessenabnahme wird u.a. die entwicklungs-bedingte Differenzierung von Interessen betrachtet (Tracey & Sodano, 2013), die auf einer zunehmend realistischeren Wahrnehmung der eigenen Person sowie der zunehmenden Bedeutung sozialer Vergleiche mit anderen Personen beruht (Rohlfing, Nota, Ferrari, Soresi, & Tracey, 2012). Eine Annahme die an einer Stichprobe mit deutschen Grundschulkindern nicht vollumfänglich bestätigt werden konnte (Maurice & Bäumer, 2015).

Im Rahmen einer Längsschnittstudie mit Schweizer Primarschülerinnen und -schülern wurde die Entwicklung kindlicher Interessen über einen Zeitraum von drei Jahren untersucht. Ziel der Studie war es erstens zu untersuchen, ob sich Geschlechtsunterschiede in den Interessen von Mädchen und Jungen tatsächlich schon so früh nachweisen lassen und

zweitens zu prüfen, ob Unterschiede in den Entwicklungsverläufen von Mädchen und Jungen zu einer altersbedingten Zunahme der Geschlechtsunterschiede beitragen. Ferner sollten mit der vorliegenden Studie methodische Schwächen bisheriger Studien überwunden werden, die entweder nur querschnittlich oder längsschnittlich mit zwei Erhebungszeitpunkten die Entwicklung beruflicher Interessen bei Kindern untersucht haben. Dazu wurden im Herbst 2015 erstmals 500 Primarschülerinnen und -schüler im Alter zwischen 8 und 13 Jahren ( $MW = 10.8$ ,  $SD = 1.00$ ) in der Deutschschweiz zu ihren Interessen befragt. Mithilfe von Multilevel Growth Curve Modellen konnte nachgewiesen werden, dass bereits zum ersten Messzeitpunkt in allen sechs Interessendimensionen des RIASEC-Modells (Holland, 1997) signifikante Unterschiede zwischen Mädchen und Jungen bestehen. Allerdings fallen diese Unterschiede geringer aus als bei Jugendlichen und Erwachsenen ( $d = .23 - .87$ ). Zudem berichten anders als bei älteren Stichproben Mädchen ein signifikant höheres Interesse an intellektuell-forschenden Tätigkeiten als Jungen ( $d = .33$ ). Die Analyse altersbedingter Veränderungen zeigt eine signifikante Abnahme des praktisch-technischen, forschenden und sprachlich-künstlerischen Interesses mit zunehmendem Alter, während das soziale, systematisierend-ordnende und unternehmerische Interesse stabil blieb. Geschlechtsunterschiede in der altersbedingten Veränderung ließen sich nur für das sprachlich-künstlerische sowie das forschende Interesse nachweisen. Analysen der Profilhöhe und Differenziertheit zeigen zwar eine altersbedingte Abnahme der Profilhöhe, aber keine signifikante Zunahme der Profildifferenziertheit und bestätigen damit die Ergebnisse von Maurice und Bäumer (2015).

## **Das Zusammenspiel von beruflichen Interessen und mathematischem Fachinteresse von der 5. bis zur 8. Klasse**

**Heide Kneißler, Hana Gaspard, Gundula Stoll, Eike Wille, Benjamin Nagengast, Ulrich Trautwein**  
Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Zwei gut erforschte und etablierte Theorien zur Erklärung von beruflichen Entscheidungen, unter anderem im Bereich der Mathematik, Informatik, Naturwissenschaft und Technik (MINT), sind Eccles' Erwartungs-Wert-Theorie (EWT; Eccles et al., 1983) und Hollands Theorie der beruflichen Interessen (RIASEC-Modell; Holland, 1997). In beiden Theorien stellt das Interessenskonstrukt den zentralen Prädiktor beruflicher Entscheidungen dar. Während in der EWT das Interesse an einem Fach als Teil der Wertüberzeugungen akademische Entscheidungen vorhersagt (Eccles, 2005), sagen in Hollands Modell dispositionale Präferenzen für bestimmte Arten von Tätigkeiten berufliche Entscheidungen vorher (Holland, 1997). Zahlreiche empirische Studien konnten die Vorhersagekraft sowohl der beruflichen Interessen (z.B. Lent, Brown & Hackett, 1994) als auch der Wertüberzeugungen (für einen Überblick, siehe Wigfield, Tonks & Klauda, 2009) für berufliche Aspirationen und Entscheidungen zeigen. Auch im MINT-Bereich finden sich dafür Belege (Humphreys & Yao, 2002; Schoon & Eccles, 2014). Zu den Zusammenhängen zwischen den verschiedenen Interessenskonstrukten ist allerdings wenig bekannt. Da in den Annahmen der EWT und des RIASEC-Modells die Interessensbereiche unterschiedlich spezifisch sind, wurden die beiden Theorien bisher kaum gemeinsam untersucht. Befunde einer ersten Studie, in der berufliche Interessen zusammen mit Fachinteresse in Mathematik an einer Stichprobe von Abiturientinnen und Abiturienten analysiert wurden, zeigen, dass ein mittlerer Zusammenhang zwischen mathematischem Interesse und den Interessensbereichen Realistic und Investigative besteht (Wille, Cambria, Stoll, Nagengast & Trautwein, under review). Auch Holland beschreibt die Nähe dieser Interessensbereiche zum Fach Mathematik (Holland, 1997). Es kann demnach angenommen werden, dass die Interessen Realistic und Investigative positiv mit dem mathematischen Fachinteresse zusammenhängen. Unklar ist jedoch, wie sich dieser Zusammenhang über die Zeit entwickelt. Bezüglich der Entwicklung beider Interessenskonstrukte ist das Jugendalter eine wichtige Phase, in der sich sowohl Fachinteressen (Wigfield et al., 2009) als auch berufliche Interessen (Tracey, Robbins & Hofess, 2005) ausdifferenzieren. Es erscheint somit plausibel, dass sich über die Zeit eine Art mathematisch-naturwissenschaftliches Interessensprofil ausbildet, welches sich durch einen zunehmenden Zusammenhang des mathematischen Interesses mit den Interessen Realistic und Investigative zeigt. Obwohl das RIASEC-Modell und die EWT wie oben dargestellt verwandte Konstrukte beinhalten, wurden sie bisher kaum gemeinsam untersucht. Es stellt sich hier somit die Frage, wie die beiden Konstrukte der beruflichen Interessen und des mathematischen Fachinteresses zusammenhängen und wie sich dieser Zusammenhang über die Zeit entwickelt.

Die vorliegende Studie untersucht zwei Fragestellungen: 1. Wie hängen die beruflichen Interessen mit dem Fachinteresse in Mathematik zusammen? 2. Wie entwickeln sich diese Zusammenhänge während der frühen Jugendzeit?

Dazu werden Daten der Längsschnittstudie TRAIN („Tradition und Innovation: Entwicklungsverläufe an Haupt- und Realschulen in Baden-Württemberg und Mittelschulen in Sachsen“) von 3880 Schülerinnen und Schülern verwendet. Diese wurden zu vier Messzeitpunkten von der fünften bis zur achten Klasse per Fragebogen befragt. Es wurden jeweils die beruflichen Interessen nach dem RIASEC-Modell (Holland, 1994) sowie das Fachinteresse für Mathematik (Wigfield & Eccles, 2000) erhoben. Die beruflichen Interessen wurden mit sechs Items pro Interessensbereich erfasst (Beispielitem für Realistic-Skala: „Wie sehr interessieren dich folgenden Dinge? – mit Maschinen und technischen Geräten arbeiten“; Cronbach's  $\alpha = .78-.89$ ). Das mathematische Fachinteresse wurde mit drei Items erfasst (Beispielitem: „Matheaufgaben machen mir einfach Spaß“; Cronbach's  $\alpha = .73-.75$ ).

Erste deskriptive Analysen zeigen niedrige bis mittlere Zusammenhänge zwischen dem mathematischen Interesse und den beruflichen Interessen. Die höchsten Korrelationen des Mathematikinteresses finden sich erwartungskonform mit den Interessen Realistic, Investigative und Conventional ( $.20 \leq r \leq .40$ ). Diese werden über die Zeit stärker, wohingegen die Zusammenhänge mit den anderen Interessensbereichen Artistic, Social und Enterprising ( $.10 \leq r \leq .20$ ) schwächer werden, was auf eine Ausdifferenzierung hindeutet.

## **The Effects of Getting a New Teacher on the Consistency of Interest, Effort, Self-Concept, and the Big Five**

**Sven Rieger<sup>1</sup>, Richard Göllner<sup>1</sup>, Marion Spengler<sup>1</sup>, Benjamin Nagengast<sup>1</sup>, Ulrich Trautwein<sup>1</sup>, Jeffrey R. Harring<sup>2</sup>, Brent W. Roberts<sup>3</sup>**

<sup>1</sup>Hector-Institut für Empirische Bildungsforschung, Universität Tübingen, <sup>2</sup>University of Maryland, <sup>3</sup>University of Illinois at Urbana-Champaign

### Theoretical Background & Research Questions

Given the impact of teachers on the development of young students (e.g., Wentzel, 2002) and the fact that students get new teachers every year in most school systems, an important question is what happens to students when they get a new teacher? Students do not know much about their new teacher's expectations and attitudes, and consequently, such a transition entails a degree of uncertainty that should putatively cause at least a temporary change in students' thoughts, feelings, or behaviors. In the present investigation, we were particularly interested in the extent to which a variety of personality constructs (e.g.,

interests, the Big Five; BF) are affected by getting a new teacher. The investigation addressed three questions. First, we asked whether the rank-order correlations of personality constructs are affected by getting a new teacher. Second, we investigated whether getting a new teacher or not getting a new teacher is associated with a change in variance over time (i.e., do individual differences become more or less pronounced) in personality constructs? Third, we asked whether social cognitive (SC) variables are more strongly influenced by getting a new teacher than the BF personality traits. For this purpose, we compared students who had the same teacher and students who got a new teacher on several SC variables (i.e., interests, self-concept, effort) and the BF personality traits in a large longitudinal study.

#### Method

**Sample.** We used two time points from a large longitudinal German study ("Tradition and Innovation in Educational Systems"; TRAIN). In sum, we used information for N = 2,458 (54.6% male) students such that n = 1,546 had different teachers and n = 912 had the same teacher at both time points.

**Instruments.** The SC constructs (i.e., self-concept, interest, and effort) were assessed with four items each in three school subjects (math, German, and English). Cronbach's Alphas ranged from .64 to .90. The BF were measured with the German version (Lang, Lüdtke, & Asendorpf, 2001) of the BF Inventory (Alphas ranged from .66 to .80).

**Statistical analyses.** We used multiple-group latent-variable models to estimate the 1-year rank-order correlations in both groups as well as all variance components. To compare the differences in the correlations, we used Fisher's z-tests (Fisher, 1925) and confidence intervals (Zou, 2007). We compared the differences in variances (from T1 to T2) between the groups with the Wald test.

#### Results

We found no differences in the rank-order correlations for the math-related SC variables (e.g., math interest) between students who got a new teacher and students who did not get a new teacher. However, on the SC constructs that are related to German and English (i.e., effort in English and German as well as interest and self-concept in German), students who got a new teacher showed significantly lower rank-order correlations (i.e., their relative placement within the group changed) than students who did not get a new teacher. On the BF personality traits, there were no significant differences in the rank-order correlations between the groups (except extraversion). In sum, our results suggest that SC variables are slightly more influenced by getting a new teacher than the BF variables.

Regarding the variance comparison, we found two (of 14) significant differences in the change in the variances over time between the groups (i.e., interest in English and conscientiousness). Moreover, there were significant increases or decreases in variances over time within the groups (e.g., math interest, effort in math and German). On the basis of these contradictory patterns, we concluded that there were no systematic patterns in the increases or decreases in variances over time within or between groups.

### **Die Bedeutung der Interessenkongruenz für die Schulzufriedenheit in der gymnasialen Oberstufe**

**Jakob Bergmann<sup>1</sup>, Gabriel Nagy<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Christian-Albrechts-Universität Kiel, <sup>2</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Christian-Albrechts-Universität Kiel

Berufliche Interessen gelten als bedeutsame Prädiktoren von Wahlentscheidungen und beruflichen Bewährungsvariablen (Rounds & Su, 2014; Volodina & Nagy, 2016). Das am besten untersuchte Modell beruflicher Interessen ist das RIASEC-Modell von Holland (1997), in dem sechs Interessenorientierungen postuliert werden: Realistic (praktisch-technisch), Investigative (intellektuell-forschend), Artistic (künstlerisch), Social (sozial), Enterprising (unternehmerisch) und Conventional (ordnend-verwaltend).

Eine Kernannahme des Modells ist, dass die Ähnlichkeitsstruktur der Interessendomänen einer Circumplex-Struktur entspricht. Eine Implikation dieses gut validierten Strukturmodells ist, dass individuelle Interessenprofile auf zwei Achsen (Dinge-Personen und Daten-Ideen; Prediger, 1982) übertragen werden können (vgl. Nagy, Marsh, Lüdtke & Trautwein, 2009). Eine weitere zentrale Annahme des RIASEC-Modells ist die Kongruenz-Hypothese, die besagt, dass sich eine Übereinstimmung zwischen Personen- und Umweltorientierungen positiv auf Zufriedenheit, Leistung und Stabilität auswirkt.

Die Kongruenzhypothese fand in vielen Studien, die in beruflichen Kontexten und im Hochschulstudium durchgeführt wurden, empirische Unterstützung (vgl. Nye, Su, Rounds & Drasgow, 2017). Bis heute ist jedoch wenig über die Rolle der Interessenkongruenz in Bildungsetappen, die als Vorläuferphasen der beruflichen Entwicklung gelten, bekannt (siehe aber Eder, 1988; Eder & Reiter, 2002). Zudem fehlt es an Studien in denen konkurrierende Erklärungsvariablen kontrolliert wurden. Berufliche Interessen sind stark mit den Big Five Persönlichkeitsmerkmalen sowie mit der kognitiven und akademischen Leistungsfähigkeit assoziiert (Ackerman & Heggestad, 1997), sodass sich die Frage stellt, inwieweit der Effekt der Interessenkongruenz robust gegenüber der Kontrolle dieser Merkmale ist (z.B. Volodina, Nagy & Köller, 2015).

Der vorliegende Beitrag untersucht die Kongruenzhypothese hinsichtlich der Zufriedenheit im Schulkontext in der Sekundarstufe II. Seit der flächendeckenden Einführung beruflicher Gymnasien bietet die Sekundarstufe II die Möglichkeit berufsbezogener Schwerpunktsetzungen. Die Richtungen der Gymnasialzweige unterscheiden sich hinsichtlich der inhaltlichen Anforderungsprofile und ziehen Schülerinnen und Schüler (SuS) mit unterschiedlichen Interessenprofilen an (Nagy & Husemann, 2010). Dementsprechend ist zu erwarten, dass sich interesseninkongruente Wahlentscheidungen negativ auf die schulische Zufriedenheit auswirken.

Zur Untersuchung der Fragestellung wurde ein Datensatz von N = 6300 SuS aus J = 155 Schulen der 13. Jahrgangsstufe (zusammengefügt aus zwei Kohorten der TOSCA-Studie; Köller, Watermann, Trautwein & Lüdtke, 2004) analysiert. Die SuS entstammten allgemeinbildenden Gymnasien sowie beruflichen Gymnasien der Ausrichtungen sozialpädagogisch, ernährungswissenschaftlich, biotechnologisch, technisch und wirtschaftswissenschaftlich. Die beruflichen Interessen wurden mit dem Allgemeinen Interessen-Struktur-Test (Bergmann & Eder, 1992) erfasst, wobei die Interessenprofile mittels personenseitiger Gewichte der Achsen Dinge-Personen und Daten-Ideen errechnet wurden. Schulzufriedenheit wurde anhand einer Kurzskaala der LAU-Studie (Lehmann, Gänßfuß & Peek, 1999) erhoben. Als weitere Variablen gingen Maße zu den Big Five Persönlichkeitstraiten (Borkenau & Ostendorf, 1991) sowie Testwerte in den Bereichen kognitive Grundfähigkeiten (Heller & Perleth, 2000), Mathematik (Baumert, Bos & Lehmann, 2000) und Englisch (Kurzversion ITP-TOEFL; vgl. Köller & Trautwein, 2004) ein.

Die Auswertung erfolgte mittels mehrbenenanalytischer Regressionen mit Schulzugehörigkeit als Gruppierungsvariable. Es fanden sich statistisch signifikante mittlere Effekte der Testleistungen (höhere Zufriedenheit bei guten Leistungen) und Big Five

Traits (positive Effekte von Belastbarkeit, Verträglichkeit und Gewissenhaftigkeit), wobei diese Effekte nicht zwischen Schulen variierten. Im Fall der Interessenprofile fielen die mittleren Effekte gering aus, aber die Effekte variierten zwischen Schulen. Die Hinzunahme der Gymnasialzweige führte zu signifikanten Cross-Level-Interaktionen mit den Effekten der Interessenprofile. Die entsprechenden Effektmuster waren kompatibel mit der Kongruenzhypothese. Im allgemeinbildenden Gymnasium, das keine spezifische Schwerpunktsetzung ausweist, lag ein Nulleffekt der Interessenprofile vor. Im Fall der beruflichen Gymnasien zeigte sich, dass die Zufriedenheit mit zunehmender Nähe der individuellen Interessenprofile zu den Interessenbereichen, die den jeweiligen inhaltlichen Schwerpunkten entsprachen, anstieg.

Die Befunde stützen somit die Kongruenzhypothese in schulischen Kontexten, wobei sie sich zudem als robust gegenüber alternativen Erklärungsvariablen erwies (vgl., Volodina, Nagy, & Köller, 2015). Berufliche Interessen steuern demnach nicht nur die mittels der inhaltlichen Wahl- und Übergangsentscheidungen eingeschlagenen beruflichen Entwicklungspfade (z.B. Nagy & Husemann, 2010), sondern bedingen auch das individuelle Erleben und Verhalten in schulischen Vorläuferetappen der beruflichen Entwicklung.

## A22

Zeit: Donnerstag, 15.02.2018: 13:15 - 15:00 · Ort: Musikwissenschaftliches Seminar, Hörsaal 006

### Migration und Integration als Herausforderung für Lehrerprofessionalität- Personen und Kontextmerkmale

Chair(s): **Meike Bonefeld** (Universität Mannheim, Deutschland), **Hannah Kleen** (Bergische Universität Wuppertal)

DiskutantIn(nen): **Birgit Becker** (Goethe Universität Frankfurt a.M.)

Im Schulkontext wird professionelles Handeln immer stärker in den Fokus der Aufmerksamkeit gerückt. Eine wichtige Säule stellen hierbei gerechte Beurteilungen der Lehrkräfte dar. Lehrkräfte sind gemäß Schulgesetz dazu verpflichtet, unabhängig von Herkunft zu handeln, damit jedem Schüler und jeder Schülerin die gleichen Chancen zukommen. Forschung konnte allerdings zeigen, dass gerade die soziale und ethnische Herkunft Faktoren für eine Benachteiligung im Schulsystem darstellen. Schülerinnen und Schüler (SuS) mit Migrationshintergrund erzielen z.B. niedrigere Leistungen als SuS ohne Migrationshintergrund und besuchen häufiger eine Hauptschule (Bos et al., 2004; OECD, 2014). Neben tatsächlicher Leistungsunterschiede, z.B. aufgrund unterschiedlicher Sprachfähigkeiten, spielt möglicherweise die Lehrkraft dabei eine Rolle. Lehrkräfte stellen einen wichtigen Interaktionspartner auf dem Bildungsweg von SuS dar: sie entscheiden über Förderung, interagieren alltäglich mit den SuS und beurteilen diese. Im Kontext gerechter Beurteilung sind daher auch Lehrermerkmale bedeutsam. Offene Fragen sind dabei unter anderem, ob sich Lehrkräfte in ihren Urteilen und ihrem Verhalten gegenüber SuS mit versus ohne Migrationshintergrund unterscheiden und unter welchen Bedingungen sie akkurater beurteilen. Beurteilungen sind unter anderem abhängig von den Einstellungen und Erwartungen von Lehrkräften. Hier wurde bereits gezeigt, dass Lehrkräfte gegenüber SuS mit Migrationshintergrund negativer eingestellt sind und geringere Erwartungen an diese haben als an deutsche SuS (Glock, 2016). Diese Einstellungen und Erwartungen und somit auch Beurteilungen und Verhalten der Lehrkräfte können jedoch durch weitere Merkmale beeinflusst werden. In Bezug auf Urteilsgenauigkeit sind sowohl Lehrer- und Schülermerkmale als auch der Kontext, in dem die Lehrer-Schüler-Interaktion stattfindet, entscheidend (Südkamp, Kaiser, Möller, 2012).

In diesem Symposium wird systematisch der Frage nachgegangen, welche verschiedenen Personen- und Kontextmerkmale im Zusammenhang mit SuS mit Migrationshintergrund eine geringere oder stärkere Benachteiligung dieser Schülergruppe zur Folge haben. Bedeutsam ist dabei, dass das Symposium die Situation dieser SuS aus unterschiedlichen erklärenden (Einstellungen und Erwartungen der Lehrkraft) und zu erklärenden Urteilskontexten (Leistung, Sozialer Ausschluss) betrachtet und somit die Komplexität und Vielschichtigkeit der Situation von SuS mit Migrationshintergrund im Schulkontext verdeutlicht. Das Symposium eröffnet die Möglichkeit der Ableitung von Faktoren zur Verringerung der Verzerrungen nach dem Migrationshintergrund und bildet damit einen wichtigen Aspekt ab, der zu einer höheren Lehrerprofessionalität gehört.

Beitrag 1 fokussiert auf die Effekte von Lehrererwartungen auf die Leistung der SuS im Sinne einer selbsterfüllenden Prophezeiung. Dabei zeigt sich, dass Lehrkräfte negativere Erwartungen gegenüber SuS mit Migrationshintergrund ausbilden und sich diese auf die Kompetenzentwicklung in den Bereichen Lesen und Mathematik auswirken. Beitrag 2 konzentriert sich auf die Einstellungen angehender Lehrkräfte und deren unterschiedliche Ausprägung je nach eigener Migrationsherkunft der angehenden Lehrkraft. Es zeigt sich, dass sich Einstellungen gegenüber SuS mit Migrationshintergrund nach dem Migrationshintergrund der Lehrenden unterscheiden. Der 3. Beitrag betrachtet den Einfluss der Integration und der Erfahrung mit Migranten auf die Einstellungen angehender Lehrkräfte. Die Befunde zeigen, dass die Einstellungen positiver sind bei einem integrierten Schüler, aber nur bei Lehrkräften mit viel Erfahrung. Beitrag 4 richtet den Blick auf Rolle des Migrationshintergrundes beim Umgang mit sozialem Ausschluss seitens der Lehrkraft. Hier zeigt sich, dass für angehende Lehrkräfte die Bewertung von sozialer Ausgrenzung nicht vom Migrationshintergrund des Betroffenen beeinflusst wird.

#### Beiträge des Symposiums

### Selbsterfüllende Prophezeiungen im Kontext ethnischer Bildungsungleichheiten

**Georg Lorenz, Sarah Gentrup**  
IQB Humboldt-Universität zu Berlin

Verschiedene Forschungsarbeiten verdeutlichen, dass Lehrkräfte geringere Erwartungen für die schulischen Kompetenzen von Kindern mit Migrationshintergrund haben und dass derartige Unterschiede im Vergleich zu einheimischen Schülerinnen und Schülern zum Teil unabhängig von tatsächlichen Kompetenzen und Fähigkeiten bestehen (z. B. Ready & Wright, 2011). Auch im deutschen Kontext wurde kürzlich gezeigt, dass Lehrkräfte negativ verzerrte Erwartungen für die Leistungen von Kindern aus türkeistämmigen Familien haben (Lorenz, Gentrup, Kristen, Stanat & Kogan, 2016).

Verzerrte Lehrkrafterwartungen können selbsterfüllende Prophezeiungen auslösen. Dabei handelt es sich um einen Prozess, bei dem sich die Interaktionen im Klassenzimmer durch Erwartungen der Lehrenden so verändern, dass sich die Kompetenzen der Schülerinnen und Schüler letztendlich an diese Erwartungen anpassen (Merton 1948; Rosenthal & Jacobson, 1968). In der internationalen Literatur wurden derartige Effekte wiederholt nachgewiesen (Jussim, Robustelli & Cain, 2009). Auch im deutschen Kontext finden sich Hinweise für dieses Phänomen (Gentrup, Lorenz, Kristen & Kogan, in Vorbereitung; Lorenz, im Erscheinen).

In Kombination betrachtet legt das geschilderte Befundmuster implizit nahe, dass Erwartungseffekte mit der Entstehung von Bildungsungleichheiten in Verbindung stehen könnten (Ready & Chu, 2015; van den Bergh, Linda, Denessen, Hornstra, Voeten & Holland, 2010). Die vorliegende Untersuchung soll klären, ob sich diese Annahme mittels weiterführender Analysen bestätigen lässt und ob selbsterfüllende Prophezeiungen im deutschen Schulkontext tatsächlich zur Aufrechterhaltung oder Verstärkung ethnischer Bildungsungleichheiten beitragen.

Die Analysen erfolgen anhand einer Stichprobe von N=1.030 (sprachlicher Bereich) bzw. N=943 (mathematischer Bereich) Schülerinnen und Schülern sowie N=65 (sprachlicher Bereich) bzw. N=60 (mathematischer Bereich) Lehrkräften. Die Lehrkräfte wurden direkt nach Beginn des ersten Grundschuljahres gebeten, ihre Erwartungen an alle Kinder ihrer Klasse anhand mehrerer Items anzugeben (5-stufig). Zum gleichen Zeitpunkt wurden die Mathematik- und Lesefähigkeiten der Kinder (FIPS; Bäuerlein et al., 2012), ihre kognitiven Grundfähigkeiten (CFT1; Weiß, Rudolf H. & Osterland, 1997) sowie ihre Motivation (FEES1-2 adaptiert; Rauer & Schuck, 2004) und ihr Selbstkonzept (Poloczek et al., 2009) erfasst. Die Messungen der Mathematik- und Lesefähigkeiten wurden am Ende des Schuljahres wiederholt. Telefonische Elterninterviews lieferten Informationen über den familiären Hintergrund der Kinder.



Ziel der Analysen war es, den Einfluss solcher Lehrkrafteerwartungen auf die Leistungsentwicklung der Kinder zu bestimmen, die systematisch nach dem ethnischen Hintergrund der Kinder verzerrt waren. Dazu wurden zunächst Regressionen der Lehrkrafteerwartungen auf fachspezifische Leistungen, kognitive Grundfähigkeiten, motivationale Merkmale und das Selbstkonzept der Kinder am Schuljahresanfang sowie ihr Geschlecht und ihren sozioökonomischen Hintergrund berechnet. Die Residuen dieser Regressionen gingen dann als Prädiktoren in Analysen der Leistungsentwicklung der Kinder anhand von Latent-Change Modellen ein. In diesen Modellen wurden das Vorwissen und der familiäre Hintergrund der Kinder kontrolliert. Alle Analysen berücksichtigten die Mehrebenenstruktur der Daten.

Die Ergebnisse unterstützen zunächst die Befunde bisheriger Studien. So zeigten Lehrkräfte häufiger negativ verzerrte Erwartungen für Kinder aus türkischstämmigen Familien als für Kinder ohne Migrationshintergrund. Außerdem wurden Effekte verzerrter Erwartungen auf die Kompetenzentwicklung der Kinder sowohl in Mathematik als auch im Lesen sichtbar. Weiterführende Analysen verdeutlichen darüber hinaus, dass diese Effekte bei Kindern aus türkischen Familien größer ausfielen als bei Kindern ohne Migrationshintergrund oder bei Kindern aus anderen Herkunftsgruppen. Auch wenn sich die beobachtete Leistungsentwicklung nicht signifikant zwischen den Herkunftsgruppen unterschied, deuten erste Ergebnisse außerdem darauf hin, dass sich die Leseleistungen von Kindern aus türkischstämmigen Familien tendenziell an die Kompetenzen von Kindern aus einheimischen Familien angleichen, wenn Nachteile durch negative Erwartungseffekte aus den statistischen Modellen herausgerechnet werden.

Insgesamt legen die Ergebnisse nahe, dass Erwartungseffekte mit einer Verfestigung der Bildungsungleichheiten zwischen Kindern mit einem türkischen Migrationshintergrund und Kindern ohne Migrationshintergrund einhergehen. Die berichteten Effekte weisen insgesamt eine geringe Effektstärke auf. Die Erklärungskraft des untersuchten Phänomens für das Zustandekommen ethnischer Leistungsdisparitäten ist demnach insgesamt als vergleichsweise gering einzustufen.

## **2. Die Rolle des eigenen Migrationshintergrundes von Lehramtsstudierenden für die Einstellungen gegenüber Schülern mit Migrationshintergrund**

**Sabine Glock<sup>1</sup>, Meike Bonefeld<sup>2</sup>, Hannah Kleen<sup>1</sup>, Oliver Dickhäuser<sup>2</sup>**  
<sup>1</sup>Bergische Universität Wuppertal, <sup>2</sup>Universität Mannheim

Schüler mit Migrationshintergrund gehören zu den Verlierern im deutschen Bildungssystem. Die Forschung zeigt, dass Lehrkräfte diese Schüler nicht nur ungünstiger beurteilen als Schüler ohne Migrationshintergrund (Bonefeld, Dickhäuser, Janke, Praetorius, & Dresel, 2017; Glock, 2016), sondern auch negativere Einstellungen gegenüber dieser Schülergruppe haben (Glock & Klapproth, 2017). Dies betrifft vor allem die impliziten Einstellungen als automatische Bewertungen, während die expliziten Einstellungen als bewusste Bewertungen oftmals positiv sind. Forschung hat gezeigt, dass gerade die impliziten Einstellungen von Lehrkräften eine Rolle für ihre Urteile und ihr unterrichtliches Handeln spielen (Kumar, Karabenick, & Burgoon, 2015; Peterson, Rubie-Davies, Osborne, & Sibley, 2016), so dass negative implizite Einstellungen von Lehrkräften gegenüber Schülern mit Migrationshintergrund zur Benachteiligung dieser Schülergruppe beitragen können. Dabei wird der kulturelle Unterschied zwischen den Lehrpersonen und den Schülern mit Migrationshintergrund als Hauptursache für negativen Einstellungen und benachteiligende Urteile gesehen (Monroe, 2005). So ist die Annahme, dass Lehrpersonen positivere Einstellungen zeigen und auch positivere Urteile gegenüber Schülern treffen, mit denen sie den sozialen Hintergrund teilen. Dies wurde bisher in Bezug auf den sozio-ökonomischen Hintergrund (Dunkake & Schuchart, 2015) und im Hinblick auf explizite Einstellungen gegenüber Schülern mit Migrationshintergrund gezeigt (Hachfeld, Schroeder, Anders, Hahn, & Kunter, 2012). Was jedoch noch nicht untersucht wurde, ist, ob die impliziten Einstellungen ebenfalls positiver ausfallen, wenn Lehrpersonen einen eigenen Migrationshintergrund aufweisen. Diese Annahme wurde in einer Studie untersucht.

116 Lehramtsstudierende nahmen an dieser Untersuchung teil. 52 der Probanden wiesen einen eigenen Migrationshintergrund auf. Zur Erfassung der impliziten Einstellungen gegenüber Schülern mit Migrationshintergrund wurde ein Implicit Association Test (IAT; Greenwald, McGhee & Schwartz, 1998) durchgeführt. Die Aufgabe der Probanden bestand darin, sowohl Namen von Schülern mit bzw. ohne Migrationshintergrund als auch positive bzw. negative Adjektive per Tastendruck in Kategorien einzuordnen. Erfasst wurde die Reaktionszeit. Nach dem IAT, wurden die expliziten Einstellungen mit Hilfe eines Fragebogens erhoben. Hier wurden die Probanden gebeten das Arbeitsverhalten der Schüler, deren sprachliche, naturwissenschaftlich-mathematische Begabung und Intelligenz in Bezug auf den Migrationshintergrund einzuschätzen. Der Studie lag ein Zwischensubjektgedesign mit dem Faktor Migrationshintergrund der Lehramtsstudierenden zugrunde.

Die Ergebnisse des IAT zeigten einen signifikanten Unterschied in den impliziten Einstellungen: Während deutsche Lehramtsstudierende eher negative Einstellungen aufwiesen, konnten die Einstellungen der Lehramtsstudierenden mit Migrationshintergrund als neutral eingeordnet werden. Bei den expliziten Aussagen konnte lediglich ein Unterschied auf den Dimensionen Intelligenz und naturwissenschaftlich-mathematischen Begabung ermittelt werden.

Dies ist die erste Studie, die zeigt, dass die impliziten Einstellungen gegenüber Schülern mit Migrationshintergrund von Lehrpersonen davon abhängig sind, ob die Lehrkraft einen eigenen Migrationshintergrund hat. Durch die eigene Zugehörigkeit zu der Gruppe der Menschen mit Migrationshintergrund in Deutschland können über die Bevorzugung der Eigengruppe die im Vergleich zu Lehrpersonen ohne Migrationshintergrund positiveren impliziten Einstellungen erklärt werden.

Interessanterweise wurden die impliziten Einstellungen eher neutral als positiv gewertet. Heraus könnte geschlossen werden, dass diese impliziten Einstellungen auch eine Rolle für die Beurteilung spielen könnte. Anzunehmen wäre hier, dass Lehrkräfte mit Migrationshintergrund weniger einer Urteilsverzerrung aufgrund des Migrationshintergrunds unterliegen. Die immer größere Anzahl von Studierenden mit Migrationshintergrund, die sich für den Beruf des Lehrers entscheiden, gestaltet die zukünftigen Kollegien kulturell heterogen. Lehrkräfte mit eigenem Migrationshintergrund können nicht nur zu einer fairen Beurteilung von Schülern mit Migrationshintergrund beitragen, sondern dienen auch als Vorbild einer gelungenen Integration.

## **Erfahrung und Integration: Einflussfaktoren auf Einstellungen gegenüber türkischen Schülern**

**Hannah Kleen, Sabine Glock**  
Bergische Universität Wuppertal

In dieser Studie wird untersucht, ob die Erfahrung von Lehrkräften mit türkischen Personen einen Einfluss auf die Einstellungen dieser Lehrkräfte gegenüber türkischen Schülerinnen und Schülern (SuS) hat. Dabei wird zusätzlich der Integrationsstatus von türkischen SuS variiert und als ein einflussnehmendes Merkmal betrachtet. Einstellungen als die positive oder negative Bewertung einer Schülergruppe, können einen Grund für die schulische Benachteiligung von Migranten darstellen. Die Forschung zeigt, dass die impliziten Einstellungen von Lehrkräften gegenüber SuS mit Migrationshintergrund negativ und mit Schülerleistungen und Lehrerverhalten korreliert sind, während die expliziten Einstellungen zumeist positiv sind (van den Bergh, 2010). Nach der Kontakthypothese von Allport (1954) können persönliche Erfahrungen mit Migranten allerdings zu

veränderten Einstellungen führen. Personen, die wenig persönliche Erfahrung mit Migranten haben, nehmen diese Personengruppe als homogener wahr. Diese Outgrouphomogenität verhindert, dass SuS, die der Outgroup angehören, individuell wahrgenommen werden. Personen mit viel Erfahrung mit der Outgroup nehmen diese jedoch heterogener wahr und somit auch individuelle Unterschiede innerhalb der Outgroup. Der Integrationsstatus von türkischen SuS kann dabei ein individuelles Merkmal dieser SuS darstellen. In bisheriger Forschung wurde allerdings nicht auf den Integrationsstatus von SuS mit Migrationshintergrund und die Erfahrung von Lehrkräften als einflussnehmende Merkmale auf Einstellungen eingegangen. Dabei ist die Annahme, dass die Einstellungen von Lehrkräften gegenüber integrierten türkischen SuS positiver sind als bei nicht-integrierten türkischen SuS, wenn die Lehrkräfte viel Erfahrung mit türkischen Personen haben. Bei Lehrkräften mit weniger Erfahrung sollte sich kein Unterschied nach Integrationsstatus zeigen.

Um diese Annahme zu überprüfen, wurden insgesamt 197 Lehramtsstudierende befragt. Es wurde ein Szenario entwickelt und validiert, welches entweder einen integrierten oder einen nicht-integrierten türkischen Schüler beschreibt. Einer Hälfte der Probanden wurde das integrierte Szenario gezeigt, der anderen Hälfte das nicht-integrierte. Im Anschluss daran wurden die impliziten Einstellungen der Versuchspersonen mit einem Impliziten Assoziationstest sowie die expliziten Einstellungen erhoben. Der Fragebogen zu den expliziten Einstellungen (Hachfeld et al., 2012) bestand aus vier Dimensionen: Enthusiasmus, Selbstwirksamkeit und Lernvorurteile gegenüber dem Unterrichten von SuS mit Migrationshintergrund sowie multikulturelle Überzeugungen. Zusätzlich zu demografischen Variablen wurde die persönliche Erfahrung mit Personen mit Migrationshintergrund erfragt, indem unter anderem der prozentuale Anteil an Migranten in der Schulzeit der Probanden als auch in ihren ersten praktischen Schulerfahrungen als Lehramtsstudierende erfragt wurde. Erste Analysen zeigen, dass die impliziten Einstellungen von Lehramtsstudierenden mit viel Erfahrung bei integrierten SuS positiver sind als bei nicht-integrierten. Dieser Unterschied zeigt sich bei den expliziten Einstellungen in der Selbstwirksamkeitserwartung der Lehramtsstudierenden. Diejenigen mit viel Erfahrung berichten eine höhere Selbstwirksamkeitserwartung beim integrierten SuS als beim nicht-integrierten. Bei den Probanden mit wenig Erfahrung zeigte sich kein Unterschied nach Integrationsstatus.

Je mehr persönliche Erfahrung mit türkischen Migranten Lehramtsstudierende haben, desto positiver ist die Einstellung gegenüber türkischen Schülern mit hohem Integrationsstatus. Die persönliche Erfahrung, vielleicht auch die Freundschaft zu türkischen Personen, kann Vorurteile gegenüber Migranten abbauen. Lehramtsstudierende mit hoher persönlicher Erfahrung mit Migranten scheinen eine differenziertere Wahrnehmung zu haben, welche sich positiv auf ihr späteres unterrichtliches Handeln auswirken kann. In weiterer Forschung sollte noch tiefergehend darauf eingegangen werden, welche Lehrer- und Schülermerkmale Einstellungen von Lehrkräften gegenüber SuS mit Migrationshintergrund beeinflussen.

## Zur Rolle von Herkunft und Kontext für die Beurteilung von sozialer Ausgrenzung

Hanna Beißert<sup>1</sup>, Meike Bonefeld<sup>2</sup>

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, <sup>2</sup>Universität Mannheim

Soziale Ausgrenzung ist ein zentrales Thema im Kindesalter, denn Heranwachsende beobachten und erleben Ausgrenzung oder Zurückweisung in ihrem alltäglichen Leben und Forschung konnte vielfältige negative Folgen sozialer Ausgrenzung dokumentieren (Killen & Rutland, 2011). Forschung zeigt, dass soziale Ausgrenzung häufig auf Gruppenzugehörigkeiten basiert (Killen, Mulvey & Hitti, 2013). Lehrkräfte haben die Möglichkeit in soziale Interaktionen von Schülern einzugreifen und möglicherweise Einfluss auf deren zukünftiges Ausgrenzungsverhalten nehmen zu können.

Nun stellt sich die Frage, ob die Herkunft von Kindern, die von Ausgrenzung betroffen sind, in diesem Zusammenhang eine relevante soziale Kategorie für Lehrkräfte darstellt. Wird die Wahrscheinlichkeit, dass eine Lehrperson in eine Ausgrenzungssituation unter Kindern eingreift, unterschiedlich eingeschätzt, wenn es sich um ein deutsches versus ein türkisches Kind handelt? Wie verhält es sich mit der Bewertung des Exklusionsverhaltens? Hat die Herkunft der ausgegrenzten Schüler einen Einfluss darauf?

Drei mögliche Positionen sind denkbar: Zum einen ist es möglich, dass aufgrund von Vorurteilen oder (impliziten) Einstellungen gegenüber Kindern mit Migrationshintergrund die Ausgrenzung eines türkischen Kindes als weniger negativ eingestuft und ein Eingreifen als weniger nötig angesehen wird. Im Gegensatz dazu könnte zum anderen angenommen werden, dass es die Teilnehmenden als besonders wichtig erachten, dass ein Kind einer Fremdgruppe integriert und nicht ausgeschlossen wird und dementsprechend den Ausschluss als negativer und die Wahrscheinlichkeit zur Intervention als höher einstufen. Und schließlich wäre es auch denkbar, dass die Herkunft des ausgeschlossenen Kindes für die Beurteilungen der Lehrkräfte keine relevante soziale Kategorie darstellt.

An weiteren Stichproben und mit weiteren Vignetten soll nun überprüft werden, ob sich die hier gefundene Unabhängigkeit der Beurteilungen sozialer Exklusionssituationen von der Herkunft der ausgeschlossenen Kinder replizieren lässt. Um zu überprüfen, für welche der drei Positionen sich empirische Evidenz finden lässt, wurden in einer Online-Studie 99 Lehramtsstudierende (67 weiblich, M=21.21 Jahre, SD=3.85) anhand von Vignetten befragt. Den Teilnehmenden wurden kurze Geschichten dargeboten, in denen sich jeweils eine Gruppe von Kindern am Ende einer Unterrichtsstunde zu einer Aktivität verabredet und ein weiteres Kind davon ausschließt. Hierbei wurde variiert, ob das Kind deutsch (Max) oder türkisch (Murat) war und ob die Aktivität freizeitbezogen (Fußball spielen) oder schulbezogen (Lerngruppe) war.

Die Teilnehmenden wurden gebeten, anzugeben wie fair/unfair, gerechtfertigt/ungerechtfertigt, in Ordnung/nicht in Ordnung das Verhalten der Schüler in den Vignetten war. Weiterhin gaben sie an, wie wahrscheinlich es ist, dass sie in eine solche Situation eingreifen würden. In einer offenen Frage wurden die Begründungen hierzu erfasst.

Vorläufige Analysen liefern erste Evidenz für die dritte oben genannte Position, denn es konnten keinerlei Unterschiede zwischen den Geschichten mit Kindern unterschiedlicher Herkunft aufgedeckt werden. Univariate Anovas zeigten weder für die Beurteilungsmaße noch für die Wahrscheinlichkeit der Intervention Unterschiede zwischen Max und Murat, weder im freizeitbezogenen noch im schulbezogenen Kontext.

Darüber hinaus zeigten die offenen Begründungen der Teilnehmenden, dass die überwiegende Mehrheit eingreifen würde, um mit den Schülern ins Gespräch zu kommen und die Gründe für den Ausschluss zu hinterfragen und anschließend gemeinsam mit den Kindern eine Lösung zu finden. Das deutet darauf hin, dass die angehenden Lehrkräfte nicht unreflektiert auf Inklusion der betroffenen Kinder bestehen, sondern auch Situationen in Betracht ziehen, in denen ein solcher Ausschluss legitim sein könnte, was mit der Wahrnehmung von Kindern in solchen Situationen in Einklang steht (Hitti, Mulvey & Killen, 2011). Interessanterweise zeigten sich auch hier keine Unterschiede in Abhängigkeit der Herkunft des Kindes.

Diese Befunde können im Sinne der dritten Position interpretiert werden und darauf hindeuten, dass die Herkunft der ausgeschlossenen Kinder für die Beurteilungen nicht relevant ist. Vor dem Hintergrund früherer Forschung, die herkunftsbezogene Urteilsverzerrungen bei der Leistungsbeurteilung aufzeigen konnte (Bonefeld & Dickhäuser, 2017), ist es

ein erfreulicher Befund, dass die Herkunft von Kindern für die Bewertung ihrer sozialen Interaktionen aus Sicht von Lehramtsstudierenden keine Rolle zu spielen scheint.

## B01

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S103

### Sozialverhalten und Schulleistungen: Erklärt das Arbeitsverhalten den Zusammenhang?

**Julia Becherer<sup>1</sup>, Olaf Köller<sup>2</sup>, Friederike Zimmermann<sup>1</sup>**

<sup>1</sup>Institut für Pädagogisch-Psychologische Lehr- und Lernforschung (IPL), CAU Kiel; <sup>2</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), CAU Kiel; [jbecherer@ipl.uni-kiel.de](mailto:jbecherer@ipl.uni-kiel.de)

#### Theoretischer Hintergrund und Fragestellung

Obwohl bisherige Forschung darauf verweist, dass dissoziales Verhalten negativ mit Schulleistungen assoziiert ist (Hinshaw, 1992) und prosoziales Verhalten hingegen positiv mit Schulleistungen assoziiert ist (Caprara, Barbaranelli, Pastorelli, Bandura & Zimbardo, 2000; Wentzel, 1993), sind die Wirkmechanismen hinter den Beziehungen bisher wenig erforscht. In der Literatur werden verschiedene Erklärungsansätze für einen Einfluss des Sozialverhaltens auf Leistungen diskutiert (wie z.B. lernförderliche oder -abträgliche Beziehungen zu den Mitschülerinnen und Mitschülern oder Lehrkräften, belohnende oder bestrafende Notenvergabe; Loveland, Lounsbury, Welsh & Buboltz, 2007; Wentzel, 1993), bisher fehlt jedoch zumeist die empirische Fundierung. In dieser Untersuchung wird das Arbeitsverhalten als weiterer möglicher Mediator näher untersucht. Eine Studie von Metsäpelto et al. (2015) weist bereits darauf hin, dass die Beziehung zwischen dissozialem Verhalten und Testleistungen über arbeitsvermeidendes Verhalten vermittelt wird. Erweiternd hierzu soll nun untersucht werden, ob auch prosoziales Verhalten über das Arbeitsverhalten vermittelt mit den Schulleistungen zusammenhängt sowie ob die Zusammenhangsmuster unterschiedlich für Noten und Testleistungen ausfallen. Es wird angenommen, dass prosoziales Verhalten mit günstigem, dissoziales Verhalten hingegen mit ungünstigem Arbeitsverhalten im Schulunterricht assoziiert ist (z.B. im Sinne eines situationsübergreifenden regel- und erwartungskonformen bzw. -verletzenden Verhaltens oder über Wirkungen des Sozialverhaltens auf die Peerbeziehungen, welche wiederum die Arbeitsmotivation beeinflussen könnten) und hierüber vermittelt auch zu schulischen Leistungen in Form von Noten und standardisierten Testleistungen beiträgt.

#### Methode

Die vorliegende Studie wurde mit einer Kohorte von Neuntklässlern verschiedener Schulformen aus dem größeren Projekt EIKA (2006) durchgeführt (N = 1024).

Die Schulnoten in Mathematik und Deutsch wurden von den jeweiligen Fachlehrkräften am Ende des Schuljahres vergeben und umcodiert, sodass höhere Werte besseren Leistungen entsprechen. Die Testleistungen in Mathematik und Deutsch wurden mittels standardisierter Tests erhoben, deren Items für internationale Schulleistungsstudien entwickelt wurden (z.B. Lehmann, Peek, Gänsfuß & Husfeld, 2002).

Das Sozialverhalten wurde zu Beginn des Schuljahres soziometrisch erfasst, wobei die Schülerinnen und Schüler jeweils alle Mitschülerinnen und Mitschüler der Klasse einschätzten. Der Indikator für prosoziales Verhalten bezog sich auf helfendes Verhalten, dissoziales Verhalten wurde mittels aggressiven Verhaltens eingeschätzt (EIKA, 2006).

Die Arbeitshaltung der Schülerinnen und Schüler wurde durch die Lehrkraft mit den meisten Schulstunden in der Klasse eingeschätzt. Die Skala wurde innerhalb des Projektes entwickelt und umfasste vier Items, z.B. der Schüler „arbeitet intensiv im Unterricht mit“ oder „erledigt Aufträge zügig“. Die Reliabilität der Skala nach Cronbachs Alpha war sehr gut (.92).

Es wurde je ein Strukturgleichungsmodell mit Noten und eines mit Testleistungen als abhängiger Variable mit dem Programm Mplus (Muthén & Muthén, 2014) berechnet, wobei für die jeweilige Vorleistung kontrolliert wurde. Sowohl die Leistungsmaße als auch das Arbeitsverhalten wurden latent erfasst.

#### Ergebnisse

Beide Modelle wiesen einen guten Modelfit auf (Noten/Testleistung):  $\chi^2(25) = 102.6/108.4$ , CFI = .98/.98, TLI = .96/.97, RMSEA = .06/.06, SRMR = .03/.03.

Es zeigten sich deutliche positive Effekte des prosozialen Verhaltens auf das Arbeitsverhalten, sowie negative Effekte des dissozialen Verhaltens auf das Arbeitsverhalten. Das Arbeitsverhalten war wiederum prädiktiv für die Schulleistungen, insbesondere für Schulnoten. Die indirekten Pfade, von den Sozialverhaltensweisen über das Arbeitsverhalten vermittelt auf die Leistungen, waren ebenfalls in beiden Modellen signifikant. Die direkten Pfade des Sozialverhaltens auf Schulleistungen waren überwiegend nicht zusätzlich prädiktiv.

Diese Ergebnisse zeigten sich, obwohl für den starken Einfluss der Vorleistung kontrolliert wurde. Des Weiteren erwiesen sich die Ergebnismuster auch unter Kontrolle des Geschlechts und sozioökonomischen Status (erfasst mit dem HISEI; Ganzeboom & Treiman, 1996) als robust. Die Ergebnismuster konnten an einer weiteren Stichprobe von Neuntklässlern (N = 1360) repliziert werden.

#### Einordnung

Bisher ist wenig über vermittelnde Mechanismen der Beziehungen zwischen Sozialverhalten und Schulleistungen bekannt. Das Ergebnis dieser Studie, dass dissoziales und prosoziales Verhalten über das Arbeitsverhalten vermittelt prädiktive Relevanz für Schulleistungen aufweist, stellt daher einen wichtigen Schritt zu tieferem Verständnis der Beziehungen dar.

### Warum melden sich Schüler? Ergebnisse einer Video-Studie in zwei Unterrichtsfächern.

**Ricardo Böheim, Maximilian Knogler, Tina Seidel**

Technische Universität München, Deutschland; [ricardo.boeheim@tum.de](mailto:ricardo.boeheim@tum.de)

#### Theoretischer Hintergrund

Schülerinnen und Schüler (SuS), die sich im Unterricht aktiv beteiligen wollen, müssen sich in aller Regel melden. Diese Tatsache spiegelt sich in den Analysen von Crombie, Pyke, and Silverthorn (2003) wider, wonach Lernende ihre aktive Unterrichtsbeteiligung hauptsächlich in Abhängigkeit ihres Meldeverhalten einschätzen. In der Literatur zu erfolgreichen Unterrichtsinteraktionen wird das Meldeverhalten jedoch kaum berücksichtigt (siehe z.B. Mercer & Dawes, 2014; Scott, Mortimer, & Aguiar, 2006). Frühere Studien konnten zeigen, dass die Häufigkeit der Meldungen zwischen einzelnen SuS stark variiert (Sacher, 1995). Die bisherige Forschung liefert jedoch kaum Erklärungen für diese Variation.

Der vorliegende Beitrag widmet sich der Analyse der Varianz dieses wichtigen Schülerverhaltens. Basierend auf Modellen zur Erklärung von Schülerbeteiligung (engl. engagement), werden motivationale Faktoren wie intrinsische Motivation (Ryan & Deci,

2000), Interesse (Ainley, 2012) und Selbstkonzept (Marsh & Martin, 2011) zur Vorhersage des Meldeverhaltens herangezogen. Insgesamt wird angenommen, dass motivationale Faktoren einen substantiellen Beitrag zur Vorhersage des Meldeverhaltens leisten. Da motivationale Konstrukte weitestgehend fachspezifisch organisiert sind (Bong, 2001; Pekrun, 2006), untersuchen wir die Zusammenhänge in zwei unterschiedlichen Fachkontexten (Deutsch und Mathematik).

Fragestellungen

1. Inwiefern variiert das Meldeverhalten zwischen einzelnen SuS, Klassen und Fächern?
2. Inwiefern leisten motivationale Faktoren einen Beitrag zur Erklärung dieser Varianz?
3. Inwiefern unterscheidet sich die Varianzaufklärung in den beiden Schulfächern Mathematik und Deutsch?

Methode

An der Untersuchung nahmen 503 SuS (59% Mädchen) aus 20 gymnasialen Klassen der achten Jahrgangsstufe teil. Zur Erfassung fachspezifischer Unterschiede wurden die Klassen jeweils in den Fächern Mathematik und Deutsch untersucht.

Um die Meldehäufigkeit pro SuS zu erfassen, wurden die Klassen in beiden Fächern in jeweils einer Schulstunde videographiert. Die Schülermeldungen wurden von zwei unabhängigen Ratern mittels der Software Interact codiert (Interrater-Reliabilität  $\kappa = .68$ ; Übereinstimmung zwischen Beurteilern: 72.7%). Im Anschluss an die videographierte Stunde wurden die SuS zu ihrer situationalen Motivation befragt. Die intrinsische Motivation wurde in Anlehnung an Mitchell (1993) in eine Catch- (3 Items,  $\alpha_{\text{Deu}} = .87$ ,  $\alpha_{\text{Ma}} = .86$ ) und Hold-Facette (6 Items,  $\alpha_{\text{Deu}} = .82$ ,  $\alpha_{\text{Ma}} = .81$ ) differenziert (Prenzel, Duit, Euler, Lehrke, & Seidel, 2001). Zusätzlich machten die SuS Angaben zu ihrer extrinsischen Motivation in der videographierten Stunde (3 Items,  $\alpha_{\text{Deu}} = .82$ ,  $\alpha_{\text{Ma}} = .78$ , Prenzel et al., 2001).

Die stabilen motivationalen Faktoren, das Fachinteresse (3/4 Items,  $\alpha_{\text{Deu}} = .82$ ;  $\alpha_{\text{Ma}} = .88$ ) und das fachspezifische Selbstkonzept (6/5 Items,  $\alpha_{\text{Deu}} = .82$ ;  $\alpha_{\text{Ma}} = .92$ ), wurden zu Beginn des Schuljahres erhoben (Deutsch: Hertel, Hochweber, Mildner, Steinert, & Jude, 2014; Mathe: OECD, 2014). Für die Auswertung der Daten wurde eine Regressionsanalyse in R gerechnet, wobei die Mehrebenen-Struktur der Daten berücksichtigt wurde (package: lme4).

Ergebnisse und Diskussion

Die deskriptiven Befunde zeigen, dass die Meldehäufigkeit sowohl zwischen den einzelnen SuS als auch zwischen den 20 Klassen stark variiert. In Mathematik melden sich 15% der SuS überhaupt nicht, während sich 4% 20-mal und öfter melden. Im Deutschunterricht ergibt sich ein vergleichbares Bild (18% der SuS ohne Meldungen, 4% der SuS mit mindestens 20 Meldungen).

Die untersuchten motivationalen Lernervoraussetzungen erklären 30% (Ma) bzw. 23% (Deu) der Varianz in der Meldehäufigkeit. Während im Matheunterricht das Selbstkonzept ( $\beta = 0.23$ ,  $p < .001$ ) einen positiven Einfluss auf das Meldeverhalten zeigt, kann dieser Effekt für den Deutschunterricht nicht bestätigt werden. Die extrinsische Motivation hingegen zeigt in beiden Fächern einen negativen Einfluss auf das Meldeverhalten ( $\beta_{\text{Ma}} = -0.24$ ,  $p < .001$ ;  $\beta_{\text{Deu}} = -0.27$ ,  $p < .001$ ). Das Fachinteresse sowie die intrinsische Motivation tragen hingegen nicht signifikant zur Varianzaufklärung der Meldehäufigkeit bei.

Im Vortrag werden diese Ergebnisse darüber hinaus durch Analysen von  $N = 14$  Schülerinterviews (achte Jahrgangsstufe) ergänzt. Diese Zusatzanalysen sollen darüber Aufschluss geben, ob neben den gefundenen Einflussfaktoren weitere Prädiktoren für das Meldeverhalten von SuS gefunden werden können.

## **Wer profitiert von der Arbeit mit Zeitzeugen im Geschichtsunterricht? Interaktionsanalysen zum Zusammenhang von Persönlichkeitsmerkmalen und Leistungs-Outcomes in einer Interventionsstudie**

**Christiane Bertram<sup>1</sup>, Wolfgang Wagner<sup>2</sup>, Lisa Henke<sup>2</sup>, Ulrich Trautwein<sup>2</sup>**

<sup>1</sup>Universität Konstanz, Deutschland; <sup>2</sup>Universität Tübingen, Deutschland; [christiane.bertram@uni-konstanz.de](mailto:christiane.bertram@uni-konstanz.de)

Die Befragung von Live-Zeitzeugen im Geschichtsunterricht wird in den letzten Jahren hinsichtlich der Chancen wie auch Risiken diskutiert und beforscht (u.a. Angvik & Borries, 1997; Henke-Bockschatz, 2014; Whitman, 2011). Eine jüngst veröffentlichte Interventionsstudie (Bertram, Wagner & Trautwein, 2017) fand bezüglich zentraler Outcomes unterschiedliche Effekte der Arbeit mit Live-Zeitzeugen im Vergleich zur Arbeit mit Zeitzeugen-Videos bzw. Transkripten. Die Schülerinnen und Schüler in der Live-Bedingung schnitten in zwei von drei Kompetenztests, die ihre Einsicht in die epistemologischen Prinzipien des Fach Geschichte erfassen, statistisch signifikant schlechter ab, waren aber deutlich überzeugter von ihren inhaltlichen und methodischen Lernfortschritten wie auch von dem motivationalen Potenzial der Zeitzeugenmethode als die Lernenden in der Video- und Textbedingung.

Im nächsten Schritt stellt sich die Frage, ob – im Sinne einer Aptitude-Treatment-Interaktion – die persönlichen Voraussetzungen der Schülerinnen den Einfluss des Treatments moderieren. Hinsichtlich der Live-Bedingung wäre es denkbar, dass sich die Schülerinnen und Schüler mit einem höheren Selbstkonzept eher aktiv an einer Befragung im Klassenraum beteiligen und mit dem Zeitzeugen direkt interagieren. Daher könnte es sein, dass sie eventuell mehr von der Live-Bedingung profitieren als die Lernenden mit einem niedrigeren Selbstkonzept. Auf der anderen Seite könnte es sein, dass die Lernenden mit einem niedrigeren Selbstkonzept eher von der Arbeit mit Zeitzeugenvideos und transkripten profitieren, da bei dieser Arbeit keine direkte Kommunikation und Interaktion mit einer fremden Person erforderlich ist und der Unterricht wie gewohnt abläuft.

Die Re-Analyse des Datensatzes basierte auf einer cluster-randomisierten Stichprobe von 769 Gymnasialschülerinnen und schülern in drei Interventionsbedingungen (Live/Video/Text). Die Lernenden in den Interventionsklassen hatten im Vortest ihr generelles Selbstkonzept, ihr soziales Selbstkonzept (Jonkmann, Trautwein, & Lüdtke, 2009; Leung, Marsh, Yeung, & Abduljabbar, 2015) wie auch ihre soziale Selbstwirksamkeit (Jerusalem & Klein-Hessling, 2009) eingeschätzt. Die Interaktionsanalysen wurden in einem Regressionsmodell auf Basis manifester Variablen (UVs: Selbstwahrnehmung / Gruppenzugehörigkeit; AVs: Einsicht in epistemologische Prinzipien) anhand von Pfadmodellen unter Berücksichtigung von Hilfsvariablen (auxiliary variables zur Optimierung der Schätzung bei fehlenden Werten) berechnet (Mplus 8.0, Type = Complex, full information maximum likelihood-Schätzung; Muthén & Muthén, 1998-2017). Die Interaktionseffekte der Interventionsbedingungen (Dummy-Kodierung: Live vs. Video/Text) und der Selbstwahrnehmung der Schülerinnen und Schüler wurden im Hinblick auf ihre Leistung in den drei Kompetenztests im Nachtest (NT) und Follow Up-Test (FT) berechnet. Die Signifikanz der Interaktionsterme wie auch der Treatmentunterschiede bei hoher (+1 SD) bzw. niedriger (-1 SD) Ausprägungen des Selbstkonzepts wurde in insgesamt 18 Modellen überprüft (für jede AV und UV separat: 1 x (Treatmentunterschied) x 6 AVs x 3 UVs). In die Regressionsanalysen wurden die Schülerleistungen im Vortest wie auch ihre kognitiven Grundfähigkeiten (KFT) einbezogen.

In 6 von 9 Tests im Nachtest und in 7 von 9 Tests im Follow Up-Test haben wir statistisch signifikante Interaktionsterme gefunden. Für Lernende mit einem niedrigen (-1 SD) Selbstkonzept zeigte sich, dass in 4 von 9 Modellen im Nachtest und in 4 von 9 Modellen im Follow-Up Test statistisch signifikant höhere Kompetenzen in der Video-/Textbedingung als in der Live-Bedingung zu erwarten sind. Für Lernenden mit einem höheren Selbstkonzept (+1 SD) haben wir im Nachtest keine statistisch signifikanten Unterschiede bezüglich der erwarteten Kompetenzen in den jeweiligen Interventionsbedingungen gefunden, allerdings drei im Follow Up-Test. Die gefunden positiven Interaktionsterme bestätigen somit nicht die Vermutung, wonach Lernende mit einem hohen Selbstkonzept besonders von der Live-Bedingung profitieren, sondern es gibt Hinweise darauf, dass die Lernenden mit einem niedrigeren Selbstkonzept mehr von der Video- und Textbedingung profitieren.

Unsere Studie bietet Hinweise auf eine Aptitude-Treatment-Interaktion in dem Sinne, dass die Effekte verschiedener Unterrichtsangebote mit Zeitzeugen (Live-Befragung versus Arbeit mit Zeitzeugen-„Konserven“) mit der Persönlichkeit der Lernenden, in diesem Fall ihrem Selbstkonzept, zusammenhängen.

## **Welche Schüler\*innen beteiligen sich an kontroversen Diskussionen im politischen Fachunterricht? Und wie wirkt sich die Einnahme einer zum persönlichen Urteil der Schüler\*innen konträren Position (pro/contra) auf die mündliche Beteiligung und die Änderung (versus Beibehaltung) des persönlichen Urteils aus?**

**Dorothee Gronostay, Sabine Manzel**

Universität Duisburg-Essen, Deutschland; [dorothee.gronostay@uni-due.de](mailto:dorothee.gronostay@uni-due.de)

Theoretischer Hintergrund: Mündliches Argumentieren stellt im Sinne des „learning-to-argue“-Ansatzes eine Facette politischer Handlungskompetenz dar. Zugleich dient es im Sinne des „arguing-to-learn“-Ansatzes auch der Förderung anderer Kompetenzdimensionen, wie zum Beispiel der politischen Urteilsfähigkeit sowie demokratischer Einstellungen (Detjen et al., 2012). Drittens ist die Teilnahme an kontroversen Unterrichtsdiskussionen ein Bestandteil der mündlichen Beteiligung am Unterricht. Daher stellt sich angesichts heterogener Lernvoraussetzungen die Frage, welche Schüler\*innen besonders zur Teilnahme an kontroversen politischen Unterrichtsdiskussionen prädisponiert sind. Der Fokus wird hier auf ein Diskussionsformat gelegt, bei dem den Schüler\*innen zufällig eine Position (pro/contra) zur Einnahme in der Diskussion zugewiesen wird. Dieses Format impliziert, dass ein Teil der Schüler\*innen eine Position vorbereitet und ggf. auch in der Diskussion vertritt, die mit dem eigenen Urteil übereinstimmt, während der andere Teil eine mit dem eigenen Urteil unvereinbare Position einnimmt (konträre Position).

Politikwissenschaftliche Studien zeigen, dass extravertierte, emotional stabilere Menschen sowie solche mit einer größeren Offenheit für Erfahrungen eher zur Diskussion politischer Themen neigen, während besonders verträgliche Menschen diese tendenziell meiden (z.B. Gerber, Huber, Doherty, Dowling, 2012). Die Argumentierfreude als Prädisposition zu argumentativem Verhalten ergibt sich dabei aus einer Annäherungskomponente (Wahrnehmung des Argumentierens als intellektuell bereichernde Tätigkeit, die mit positiven Emotionen verbunden ist) und einer Vermeidungskomponente (geringes Vertrauen in die eigene Argumentationskompetenz, verbunden mit negativen Emotionen, Infante & Rancer, 1982). Aus Theorien des kritischen Denkens lässt sich zudem eine Relevanz epistemologischer Überzeugungen für die Bereitschaft zum Argumentieren ableiten (Kuhn, Cheney und Weinstock, 2000). Bislang ist kaum erforscht, inwiefern diese Personenmerkmale einen Beitrag zur Vorhersage von Diskussionsverhaltens im Unterricht leisten.

Ziele: Die Studie zielt darauf, die angenommenen Wirkzusammenhänge zwischen den fünf Faktoren der Persönlichkeit und den epistemologischen Überzeugungen auf die Argumentierfreude zu prüfen sowie die Prädiktionskraft dieser Konstrukte auf die Beteiligung von Schüler\*innen an einer kontroversen Unterrichtsdiskussion zu untersuchen. Zusätzlich wird die Wirkung der Einnahme einer konträren Position auf die mündliche Beteiligung an einer kontroversen Diskussion und auf die Wahrscheinlichkeit der (Nicht-)Änderung des persönlichen Urteils in den Blick genommen.

Methode: Die vorliegende Studie ist eingebettet in die Videostudie „Argumentative Lehr-Lernprozesse“ (Gronostay, 2016), die in zehn Schulklassen der Jahrgangsstufen 8/9 (Gymnasium) in Nordrhein-Westfalen durchgeführt wurde. Die Beteiligung an der Fishbowl-Diskussion wird über die mittlere Anzahl der Wörter der Redebeiträge je Schüler\*in operationalisiert. Zudem wurden epistemologische Überzeugungen“ (Krettenauer, 2005), die fünf Faktoren der Persönlichkeit (Rammstedt & John, 2005) und die Argumentierfreude (in Anlehnung an Infante & Rancer, 1982) erfasst. Das persönliche Urteil der Schüler\*innen zur Streitfrage wurde zu Beginn und erneut am Ende der Unterrichtslektion erhoben. Die Auswertung erfolgte quantitativ mittels hierarchischer Regressions- und Pfadanalysen.

Ergebnisse: Es konnten jeweils ca. 40% der Varianz der Annäherungs- sowie der Vermeidungskomponente der Argumentierfreude aufgeklärt werden. Die Befunde stehen im Einklang mit einer Studie von Nussbaum und Bendixen (2003) und komplementieren diese u.a. insofern, als dass zusätzlich die tatsächliche Beteiligung an einer argumentativen Diskussion in die Analysen miteinbezogen wurde. Pfadanalysen zeigen, dass Effekte der Persönlichkeitsmerkmale sowie auch der epistemologischen Überzeugungen auf die Diskussionsbeteiligung fast vollständig über die Argumentierfreude vermittelt werden. Dabei stellt sich die Vermeidungskomponente als besserer Prädiktor für die Diskussionsbeteiligung heraus als die Annäherungskomponente. Dies kann durch den „Bühnencharakter“ kompetitiver Diskussionen mit Rollenvorgabe erklärt werden. Die Einnahme einer konträren Position wirkte sich negativ auf die Diskussionsbeteiligung aus. Offen bleibt die Frage, ob dieser Effekt auf höhere kognitive Anforderungen oder/und eine niedrigere Teilnahmemotivation der Schüler\*innen zurückzuführen ist. Zudem zeigten Schüler\*innen mit konträrer zugewiesener Position deutlich häufiger eine Änderung ihres persönlichen Urteils (40%) als solche, denen eine mit ihrem persönlichen Urteil übereinstimmende Position zugeteilt wurde (4%),  $\chi^2(1, N = 156) = 34.03, p < .001, Cramers V = .47$ . Dieses Ergebnis bestätigt Befunde von Lilly (2012) und verdeutlicht die Wirkmacht von Rollen-/Positionseinnahmen auf die Urteilsbildung.

## B02

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S104

### Interesse an Mathematik: Welche Facetten beeinflussen den Studienerfolg in der mathematischen Studieneingangsphase?

**Stefanie Rach<sup>1</sup>, Stefan Ufer<sup>2</sup>, Timo Kosiol<sup>2</sup>**

<sup>1</sup>Universität Paderborn, Deutschland; <sup>2</sup>LMU München, Deutschland; [rach@math.uni-paderborn.de](mailto:rach@math.uni-paderborn.de)

#### Theoretischer Hintergrund

Über 30% der Studierenden brechen ihr Mathematikstudium schon in der Studieneingangsphase ab (Dieter, 2012). Individuelles, fachbezogenes Interesse im Sinne einer Person-Objekt-Beziehung (Krapp, 2002) wird als wichtiger Prädiktor von Studienerfolg (Schiefele, Streblo & Brinkmann, 2007) angenommen, da es tiefergehende Lernhandlungen fördert (Hidi & Renninger 2006). Ein Einfluss von Interesse auf Studienerfolg zeigte sich für Klausurergebnisse in empirischen Studien zur Studieneingangsphase bisher jedoch nicht (Rach & Heinze, 2016).

Diese Diskrepanz zwischen Theorie und Empirie liegt möglicherweise in der Konzeption der verwendeten Instrumente begründet, die individuelles Interesse unspezifisch messen. Am Übergang von der Schule in die Hochschule verändert sich der Charakter von Mathematik: Während in der Schule Anwendungen mit einem Fokus auf Berechnungen im Vordergrund stehen, sind Beweise und formale Begriffsrepräsentationen für die Hochschulmathematik charakteristisch (Engelbrecht, 2010). Wenn Studierende allgemein zu ihrem Interesse an Mathematik befragt werden, ist nicht klar, auf welchen Charakter sie ihre Antwort beziehen. Möglicherweise berichten sie ein hohes Interesse an Berechnungs- und Anwendungsaufgaben, die jedoch nicht zentrale Bestandteile mathematischer Lernprozesse im ersten Semester sind. Es liegt nahe, dass insbesondere Interesse an der Hochschulmathematik wichtig für erfolgreiche Lernprozesse in der Studieneingangsphase ist.

#### Fragestellung

Welche Interessensfacetten (Interesse an Schulmathematik bzw. Hochschulmathematik) sagen den Studienzufriedenheit und Klausurergebnisse im ersten Semester eines Mathematikstudiums vorher?

#### Methode

Es wurden Skalen zum Interesse an Mathematik im Allgemeinen (6 Items), an typischen Aktivitäten der Schulmathematik wie Berechnen (6 Items, „Mir ist es wichtig, mathematische Rechenregeln richtig anwenden zu können.“) oder Anwenden (6 Items, „Ich finde es interessant, reale Anwendungsprobleme mit Hilfe von Mathematik zu lösen.“) sowie zum Interesse an typischen Aktivitäten der Hochschulmathematik (8 Items, „Mathematische Beweise zu lesen, macht mir Spaß.“) entwickelt (vgl. Ufer, Rach & Kosiol, 2017). Alle Skalen beziehen die emotions- und wertbezogene sowie intrinsische Komponente von Interesse ein (Krapp, 2002). Die Studierenden schätzten alle Items auf einer vierstufigen Likert-Skala von 0 (trifft nicht zu) bis 3 (trifft zu) ein.

Zu Studienbeginn berichteten Mathematikstudierende (128 Fachbachelor, 74 Lehramt; 95 weiblich) ihre Interessensfacetten. Als kognitive Eingangsvoraussetzungen wurden das mathematische Vorwissen (8 Items) und die Abiturnote erhoben (Rach & Heinze, 2016). Nach sechs Wochen, berichteten die Studierenden ihre Studienzufriedenheit (7 Items, N=202,  $\alpha=.83$ ). Am Ende des Semesters nahmen die beiden Studienganggruppen an unterschiedlichen Klausuren zur Analysis teil (Fachbachelor bzw. Lehramt: jeweils 6 Aufgaben, N=69 bzw. N=51,  $\alpha=.67-.85$ ).

#### Ergebnisse

Die entwickelten Skalen zeigen eine akzeptable bis gute interne Konsistenz ( $\alpha=.70-.87$ ). Für die Klausurergebnisse im ersten Semester zeigten Regressionsmodelle unter Einschluss aller Eingangsvoraussetzungen nur Zusammenhänge mit kognitiven Lernvoraussetzungen (Fachbachelor: N = 122,  $R^2=.48$ ; Lehramt: N=73,  $R^2=.29$ ), jedoch keine zusätzliche Erklärungskraft der Interessensfacetten ( $|\beta|<.15$ ,  $p>.26$ ). Im Gegensatz dazu erklärt neben dem allgemeinen Interesse an Mathematik ( $\beta=.21$ ,  $p<.05$ ) das Interesse an zentralen Tätigkeiten der Hochschulmathematik ( $\beta=.30$ ,  $p<.001$ ) zusätzlich Varianz an der Studienzufriedenheit ( $R^2=.23$ ,  $|\beta|<.10$ ,  $p>.12$  für die anderen Eingangsvoraussetzungen). Studienzufriedenheit und Studienerfolg sind kaum korreliert (Fachbachelor:  $r=.08$ ; Lehramt:  $r=.07$ ).

#### Diskussion

Es konnte kein Zusammenhang zwischen Interesse zu Studienbeginn und dem objektiven Studienerfolg belegt werden, auch nicht für spezifisches Interesse an der Hochschulmathematik. Weitere Forschung sollte aufzeigen, an welcher Stelle die erwartete Wirkungskette von individuellem Interesse über situationales Interesse und Lernaktivitäten hin zum Lernerfolg im ersten Semester nicht trägt. Allerdings könnten auch Interaktionen mit anderen Lernvoraussetzungen, z.B. extrinsischer Motivation, die erwarteten Effekt verdecken. Erwartungskonform geht Interesse an Tätigkeiten, die zentral für die Hochschulmathematik sind, mit höherer Studienzufriedenheit in der Semestermitte einher. Dies zeigt, dass die spezifischen Interessensmaße bestimmte Facetten des Studienerfolgs über allgemeines Interesse hinaus vorhersagen können. Die Studienergebnisse zeigen insbesondere auf, dass zukünftige Studierende auf den veränderten Charakter von Mathematik hingewiesen werden sollten. Unsere Ergebnisse weisen darauf hin, dass die individuellen Interessensprofile der Studierenden die Wahrnehmung des Studienangebots beeinflussen und somit von Lehrpersonen berücksichtigt werden sollten.

### Effekte ökonomischer Kompetenzen auf die Studienfachwahl

**Michael Jüttler, Stephan Schumann**

Universität Konstanz, Deutschland; [Michael.Juettler@uni-konstanz.de](mailto:Michael.Juettler@uni-konstanz.de)

#### Theoretischer Hintergrund

Die Vergabe von Hochschulberechtigungen an Lernende ist in Ländern wie der Schweiz oder Deutschland mit der Annahme verknüpft, dass zum Zeitpunkt der Abschlüsse eine lückenlose allgemeine oder domänenbezogene Studierfähigkeit vorliegt (Eberle, 2013; Huber, 1997; Trautwein & Lüdtke, 2004). Dass diese für einen Teil der Absolvent/-innen nicht gegeben ist, lässt sich aus den zwar tendenziell rückläufigen, jedoch weiterhin hohen Studienabbruchzahlen, aber auch aus Untersuchungen wie z.B. der Studie EVAMAR II für die Schweiz (Eberle et al., 2008) ableiten. Dabei spielen bei der Studienfachwahl neben Merkmalen wie Interessen, sozioökonomischer Hintergrund, etc. besonders auch die individuellen domänenspezifischen Fähigkeiten eine bedeutsame Rolle (z.B. Nagy, 2007; Oepke & Eberle, 2014). Für die in der Studie im Fokus stehenden ökonomischen Kompetenzen liegen zumindest für den deutschsprachigen Raum keine Befunde zur Zusammenhangskette „Ökonomische Kompetenzen - Studienaspiration - Studienfachwahl - Studienerfolg“ vor. Dies überrascht angesichts der

Tatsache, dass wirtschaftswissenschaftliche Studiengänge zu den meistgewählten Studienfächern gehören (Bundesamt für Statistik, 2017; Statistisches Bundesamt, 2017) und ökonomische Kompetenzen eine substantielle Bedeutung hinsichtlich einer sozialen Teilhabe in modernen Wirtschaftsgesellschaften zugeschrieben werden kann (Dubs, 2001, 2011; Eberle & Schumann, 2011; Schumann & Jüttler, 2015; Jüttler et al., 2016). Den ökonomischen Kompetenzen wird in der Studie in Anlehnung an Weinert (2001) und Beck (1989) ein umfassendes Begriffsverständnis zugrunde gelegt, wobei sowohl kognitive (ökonomisches Wissen und Können) als auch nicht-kognitive Kompetenzfacetten (intrinsische Motivation, Interesse, Einstellung und Werthaltung) betrachtet werden.

#### Fragestellung

Vor diesem Hintergrund wird der Fragestellung nachgegangen, welche Vorhersagekraft die am Ende der Gymnasialzeit verfügbaren individuellen ökonomischen Kompetenzen auf die Studienfachwahl unter Kontrolle individueller und kontextueller Merkmale haben.

#### Methode

Als Datengrundlage der Untersuchung dient die längsschnittlich angelegte Studie OEK-Transition („Ökonomische Kompetenzen und Studienerfolg“). Die Stichprobe umfasst  $N_2 = 520$  Lernende aus Gymnasien und Berufsmaturitätsschulen der Deutschschweiz. Der erste Messzeitpunkt lag am Ende des letzten Schulbesuchsjahres ( $N_1 = 2.328$ ; Schuljahr 2010/11; vgl. Schumann & Eberle, 2014). Die Folgeuntersuchung wurde ca. fünf Jahre später im Frühjahr/Sommer 2016 durchgeführt. Das Instrumentarium des ersten Messzeitpunktes setzt sich aus Leistungstests zu den Domänen „Wirtschaft“, „Mathematik“ und „Deutsch“, einem kognitiven Fähigkeitstest und einem Fragebogen (u.a. zur Erfassung der Studienaspiration) zusammen. Die EAP/PV-Reliabilitäten der Tests liegen zwischen 0.75 und 0.83. Die Daten der Tests wurden auf Basis der Item-Response-Theorie mittels des Programms „ConQuest“ (Wu et al. 2007) ausgewertet. In der Folgeuntersuchung wurden mit Hilfe eines Fragebogens und eines Computer-Assisted-Telephone-Interviewing (CATI) die Bildungsverlaufsdaten der letzten fünf Jahre, Prädiktoren und Indikatoren zum Studienerfolg (z.B. Maturanoten, Studienzufriedenheit (Westermann et al. (1996)) sowie weitere relevante Individual- und Kontextfaktoren (z.B. Persönlichkeitsmerkmale (Rammstedt et al. (2004)) erfasst. Auf Grundlage einer Mehrebenenimputation (R Paket jomo:jomo2com) und logistischer Regressionen wurden die Daten in einem weiteren Schritt gewichtet, um den Ausfall zu t2 (Follow-Up) zu kompensieren (Brick & Montaquila, 2009; OECD, 2014). Fehlende Werte in den Leistungsvariablen wurden mittels Mehrebenenimputation ergänzt (R Paket mice::mice.2l.pan (Van Buuren & Groothuis-Oudshoorn, 2011).

#### Ergebnisse

Die Analysen zeigen deutliche Unterschiede zwischen Studierenden mit unterschiedlichem Studienprofil und den am Ende der Gymnasialzeit verfügbaren Kompetenzen. Dabei zeigt sich, dass Lernende mit höheren ökonomischen Kompetenzen im Anschluss an die (Berufs-)Maturität häufiger ein wirtschaftswissenschaftliches als ein anderes Studium wählen (ökonomisches Wissen und Können:  $d=0.78$ ; Motivation:  $d=0.30$ ; Interesse:  $d=0.34$ ; Werthaltung:  $d=0.37$ ; Einstellung:  $d=.59$ ). Bzgl. der Mathematik- und Deutschleistung zeigen sich umgekehrte Effekte (Mathematikleistung:  $d=0.42$ ; Deutschleistung:  $d=0.36$ ). Für die kognitiven Grundfähigkeiten finden sich keine signifikanten Unterschiede. Differenziert nach den einzelnen Fächergruppen lassen sich die obigen Ergebnisse im Rahmen einer multinomialen multivariaten Regression weitestgehend bestätigen. Insgesamt wird durch die Ergebnisse die Bedeutsamkeit einer ökonomischen (Schul-)Bildung für den weiteren Bildungsverlauf unterstrichen. Dabei erweisen sich die während der Schulzeit erworbenen ökonomischen Kompetenzen (insb. ökonomisches Wissen und Können und die Einstellung ggü. wirtschaftlichen Problemstellungen) als zentraler Prädiktor für die Wahl ein wirtschaftswissenschaftliches Studium zu beginnen.

### Logfile Analyse individueller Unterschiede im Lernverhalten und Lernleistung über den Verlauf eines Semesters

**Maria Theobald, Henrik Bellhäuser, Margarete Imhof**

Universität Mainz, Deutschland; [maria.theobald@uni-mainz.de](mailto:maria.theobald@uni-mainz.de)

Digitale Lehrangebote an Universitäten, wie z.B. Online-Vorlesungen, ermöglichen einen flexiblen Zugriff auf Kursinhalte unabhängig von zeitlichen oder örtlichen Beschränkungen. Diese Flexibilität setzt jedoch eigenverantwortliche Planung, Steuerung und Überwachung des Lernprozesses auf Seiten der Lernenden voraus. Verteiltes Lernen und das Überprüfen des eigenen Wissensstandes stellen hoch effektive Lernstrategie dar (Dunlosky, Rawson, Marsh, Nathan, Willingham, 2013). Bisherige Studien konnten bereits zeigen, dass man den Lernerfolg in Online-Kursen anhand des Verhaltens auf der Lernplattform vorhersagen kann (Imhof & Spaeth Hilbert, 2013; Morris, Finnegan & Wu, 2005). Allerdings wurde der Effekt zeitlich verteilten Lernens bislang noch nicht untersucht. Außerdem mangelt es an Studien, die sich mit dem Einfluss individueller Unterschiede auf das Lernverhalten und die Leistung in Online-Vorlesungen befassen (Morris et al., 2005). Daraus ergeben sich folgende Forschungsfragen:

- (1) Können die Studienleistungen anhand objektiven Lernverhaltens vorhergesagt werden?
- (2) Welche Rolle spielen Unterschiede in der Abiturnote und der Gewissenhaftigkeit für das Lernverhalten und schließlich die Leistung?

Wir testen die Effektivität verteilten Lernens und der Bearbeitung von Selbsttests hinsichtlich des Klausurerfolgs in einem ökologisch validen, benoteten Online-Kurs. Ziel ist es herauszufinden, welches Lernverhalten mit Klausurerfolg einhergeht, um Lernende künftig besser zu unterstützen. Darüber hinaus untersuchen wir, inwiefern die Abiturnote und die Gewissenhaftigkeit über das Lernverhalten hinaus Varianz in den Klausurleistungen aufklären (Richardson, Abraham & Bond, 2012). Dabei nehmen wir an, dass der Effekt der Gewissenhaftigkeit auf die Note über das Lernverhalten vermittelt wird. Dieser Wirkzusammenhang wurde bislang noch nicht erforscht.

Die Studie wurde im Rahmen einer bildungswissenschaftlichen Vorlesung für Lehramtsstudierende durchgeführt ( $N = 424$ ; 253 weiblich). Es handelte sich um eine reine Online-Vorlesung, das heißt sämtliche Lernmaterialien (Vorlesungsfolien, Podcasts und Selbsttests) wurden ausschließlich auf der Lernplattform Ilias dargeboten. Das Lernverhalten (Minuten, Zugriffe, Selbsttests) wurde über den Verlauf eines Semesters (12 Wochen) mithilfe wöchentlicher Logfiles dokumentiert. Die Anzahl aktiver Lernwochen, in denen ein Studierender die Lernplattform genutzt hat, wurde als Indikator für verteiltes Lernen angesehen. Studierende, die über mehrere Wochen verteilt auf die Lerninhalte zugegriffen haben, erzielten höhere Werte auf der Variablen als Studierende, die beispielsweise nur einmalig in der Woche vor der Klausur aktiv waren. Am Ende des Semesters wurde die Anzahl bearbeiteter Selbsttests (maximal 16) sowie die Note in der Abschlussklausur erfasst. Etwa die



Hälfte der Studierenden (n = 197) bearbeitete zu Beginn des Semesters zusätzlich einen Fragebogen, in dem die Gewissenhaftigkeit (BFI-2, Danner et al., 2016) und die Abiturnote erhoben wurden.

Es zeigte sich eine starke Heterogenität in den Lernverläufen zwischen Studierenden. Über den Verlauf des Semesters waren die Studierenden durchschnittlich 344 Minuten online (SD = 351.54, [0; 3255]) und haben 122 Mal auf die Lernplattform zugegriffen (SD = 82.20, [0; 483]). Einige Studierende waren in allen Wochen aktiv, während andere die Lernplattform gar nicht besucht haben (M = 5.59, SD = 2.93). Auch die Selbsttests wurden in unterschiedlichem Maß genutzt (M = 12.74, SD = 5.57, [0; 16]).

Die Ergebnisse einer multiplen, hierarchischen Regressionsanalyse ergeben, dass sich sowohl das zeitlich verteilte Lernen als auch die Anzahl bearbeiteter Selbsttests in besseren Klausurleistungen widerspiegelt ( $F(2, 421) = 73.03, p < .001, R^2 = .26$ ). Nimmt man die Abiturnote und Gewissenhaftigkeit in die Regression auf, stellt nur die Abiturnote einen signifikanten Prädiktor neben dem Lernverhalten dar ( $\beta = .44, p < .001$ ; Gesamt  $R^2 = .39$ ). Eine Mediationsanalyse mittels Bootstrapping zeigt jedoch, dass die Gewissenhaftigkeit die Note indirekt über das Lernverhalten beeinflusst ( $\beta = -.14, 95\% \text{ CI } [-.22, -.08]$ ). Eine höhere Gewissenhaftigkeit ist mit verteiltem Lernen über das Semester assoziiert, was wiederum mit besseren Klausurnoten zusammenhängt.

Die Ergebnisse der Studie zeigen, dass verteiltes Lernen und das Überprüfen des eigenen Verständnisses mithilfe von Selbsttests im Kontext von Online-Vorlesungen geeignete Lernstrategien darstellen. Außerdem trägt die Studie zu einem besseren Verständnis der Wirkmechanismen bei, die dem Zusammenhang zwischen Gewissenhaftigkeit und Studienerfolg zu Grunde liegen.

## **Welche Rolle spielen familiäre Verpflichtungen und die Nützlichkeit von gegenwärtigen Lernaufgaben für zukunftsbezogene Karriereziele von kollektivistisch-orientierten Studierenden?**

**Sog Yee Mok<sup>1</sup>, Cafer Bakaç<sup>1</sup>, Tina Seidel<sup>1</sup>, Laura Froehlich<sup>2</sup>**

<sup>1</sup>Technische Universität München, Deutschland; <sup>2</sup>FernUniversität in Hagen, Deutschland; [sogyee.mok@tum.de](mailto:sogyee.mok@tum.de)

### Theoretischer Hintergrund

Die Familie kann auf motivationale Ziele von Lernenden Einfluss nehmen. Familiäre Verpflichtungen sind nach Fuligni und Kollegen (1999) durch Werte und Verhaltensweisen definiert, die durch Unterstützung und Respekt von Kindern gegenüber ihren Eltern gekennzeichnet sind. Bei kollektivistisch-orientierten Familien mit Migrationshintergrund sind diese familiären Verpflichtungen besonders stark ausgeprägt (Fuligni, Tseng, & Lam, 1999). Im Leistungskontext können sich solche familiären Verpflichtungen in Form von hohen elterlichen Erwartungen an die Leistung und Leistungsmotivation der Kinder äußern (Phalet & Claeys, 1993). Diese Erwartungen können auch spätere Karriereziele beeinflussen (Kao & Tienda, 1995).

Der „Future time perspective“ Theorie (De Volder & Lens, 1982; Nuttin & Lens, 1985) zufolge vergrößert sich die „Future time perspective“ (FTP) durch die Formulierung von Zielen in der entfernten Zukunft, was die Motivation zur Zielerreichung fördert. Ein wichtiger angenehmer kognitiver Prozess der FTP ist die Vermittlung der Nützlichkeit eine gegenwärtige Aufgabe zu bearbeiten, um ein Zukunftsziel zu erreichen. So können Lernende die Folgen ihrer Handlung sowohl in der Gegenwart als auch in der Zukunft kognitiv antizipieren (Phalet, Andriessen, & Lens, 2004). Studien zeigen hierzu positive Zusammenhänge zwischen den als nützlich wahrgenommenen, in der fernen Zukunft liegenden (zukunftsbezogenen) Zielen und der Motivation Aufgaben zu bearbeiten, wobei sich letzteres wiederum positiv auf das Lernen bei kollektivistisch-orientierten SchülerInnen mit Migrationshintergrund auswirkt (Andriessen, Phalet, & Lens, 2006). Diese Zusammenhänge sind bislang bei Studierenden wenig erforscht.

Basierend auf der FTP Theorie und den bisherigen Befunden kann für kollektivistisch-orientierte Studierende angenommen werden, dass sie zukunftsbezogene Ziele, die mit den Leistungserwartungen der Eltern konform sind, als nützlich für sich und die Familie einschätzen. Diese zukunftsbezogenen Ziele können sich positiv auf die gegenwärtige Aufgabennützlichkeit und zugleich indirekt auf die Aufgabenmotivation auswirken. Durch die familiären Verpflichtungen sollte besonders für kollektivistisch-orientierte Studierende wichtig sein, zukunftsbezogene Ziele zu erreichen.

Daher wird für kollektivistisch-orientierte Studierende angenommen, dass familiäre Verpflichtungen die Verfolgung eines zukunftsbezogenen Ziels (wie eine mögliche Notenverbesserung zur Erhöhung von Karrierechancen), das auch konform mit den elterlichen Leistungserwartungen ist, indirekt durch die wahrgenommene Nützlichkeit einer Lernaufgabe positiv beeinflussen.

### Fragestellung

Wird der Zusammenhang zwischen familiären Verpflichtungen und dem zukunftsbezogenen, leistungserwartungsorientierten Ziel (Notenverbesserung zur Erhöhung der Karrierechancen) durch die Nützlichkeit einer momentanen Lernaufgabe bei kollektivistisch-orientierten Studierenden mediiert?

### Methode

Stichprobe umfasste 85 internationale Studierende aus verschiedenen kollektivistisch-orientierten Ländern, die an einer Online-Befragung teilnahmen. Die Zustimmung zu zwei zukunftsbezogenen Zielen (Notenverbesserung zur Erhöhung der Karrierechancen versus keine Notenverbesserung) wurde am Ende des Fragebogens mittels einer Vignette erfragt. In der Vignette wurden zwei Studierende am Ende ihres Studiums beschrieben, die die zweitbeste Note in einem Kurs haben und die Möglichkeit angeboten bekommen, zusätzliche Übungen zu bearbeiten, was ihre Note beeinflussen kann. In der Vignette entscheidet sich ein Studierender keine zusätzlichen Übungen zu bearbeiten, da er seine Note ausreichend für eine gute Karriere findet. Ein anderer Studierender entscheidet sich für die Bearbeitung der Übungsaufgaben, um seine Karrierechancen zu erhöhen. Diese Wahl ist konform mit den Leistungserwartungen der Familie. Die Studierenden sollten sich im Anschluss vorstellen, sie wären in der beschriebenen Situation und sollten angeben, wie hoch die Wahrscheinlichkeit ist, dass sie die zukunftsbezogenen Ziele der beiden Studierenden jeweils wählen würden. Als abhängige Variable wurde die Wahrscheinlichkeit das zukunftsbezogene, leistungserwartungskonforme Ziel (Notenverbesserung zur Erhöhung der Karrierechancen) zu wählen untersucht. Die Mediationsanalyse wurde nach dem Bootstrapping-Verfahren nach Hayes (2009) berechnet.

### Ergebnisse

Die Ergebnisse der Mediationsanalyse zeigen einen signifikanten positiven indirekten Effekt ( $a*b = .119, SE = .073, 95\% \text{ CI } [.003; .289]$ ) der familiären Verpflichtungen auf das zukunftsbezogene, leistungserwartungsorientierte Ziel via der wahrgenommenen Aufgabennützlichkeit: Das heißt die Aufgabennützlichkeit mediiert den Zusammenhang zwischen familiären

Verpflichtungen und dem zukunftsbezogenen, leistungserwartungsorientierten Ziel der Notenverbesserung zur Erhöhung der Karrierechancen bei kollektivistisch-orientierten Studierenden.

## B03

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S105

### Professional decision-making by means of a developmental assessment

**Miriam Balt<sup>1</sup>, Antje Ehlert<sup>1</sup>, Annemarie Fritz<sup>2</sup>**

<sup>1</sup>Universität Potsdam, Deutschland; <sup>2</sup>Universität Duisburg Essen; [mbalt@uni-potsdam.de](mailto:mbalt@uni-potsdam.de)

#### Theoretical background

Assessment is a central activity in the work of teachers, performed through regular formal and informal monitoring of student performance via testing and observation. Pedagogical assessment, however, should reach beyond student marks and progress towards curriculum-driven goals. The application of a developmental approach has shown strong evidence (Woods & Griffin, 2013) that teachers can use individual learning progressions to better understand the abilities of their students, as well as to target their teaching to their students' needs. This study describes the construction of a formative assessment tool which aims to measure the development of numeracy in first-grade students in Germany. The tool consisted of items which were based on the development model of numerical concepts (Fritz, Ehlert, & Balzer, 2013), and drawn from an item pool developed during previous pilot studies.

#### Research questions

This study addressed two main research questions. First, to what extent do the items fulfil the quality criteria of the Rasch model? Second, to what extent is it possible to track the students' numeracy development by means of the tool?

#### Methods

A longitudinal study with  $n = 135$  first grade students (54.1% female) from six classes of two German primary schools was conducted. The student's numerical skills were assessed four weeks after the beginning of school using the MARKO-D test (pre-test) (Ricken, Fritz, & Balzer, 2013) and re-assessed at the end of the school year using the MARKO-D1 test (post-test) (Fritz, Ehlert, Ricken, & Balzer, 2017). The trial formative tool was applied over nine measurement points between the two summative assessments.

The trial formative assessment tool consisted of short tests (15 items each) in five different versions. The test versions were targeted to the students' numerical skills, i.e., each test was adapted to the students' performance on the previous test. The instructions were read aloud by trained Masters students from a local university whilst the first grade students attempted the given assessment tasks in their own test books.

The students' performance data was Rasch scaled using a simple dichotomous Rasch model. An item-fit analysis was conducted to evaluate the item quality. Person ability values were analysed through an analysis of variance (ANOVA) for repeated measures. These values were used to describe the progression of student learning in the whole group, with sub-groups compared against each other for differences in performance.

#### Results

The item-fit analysis showed acceptable to very good Infit values between .65 and 1.25 with an overall item reliability of .87. A significant increase in performance over time could be reported ( $F(8,80) = 34.4, p < .001$ ) with significant contrasts between the measurement points on the learning progression. There was no interaction effect between measurement time and gender. Different learning progressions were identified within groups who had varying levels of entry skills as assessed through the MARKO-D test (pre-test).

This study presents a step towards the construction of a developmental assessment tool which aims to support teachers in their professional decision-making supported by an empirically based learning progression.

### Einstellungen und Selbstwirksamkeitsüberzeugungen von Lehrkräften als Prädiktoren für ihre Sichtweisen auf die Qualität von inklusivem Unterricht in der Grundschule

**Gamze Görel, Frank Hellmich**

Universität Paderborn, Deutschland; [gamze.goerel@uni-paderborn.de](mailto:gamze.goerel@uni-paderborn.de)

Mit der Ratifizierung der UN-Behindertenrechtskonvention durch die Bundesrepublik Deutschland stehen Lehrerinnen und Lehrer vor neuen unterrichtlichen Herausforderungen. Im Rahmen der Implementierung von Inklusion in der Schule ist es ihre Aufgabe, guten Unterricht für alle Schülerinnen und Schüler zu gestalten und den Bedürfnissen aller Kinder unabhängig von ihren je spezifischen Lernvoraussetzungen gerecht zu werden. Mit dieser Forderung sind Aspekte der Lehrerprofessionalität erneut als bedeutsame Faktoren in den Vordergrund der Forschung gerückt. Neben fachlich-didaktischem sowie pädagogischem Wissen umfasst Lehrerprofessionalität dabei auch Facetten wie Überzeugungen und Motivation (Baumert & Kunter, 2006). Im Zusammenhang mit der Frage nach Gelingensbedingungen für den inklusiven Unterricht in der Grundschule wurde in der vergangenen Zeit ein besonderer Fokus auf die Einstellungen von Lehrkräften zum inklusiven Unterricht gelegt, da angenommen wird, dass diese einen wichtigen Prädiktor für die erfolgreiche Umsetzung von inklusiven Konzepten in Schule und Unterricht bei Lehrkräften darstellen (Sze, 2009). Unter Einstellungen werden dabei psychologische Tendenzen verstanden, die sich in Form einer Annahme oder Ablehnung bei der Bewertung eines Einstellungsobjekts ausdrücken (Eagly & Chaiken, 1993). In vielen Studien wurden Einstellungen von Lehrkräften zur Inklusion beispielsweise im Zusammenhang mit Selbstwirksamkeitsüberzeugungen bezüglich des inklusiven Unterrichts untersucht. Dabei konnte gezeigt werden, dass Einstellungen und Selbstwirksamkeitsüberzeugungen von Lehrkräften miteinander zusammenhängen (Savolainen, Engelbrecht, Nel & Malinen, 2012). Selbstwirksamkeit wird hierbei als Überzeugung darüber verstanden, ein spezifisches Verhalten zur Erreichung von Zielen ausführen zu können (Bandura, 1997). Gegenwärtig ist noch weitgehend ungeklärt, welche Bedeutung Einstellungen von Lehrkräften zur Inklusion und Selbstwirksamkeitsüberzeugungen in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts für ihre Sichtweisen auf Qualitätsmerkmale inklusiven Unterrichts haben.

Vor diesem Hintergrund gehen wir im Rahmen unserer Studie der Frage nach, ob und inwiefern persönliche Ressourcen von Lehrkräften wie Einstellungen und Selbstwirksamkeitsüberzeugungen prädiktiv für ihre prospektiven Sichtweisen auf Qualitätsmerkmale von inklusivem Unterricht (Umgang mit Heterogenität, Strukturierung, Klarheit und lernförderliches Klima) sind. Wir nehmen an, dass Einstellungen zur Inklusion, Selbstwirksamkeitsüberzeugungen in Hinblick auf die eigene Gestaltung inklusiven Unterrichts sowie die Motivation von Lehrkräften zur Beschäftigung mit dem Thema Inklusion ihre Sichtweisen auf Qualitätsmerkmale von inklusivem Unterricht erklären. Im Detail gehen wir davon aus, dass die Einstellungen der Lehrkräfte zur Inklusion und ihre Motivation zur Beschäftigung mit dem Thema Inklusion durch ihre

Selbstwirksamkeitsüberzeugungen signifikant erklärt werden. Diese persönlichen Ressourcen stellen wiederum signifikante Prädiktoren für die prospektiven Sichtweisen von Lehrkräften auf Qualitätsmerkmale von inklusivem Unterricht (Umgang mit Heterogenität, Strukturierung, Klarheit und lernförderliches Klima) in der Grundschule dar.

Im Rahmen unserer Studie wurden N=168 Grundschullehrkräfte mit einem 'Paper and Pencil'-Fragebogen zu ihren Einstellungen zur Inklusion, zu ihrer Motivation für die Beschäftigung mit dem Thema Inklusion sowie zu ihren Selbstwirksamkeitsüberzeugungen befragt. Darüber hinaus wurden mit dem Fragebogen ihre prospektiven Sichtweisen zu verschiedenen Qualitätsmerkmalen von inklusivem Unterricht (Umgang mit Heterogenität, Strukturierung, Klarheit und lernförderliches Klima) erhoben. Zur Überprüfung der Hypothesen wurde ein Strukturgleichungsmodell in Amos berechnet.

Die Ergebnisse aus dem Strukturgleichungsmodell verdeutlichen, dass die Selbstwirksamkeitsüberzeugungen prädiktiv für die Einstellungen und die Motivation der an der Studie beteiligten Grundschullehrkräfte sind. Darüber hinaus wird die Motivation der Lehrkräfte zur Beschäftigung mit dem Thema Inklusion durch ihre Einstellungen zur Inklusion erklärt. Die Motivation der Lehrkräfte trägt zur Vorhersage ihres prospektiven Umgangs mit Heterogenität bei. Die Einschätzung der Lehrkräfte zum Umgang mit Heterogenität im inklusiven Klassenzimmer ist schließlich jeweils prädiktiv für ihre Sichtweisen auf Qualitätsmerkmale inklusiven Unterrichts (Strukturierung, Klarheit und lernförderliches Klima). Die Ergebnisse unserer Studie deuten darauf hin, dass persönliche Ressourcen von Grundschullehrkräften als wichtige Bedingungen für ihre Einschätzungen in Hinblick auf die Gestaltung von inklusivem Unterricht gesehen werden können. Die ermittelten Befunde können in Fort- und Weiterbildungen von Lehrkräften Berücksichtigung finden.

## Welches Wissen brauchen LehrerInnen für inklusiven Unterricht? Implikationen für die Professionalisierung von Lehrkräften

Franziska Greiner<sup>1</sup>, Stefanie Czempiel<sup>1</sup>, Sabine Sommer<sup>1</sup>, Bärbel Kracke<sup>1</sup>, Ada Sasse<sup>2</sup>

<sup>1</sup>Friedrich-Schiller-Universität Jena, Deutschland; <sup>2</sup>Humboldt-Universität zu Berlin, Deutschland; [franziska.greiner@uni-jena.de](mailto:franziska.greiner@uni-jena.de), [stefanie.czempiel@uni-jena.de](mailto:stefanie.czempiel@uni-jena.de), [sabine.sommer@uni-jena.de](mailto:sabine.sommer@uni-jena.de), [baerbel.kracke@uni-jena.de](mailto:baerbel.kracke@uni-jena.de)

### THEORETISCHER HINTERGRUND

Mit der Ratifizierung der UN-Behindertenrechtskonvention geht die Anforderung an alle Lehrkräfte einher, inklusive Lehr-Lern-Settings professionell gestalten zu können. Dafür müssen künftige Lehrkräfte bereits in ihrer ersten Ausbildungsphase qualifiziert werden. Um entsprechende universitäre Lernangebote konzipieren und implementieren zu können, ist es zunächst notwendig, die für die professionelle Gestaltung inklusiven Unterrichts erforderlichen Kompetenzen (Einstellungen, Wissen und Fähigkeiten) zu identifizieren (Bosse et al., 2015). In bisheriger Forschung wurde den Einstellungen zum inklusiven Lernen besondere Aufmerksamkeit geschenkt (Gebhardt et al., 2015) und festgestellt, dass es sich bei diesen Einstellungen um ein stabiles, schwer veränderbares Konstrukt handelt (Bosse et al., 2016). Vor diesem Hintergrund erscheint die Fokussierung inklusionsrelevanten Professionswissens als veränderbare Kompetenzkomponente umso wichtiger für die Gestaltung universitärer Lerngelegenheiten (Melzer et al., 2015; Kunter et al., 2011).

Obwohl inzwischen Anforderungsbereiche von Lehrkräften in inklusiven Settings definiert wurden (u.a. Wittek, 2016; Moser, 2013), steht eine systematische Beschreibung der Bereiche und Facetten des Professionswissens weiterhin aus (Hillenbrand et al., 2013).

### FRAGESTELLUNG UND HYPOTHESEN

Der vorliegende Beitrag fokussiert daher die Frage, welche Wissensbestände notwendig sind, um inklusive Settings professionell gestalten zu können. Es werden Ergebnisse der Analyse nationaler und internationaler bildungspolitischer Dokumente mit Befragungen von Lehrkräften, die in Bezug auf Inklusion in unterschiedlich weit entwickelten Schulen tätig sind, verknüpft. Es wird erwartet, dass inklusionserfahrene Lehrkräfte andere Wissensbestände fokussieren als weniger erfahrene. Die Erkenntnisse dieser Synthese sollen in ihrer Bedeutung für die LehrerInnenbildung diskutiert werden.

### METHODEN

Aus der Dokumentenanalyse konnte extrahiert werden, dass "Kooperation", "Diagnostik", "adaptive Unterrichtsgestaltung" und Wissen über "Sonderpädagogische Förderbedarfe" vier zentrale Bereiche des erforderlichen professionellen Wissens darstellen. Diese Wissensbereiche wurden als deduktive Kategorien für die Auswertung von teilstandardisierten Interviews mit LehrerInnen herangezogen, die im Rahmen des Projekts „Expertise zum Gemeinsamen Unterricht in der Stadt Jena“ durchgeführt wurden. In einem Expertenratung (Sommer et al., 2017) wurde der Entwicklungsstand der teilnehmenden Schulen bezüglich Inklusion basierend auf Prinzipien guten gemeinsamen Unterrichts in „Anfänger“, „Mittelgruppe“ und „Fortgeschrittene“ eingeteilt.

Da wenig über Inklusion im Bereich der weiterführenden Schulformen bekannt ist, wurden für die vorliegende Studie gezielt die Aussagen von 21 LehrerInnen (♂ 4 ♀ 17), die an 13 verschiedenen weiterführenden Schulen (6 Gemeinschaftsschulen, 3 Gesamtschulen und 4 Gymnasien) unterrichteten, in den Fokus genommen.

Die Antworten der Lehrkräfte auf die Frage „Welches Wissen brauchen LehrerInnen, um Gemeinsamen Unterricht zu gestalten?“ wurden gemäß der qualitativen Inhaltsanalyse nach Mayring (2010) ausgewertet.

### ERGEBNISSE

Die vier aus der Dokumentenanalyse abstrahierten Wissensbestände wurden auch von den Lehrkräften als wichtige Voraussetzung für die Gestaltung inklusiven Unterrichts genannt. Allerdings werden abhängig von dem Entwicklungsstand der Schule unterschiedliche Bereiche als relevant erachtet bzw. priorisiert. Es zeigte sich, dass die „Anfänger“ am häufigsten „Wissen über sonderpädagogische Förderbedarfe“ nannten. Die „Fortgeschrittenen“ dagegen erachteten „Wissen über Kooperation“ und „Adaptive Unterrichtsgestaltung“ als besonders zentral.

Diese Ergebnisse weisen darauf hin, dass sich Lehrkräfte mit wenig Inklusionserfahrung vor allem durch spezifische Kenntnisse über sonderpädagogische Förderbedarfe von SchülerInnen für inklusiven Unterricht gerüstet fühlen. Hingegen scheinen Lehrkräfte mit mehr Erfahrungen einen systemischen Blick auf Schule und Unterricht zu haben und nennen dementsprechend komplexere Wissensbereiche.

Es lässt sich resümieren, dass sich inklusiver Unterricht aktuell auf verschiedenen Entwicklungsstufen bewegt. Daher sollten sich Studierende – insbesondere in der Praxisphase – mit dem Entwicklungsstand der Schulen zu Inklusion auseinandersetzen und reflektieren, welches pädagogische Handeln für ihre professionelle Entwicklung beispielhaft ist. Für eine derartige reflexive Auseinandersetzung benötigen Studierende „Inklusionswissen“, das sie bereits im Laufe ihres Studiums erwerben sollten. Zu diskutieren ist, ob der Wissenserwerb im Rahmen universitärer Lerngelegenheiten an den Entwicklungsstand der Studierenden

anzupassen ist, indem beispielsweise Lehrveranstaltungen auf verschiedenen Anforderungsstufen angeboten werden. Überdies ist zu klären, ob sich die vorgestellten Wissensbereiche dem pädagogischen Wissen zuordnen lassen und beispielsweise als Differenzierung der Facette "Wissen über Heterogenität" (Voss, 2015) modelliert werden können.

## **Belastungserleben von Lehrkräften in inklusiven Schulsettings**

**Milena Peperkorn, Katharina Müller, Peter Paulus**

Leuphana Universität Lüneburg, Deutschland; [milena.peperkorn@leuphana.de](mailto:milena.peperkorn@leuphana.de)

### Theoretischer Hintergrund und Fragestellung

In Deutschland liegen bei etwa 20% aller Lehrkräfte belastungsinduzierte Gesundheitsgefährdungen vor (Rothland, 2013; Schaarschmidt & Kieschke, 2007, 2012), wobei insbesondere die Anzahl psychischer Belastungen gestiegen ist (vgl. Lehr, 2014). Es ist davon auszugehen, dass sowohl interne als auch externe Einflüsse einer Person, wie etwa schulische Rahmenbedingungen, zu Belastungen führen.

Aufgrund der Maßnahmen zur Umsetzung der UN-Behindertenrechtskonvention deutlich veränderten schulischen Rahmenbedingungen und der auf Landesebene vorgenommenen rechtlichen Verordnungen Schulen „inklusiv“ auszurichten (z.B. §4 NSchG), ist von Mehrbelastungen für Lehrkräfte inklusiver Lerngruppen auszugehen (Amrhein, 2011). Internationale (Ervasti et al., 2012; Talmor et al. 2005) und nationale (Peperkorn et al., 2017) Forschungsarbeiten konnten inklusives Unterrichten als gesundheitsbeeinflussenden Faktor identifizieren. An diese Ergebnisse anknüpfend steht im Mittelpunkt des Beitrags die Frage: Welche Faktoren beeinflussen das Belastungserleben von Lehrkräften im Kontext inklusiver Unterrichtstätigkeiten?

Da die Forschungsarbeiten zu Belastungen von Lehrkräften auf Grund der starken Fokussierung auf personen- bzw. verhaltensbezogene Faktoren (Wesselborg, 2015) in der Kritik stehen (Rothland, 2009), wurde in der vorliegenden Untersuchung als theoretische Grundlage das empirisch überprüfte, systemische Anforderung-Ressourcen-Modell (Becker, 2006) zur Erklärung der Entstehung und Erhaltung von Gesundheit gewählt. Im Modell stehen externe Ressourcen und externe Anforderungen als Umweltmerkmale sowie interne Ressourcen und interne Anforderungen als psychophysische Merkmale einer Person in wechselseitiger Beziehung miteinander.

### Studiendesign und methodisches Vorgehen

2016 wurde niedersachsenweit die Befragung „Mehr Zeit für gute Schule“ als Fragebogenerhebung online mit Lehrkräften, Schulleitungen und pädagogisch Mitarbeitenden durchgeführt (n=9868). Eine kombinierte Gewichtung anhand soziodemografischer Merkmale gewährleistet die Repräsentativität der Stichprobe. Ziel der Studie war der Informationsgewinn zum Einfluss bildungspolitischer Reformen sowie deren Umsetzung am Arbeitsplatz Schule. Teilnehmende wurden gebeten Tätigkeiten, die sich aus geltenden Rechts- und Verwaltungsvorschriften sowie administrativen Tätigkeiten zusammensetzen, in den Dimensionen „Verstehbarkeit“, „Sinnhaftigkeit“, „Bewältigbarkeit“ (in Anlehnung an das Konstrukt des Kohärenzgefühls, vgl. Antonovsky, 1997), in der wahrgenommene „Unterstützung“ sowie deren Vorkommen über die Dimension „Häufigkeit“ einzuschätzen. Acht der Tätigkeitsbeschreibung beziehen sich auf inklusive Bildung. Aus den Werten der erlebten Unterstützung sowie den drei Kohärenzwerten, als Konstrukten gesundheitsförderlicher Ressourcen (Paulus et al., 2017), wird im Studiendesign die erlebte Belastung über Mittelwertbildung berechnet. Die Skala zur Gesamtbelastung im inklusiven Unterricht zeigt eine gute Reliabilität (Cronbachs  $\alpha = .86$ ).

Zusätzlich zu den Arbeitstätigkeiten wurden soziodemografische Angaben der Teilnehmenden sowie Skalen zu arbeitsbedingten Be- und Entlastungsfaktoren, wie beispielsweise „Lehrer-Selbstwirksamkeit“ (10 Items; Cronbachs  $\alpha = .81$ ) (Schwarzer & Schmitz, 1999), „Mitarbeiterorientierte Führung“ (4 Items; Cronbachs  $\alpha = .92$ ), „Handlungsspielraum“ (4 Items; Cronbachs  $\alpha = .68$ ), „Gemeinsame pädagogische Überzeugungen“ (5 Items; Cronbachs  $\alpha = .91$ ) (alle Kaempff & Krause, 2004) und „Irritation“ (7 Items; Cronbachs  $\alpha = .87$ ) (Mohr, 1986) erhoben. Die zur Verfügung stehenden Skalen und Variablen werden forschungsbasiert den Aspekten des theoretischen Rahmenmodells zugeordnet. Regressionsanalytisch wurde in verschiedenen Modellen der Einfluss interner und externer Anforderungen und Ressourcen auf das „Belastungserleben Inklusive Bildung“ untersucht.

### Ergebnisse und Diskussion

Die Konzeption der Untersuchung nimmt in der schulischen Gesundheitsförderungsforschung einen besonderen Platz ein: Erstmals wurden umfassend administrative Tätigkeitsvorgaben untersucht. Die Daten ermöglichen auf diese Weise verhältnisbezogene Aussagen zur Lehrkräftegesundheit.

Erste Ergebnisse zeigen, dass insgesamt 22% ( $R^2 = .21$ ,  $F(7,7685) = 299,264$ ,  $p < .001$ ) der Varianz des Belastungserlebens durch die untersuchten Variablen aufgeklärt werden können. Die wahrgenommenen Ressourcen klären dabei mehr Varianz auf ( $R^2 = .18$ ,  $F(4,7742) = 438,072$ ,  $p < .001$ ) als Anforderungen ( $R^2 = .07$ ,  $F(3,7727) = 215,092$ ,  $p < .001$ ). Lehrer-Selbstwirksamkeit hat als interne Ressource die stärkste Vorhersagekraft ( $R^2 = .13$ ,  $F(1,7745) = 1168,580$ ,  $p < .001$ ). Diese Werte verdeutlichen die Wichtigkeit der Ressourcenstärkung in der schulischen Gesundheitsförderung und legen eine Fokussierung auf ebendiese nahe.

Das Modell „Umweltbedingungen“ ( $R^2 = .13$ ,  $F(5,7734) = 243,781$ ,  $p < .001$ ) und das Modell „Psychophysische Merkmale“ ( $R^2 = .13$ ,  $F(2,7739) = 606,625$ ,  $p < .001$ ) zeigen einen gleichen  $R^2$ -Wert zur Varianzaufklärung. Dieser Befund verdeutlicht die Wichtigkeit der Berücksichtigung verhältnisorientierter Merkmale 1) in der Konzeption von Forschungsarbeiten im Kontext der Lehrkräftegesundheitsforschung sowie 2) bei der Planung und Umsetzung von Maßnahmen der gesundheitsbezogenen Aus- und Weiterbildung von Lehrkräften.

**Test Fairness of an Entrance Assessment in Economics - a Large-Scale-Study in Germany****Judith Jitomirski<sup>1</sup>, Olga Zlatkin-Troitschanskaia<sup>2</sup>, Carla Kühling-Thees<sup>2</sup>, Jasmin Schlax<sup>2</sup>, Roland Happ<sup>2</sup>, Sebastian Brückner<sup>2</sup>, Manuel Förster<sup>2</sup>, Dimitar Molerov<sup>1</sup>, Hans Anand Pant<sup>1</sup>**<sup>1</sup>Humboldt Universität zu Berlin, Deutschland; <sup>2</sup>Johannes Gutenberg Universität Mainz; [judith.jitomirski@hu-berlin.de](mailto:judith.jitomirski@hu-berlin.de)

An increasing student body in the study domain of business and economics (B&E) in higher education in a context of a growing migration, internationalization and diversification of entry conditions has also led to more heterogeneity among applicants (authors 2016), which requires a reevaluation of existing assessment practices. So far, research was conducted especially on validity as one of the most important criteria in educational assessment (cf., the Special Issue on Validity in Journal for Educational Measurement, 2013). In contrast, only few publications have focused on test fairness even though it is crucial in the AERA, APA and NCME Standards for Educational and Psychological Testing (AERA et al., 2014, Chapter 3 with 21 standards).

Entrance exams are not compulsory in B&E, but if administered have an impact on the planning of learning and teaching processes (authors 2017). Thus, besides validity, test fairness needs to be analyzed explicitly. Unfairness can be described as construct-irrelevant variance in items, and test scores between different subgroups of test takers differing in gender, migration background, etc. are put at a systematic disadvantage (Crocker, 2003). A deficit of test fairness and consequently biased test scores may negatively affect learning and teaching B&E in higher education (authors 2015). Assessments of students' prior domain-specific knowledge at the beginning of studies are sensitive instruments because they enable predictions on the development of content knowledge during the course of studies and offer diagnostic and prognostic insights for learning and teaching (authors 2017).

We assessed prior economic knowledge using the items from the Test of Understanding in College Economics (TUCEIV, Walstad, et al., 2007) adapted for Germany according to the Test Adaptation Guidelines by the International Test Commission (2016) and validated according to the Standards of Testing (AERA et al., 2014). To test for convergent validity, we also included the German adaptation of the Test of Economic Literacy [TELIV] (Walstad et al., 2013). The Berlin test for assessing crystalline and fluid intelligence (BEFKI 11+) was used to control for intelligence (Schipolowski et al., 2016). We analyzed the correlation between these variables and students' test scores and differentiated between construct-relevant and construct-irrelevant variance in students' test performance (Davies et al., 2005). The Germany-wide quasi-experimental study in 2016 was conducted with 7,621 entrance students of B&E from 54 universities.

The confirmatory factor analysis confirmed a theoretical hypothesized 2-factor structure based on economic knowledge and intelligence. The results indicate differences in students' test performance related to gender and native language; gender accounted for a larger part of the variance in addition to prior education and intelligence.

We analyzed measurement invariance and differential item functioning (DIF) to examine whether test takers achieved unusually high or low performance on items which cannot be explained by their controlled construct-relevant abilities. Analyses of measurement invariance and DIFs revealed construct-irrelevant influences, which should be reduced to avoid systematic disadvantages for certain student groups in score-based decisions.

Further, we thoroughly analyzed test performance related to students' native language on the item level. The results will be critically discussed in terms of test fairness and valid test score interpretations.

Implications discussed will include a comprehensive examination of test components and items systematically disadvantaging groups of students and consequences for redrafting test interpretation guidelines. Increasing heterogeneity of students' study preconditions in higher education economics requires fair teaching and learning opportunities and assessments. A clarification of construct-relevant and -irrelevant influences in students' test scores offers the higher education institutions more differentiated information to qualify measurement results and develop efficient teaching and learning opportunities related to the diverse demands of different students' groups.

**Entwicklung und Skalierung eines Tests zur Erfassung des Verstehens multipler Dokumente von Studierenden****Cornelia Schoor<sup>1</sup>, Carolin Hahnel<sup>2,3</sup>, Cordula Artelt<sup>1,4</sup>, Ulf Kröhne<sup>2</sup>, Frank Goldhammer<sup>2,3</sup>**<sup>1</sup>Otto-Friedrich-Universität Bamberg, Deutschland; <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt, Deutschland; <sup>3</sup>Zentrum für internationale Vergleichsstudien (ZIB), Frankfurt am Main, Deutschland; <sup>4</sup>Leibniz-Institut für Bildungsverläufe e.V., Bamberg, Deutschland; [cornelia.schoor@uni-bamberg.de](mailto:cornelia.schoor@uni-bamberg.de)

Insbesondere im Hochschulkontext werden Studierende mit mehreren Quellen und somit (Text-) Dokumenten konfrontiert (z.B. als Ergebnis einer Datenbankrecherche), aus denen sie relevante Informationen herausfiltern und miteinander integrieren müssen. Dabei ist es notwendig, Informationen aus verschiedenen Quellen zueinander in Bezug zu setzen und unter Umständen Quellen (z.B. hinsichtlich ihrer Glaubwürdigkeit) zu bewerten (Documents Model Framework: z.B. Britt & Rouet, 2012; Strategien von Experten: Wineburg, 1991). Das Verstehen multipler Dokumente (Multiple Document Literacy, MDL) definieren wir daher als die Fähigkeit, aus verschiedenen Informationsquellen eine integrierte Repräsentation eines bestimmten Gegenstandsbereichs oder Sachverhalts zu konstruieren (vgl. Anmarkrud, Bråten & Strømsø, 2014, S. 64).

Bereits Abiturienten sollten wissenschaftliche Fachtexte für den eigenen Erkenntnisprozess nutzen können (Kultusministerkonferenz, 2012) bzw. diesbezüglich ihre Fähigkeiten im Laufe eines Studiums entwickeln (Arbeitskreis Deutscher Qualifikationsrahmen, 2011). Derzeit liegt jedoch kein Instrument vor, das MDL bei Studierenden angemessen und theoretisch fundiert erfasst. Es wurde daher auf der Basis des Documents Model Frameworks (z.B. Britt & Rouet, 2012) sowie der von Wineburg (1991) postulierten Strategien ein Test entwickelt, der vier kognitive Anforderungen von MDL abdeckt (Anforderungstypen): 1. Inhalte über Texte hinweg vergleichen, 2. Inhalte über Texte hinweg miteinander verknüpfen, 3. Quellen einschätzen und vergleichen (source-source links), 4. Inhalte als von Quellen generiert repräsentieren und Quelle-Inhalt-Kombinationen vergleichen (source-content links vergleichen).

Dieser Test wird im Rahmen des vorliegenden Beitrags auf seine Passung zu einem eindimensionalen Rasch-Modell sowie auf Differential Item Functioning (DIF) hin überprüft werden. Zudem wird die intendierte Konstruktinterpretation des Testwerts u.a. durch folgende Validierungshypothesen überprüft:

Der MDL-Testwert sollte (H1) positiv mit der Abiturnote, (H2) höher mit der Abiturnote von Bachelorstudierenden als mit der Abiturnote von Masterstudierenden korrelieren, (H3) für Masterstudierende höher ausfallen als für Bachelorstudierende. Da MDL als generische Kompetenz verstanden wird, werden keine bedeutsamen Unterschiede im MDL-Testwert in Abhängigkeit vom Studienfach (unter Kontrolle der Abiturnote) erwartet (H4). Im Mittel sollten Items des Anforderungstyps 1 eine größere Lösungswahrscheinlichkeit haben als Items (H5) des Anforderungstyps 4 (da nicht nur Einzelinhalte sondern Quelle-Inhalt-Kombinationen verglichen werden müssen) und (H6) des Anforderungstyps 2 (da Inhalte nicht nur gefunden, sondern zusätzlich verknüpft werden müssen).

An der Studie haben N = 310 Studierende im ersten Bachelor- oder Master-Semester zweier deutscher Universitäten verschiedener geistes- und sozialwissenschaftlicher Studienfächer (größte Gruppen: Pädagogik, Psychologie, Soziologie, Germanistik, Politik, Kommunikationswissenschaft) teilgenommen. In einem Multi-Matrix-Design wurden den Studierenden 3 aus 6 Units (Kombinationen aus 2-3 Texten und bis zu 16 Items) des Tests computerbasiert vorgelegt.

Aufgrund mangelnder Rasch-Konformität und DIF musste u.a. eine Unit ausgeschlossen werden. Der resultierende Test umfasst 67 Items in 5 Units, ist eindimensional und weist eine EAP Reliabilität von .69 auf. Die Abiturnote korrelierte in der Gesamtstichprobe signifikant mit dem MDL-Testwert ( $r = -.44$ ,  $p < .001$ ), d.h. mit einer besseren (kleineren) Abiturnote ging ein besserer (größerer) MDL-Testwert einher (H1). Der Zusammenhang war für Bachelorstudierende im ersten Semester ( $r = -.44$ ,  $p < .001$ ) größer als für Masterstudierende im ersten Semester ( $r = -.35$ ,  $p < .001$ ), allerdings war dieser Unterschied nicht signifikant ( $z = -0.81$ ,  $p = .42$ , H2). Masterstudierende erreichten signifikant höhere MDL-Testwerte als Bachelorstudierende (H3,  $M(\text{Master}) = 0.19$ ,  $SD(\text{Master}) = 0.69$ ,  $n(\text{Master}) = 93$ ,  $M(\text{Bachelor}) = -0.08$ ,  $SD(\text{Bachelor}) = 0.76$ ,  $n(\text{Bachelor}) = 204$ ,  $t(197.03) = -3.05$ ,  $p < .01$ ). Der Testwert war unter Kontrolle der Abiturnote unabhängig vom Studienfach (H4,  $F(10; 291) = 0.67$ ,  $p = .76$ ). Mittels eines generalisierten linearen gemischten Modells konnte gezeigt werden, dass Items des Anforderungstyps 2 und des Anforderungstyps 4 signifikant geringere Lösungswahrscheinlichkeiten als Items des Anforderungstyps 1 aufwiesen (H5, H6). Insgesamt liefern die Ergebnisse Belege dafür, dass der Testwert aus dem entwickelten MDL-Test die generische Fähigkeit von Studierenden wiedergibt, multiple Dokumente zu verstehen.

### **Valide Erfassung des studienrelevanten Fachwissens bei Studieneinsteigern in ein wirtschaftswissenschaftliches Masterstudium**

**Roland Happ, Olga Zlatkin-Troitschanskaia, Susanne Schmidt, Marie-Theres Nagel, Manuel Förster, Sebastian Brückner, Hannes Saas**

Johannes Gutenberg-Universität Mainz, Deutschland; [roland.happ@uni-mainz.de](mailto:roland.happ@uni-mainz.de), [susanne.schmidt@uni-mainz.de](mailto:susanne.schmidt@uni-mainz.de), [marie.nagel@uni-mainz.de](mailto:marie.nagel@uni-mainz.de)

#### Theoretischer Hintergrund

Die nationalen und internationalen Studien zur Erfassung des wirtschaftswissenschaftlichen Fachwissens Studierender fokussieren v.a. auf Bachelorstudierende (z.B. Kim et al. 2013; Schmidt et al. 2015; author 2016a). Zum Fachwissen bei Studieneinsteigern in ein wirtschaftswissenschaftliches Masterstudium, das in Deutschland zu den beliebtesten Studiengängen zählt (Statistisches Bundesamt 2016), lassen sich kaum testbasierte Befunde finden. Insbesondere mangelt es an Studien auf Basis validierter Instrumente, wie es in den aktuellen AERA Standards for educational and psychological testing (AERA et al. 2014) gefordert wird.

Für die Zulassung zu einem Masterstudium greifen die Hochschulen i.d.R. auf (formale) Leistungskriterien aus dem Bachelorstudium zurück. Hierzu zählt die Bachelorabschlussnote, die sich summativ aus den Modulprüfungsleistungen aus dem Bachelorstudiengang zusammensetzt. Dabei wird die Bachelornote als ein geeigneter Prädiktor für den Studienerfolg im Masterstudium angesehen, obgleich es diesbezüglich an empirischen Studien in den Wirtschaftswissenschaften mangelt (authors 2016b). Hingegen zeigte sich in mehreren Studien das studienrelevante Fachwissen als bester Prädiktor auf den Studienerfolg im Bachelorstudium (z.B. Dochy et al. 2002; author 2016c), wobei es auch hier an Analysen zum Masterstudium mangelt.

In unserer Studie haben wir ein inhaltsdimensions- sowie kognitionsniveaubezogenes Strukturmodell des wirtschaftswissenschaftlichen Fachwissens entwickelt (authors 2014). Auf Basis von bundesweiten curricularen Analysen sowie Expertenratings kann das Fachwissen in betriebs- und volkswirtschaftliche Inhaltsbereiche untergliedert werden (authors 2015).

#### Fragestellung

Im Vortrag gehen wir der Frage nach, inwieweit der Test zur Erfassung des wirtschaftswissenschaftlichen Fachwissens, der bislang im Bachelorstudium eingesetzt und nach den AERA et al. Standards (2014) umfassend validiert wurde (authors 2014), auch valide Rückschlüsse auf das vorhandene Fachwissen bei Studierenden zu Beginn des Masterstudiums erlaubt. Gemäß den AERA et al. Standards (2014) kann die Validität eines Tests u.a. anhand seines Zusammenhangs mit konzeptuell verwandten Konstrukten aus einem nomologischen Netzwerk beurteilt werden. Im Vortrag fokussieren wir auf den Zusammenhang mit möglichen weiteren Indikatoren des wirtschaftswissenschaftlichen Fachwissens: HZB-Note, Bachelor-Abschlussnote, besuchte Lehrveranstaltungen im Bachelor-Studium sowie die Modulprüfungsnoten. Anzunehmen ist, dass die Bachelorabschlussnote sowie die Anzahl an und die Note in besuchten Fachveranstaltungen in einem positiven Zusammenhang mit dem wirtschaftswissenschaftlichen Fachwissen stehen. Dabei sollte das betriebswirtschaftliche Fachwissen einen stärkeren Zusammenhang mit der Note in und der Anzahl an betriebswirtschaftlichen Lehrveranstaltungen aus dem Bachelorstudiengang vorweisen; analog gilt es für die volkswirtschaftlichen Lehrveranstaltungen sowie Noten und das volkswirtschaftliche Fachwissen.

#### Methode

Zur Erfassung des volkswirtschaftlichen Fachwissens wurde auf 30 Items aus der deutschsprachigen Adaptation des Test of Understanding College Economics (TUCE IV; Walstad et al. 2008) zurückgegriffen. Bei dem betriebswirtschaftlichen Fachwissen wurden 28 Items für den Bereich Finanzierung sowie Rechnungswesen aus der deutschsprachigen Adaptation des Examen General para el Egreso de la Licenciatura (EGEL; CENEVAL 2010) eingesetzt (Zlatkin-Troitschanskaia et al. 2016a). Der gesamte Test besteht aus 58 Items, die auf sieben Testhefte im Booklet-Design verteilt wurden. In der Studie wurden an 40 Universitäten und Fachhochschulen deutschlandweit die Daten von insg. 1.492 Masterstudierenden zu Studienbeginn der Wirtschaftswissenschaften erfasst. Für die Auswertung der Daten im Testheftdesign wurden Plausible Values (PV) (Wu 2005) geschätzt und IRT-, SEM- und MEM-Modelle berechnet, wobei die Einflüsse weiterer personeller Faktoren wie der Migrationshintergrund, Gender etc. als Kontrollvariablen berücksichtigt wurden.

#### Ergebnisse

In der Präsentation werden die Analysen zur Modellierung des wirtschaftswissenschaftlichen Fachwissens von Masterstudierenden und zur Validierung des Fachtests vorgestellt, wobei die CFAs mit guten Model-Fit-Indizes die theoriekonforme interne Teststruktur bestätigen; dies gilt auch für die Faktorreliabilitäten. Die positiven Zusammenhänge zwischen den PVs und den kontrollierten Indikatoren (Bachelornote, Art und Anzahl der besuchten Lehrveranstaltungen) in den HLM-Modellen liefern weitere Evidenzen für die Validität des Tests, wobei die Validität der Testwertinterpretationen beim Einsatz im wirtschaftswissenschaftlichen Masterstudium mit Blick auf die weiteren AERA et al. Standards kritisch diskutiert wird.

### **BilWiss-Stud: Ein optimierter Test zur Erfassung des bildungswissenschaftlichen Wissens bei Lehramtsstudierenden**

**Olga Kunina-Habenicht<sup>1</sup>, Christina Maurer<sup>2</sup>, Hendrik Lohse-Bossenz<sup>3</sup>, Theresa Dicke<sup>4</sup>, Doris Holzberger<sup>5</sup>, Tina Seidel<sup>5</sup>, Detlev Leutner<sup>6</sup>, Mareike Kunter<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Karlsruhe, Deutschland; <sup>2</sup>Johann Wolfgang Goethe-Universität, Deutschland; <sup>3</sup>Pädagogische Hochschule Heidelberg, Deutschland; <sup>4</sup>Australian Catholic University; <sup>5</sup>Technische Universität München, Deutschland; <sup>6</sup>Universität Duisburg-Essen, Deutschland; [Olga.Kunina-Habenicht@dipf.de](mailto:Olga.Kunina-Habenicht@dipf.de)

**THEORETISCHER HINTERGRUND:** Es besteht Konsens, dass das professionsspezifische Wissen einen wichtigen Aspekt der professionellen Lehrerkompetenz darstellt und dass dieses Wissen in einem berufsbiografischen Prozess erworben werden kann (Baumert & Kunter, 2006). In Anlehnung an Shulman (1986) hat sich die Unterscheidung zwischen Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen durchgesetzt. In diesem Beitrag liegt der Fokus auf fachunspezifischem bildungswissenschaftlichem Wissen. Wir verstehen unter bildungswissenschaftlichem Wissen fachunspezifische deklarative (Wissen über Theorien Konzepte oder Modelle) und konzeptuelle Wissensanteile (tiefes Verständnis von grundlegenden Prinzipien und Verknüpfungen zwischen Wissenskomponenten). In Anlehnung an Shulmans Argumentation umfasst bildungswissenschaftliches Wissen nach diesem Konzept neben schulnahen Inhalten auch unterrichtsfernere theoretische fachunspezifische Kenntnisse bspw. über Bildungstheorie, Schulorganisation oder Wissen über den Lehrerberuf.

**FRAGESTELLUNG:** In diesem Beitrag wird auf der Grundlage des bereits bestehenden BilWiss-Wissenstests (Kunina-Habenicht et. al, 2013; Linninger et al. 2015) ein optimierter Test zur Erfassung dieses theoretischen bildungswissenschaftlichen Wissens vorgestellt, der unter anderem die Wissensentwicklung in den Bildungswissenschaften im Verlauf des Lehramtsstudiums abbilden soll. In der optimierten Testfassung wurde insbesondere der Inhaltsbereich „Unterrichtsgestaltung“ grundlegend überarbeitet. Im Beitrag steht die Untersuchung der psychometrischen Eigenschaften des Tests bzw. der Testitems im Vordergrund. Darüber hinaus sollen Hinweise für die Validität der Interpretationen der Testwertinterpretationen untersucht werden, indem die Zusammenhänge zwischen den Leistungen in den Teilskalen bzw. im Gesamttest mit dem Fortschritt im bildungswissenschaftlichen Teilstudium sowie Unterschiede zwischen verschiedenen Lehramtszugängen betrachtet werden.

**METHODEN:** Die Kurzform des optimierten Wissenstests beinhaltet 65 Items, die Langform des Tests 119 Items aus sechs verschiedenen Inhaltsbereichen. Die sechs Inhaltsbereiche sind Unterrichtsgestaltung (23 Items in der Langform), Lernen & Entwicklung (24 Items), Diagnostik & Evaluation (21 Items), Bildungstheorie (14 Items), Schule als Bildungsinstitution (19 Items), Lehrerberuf als Profession (18 Items). Der Fortschritt im bildungswissenschaftlichen Teilstudium wurde über die Anzahl der belegten Semesterwochenstunden und Credit Points erfasst. Test wurde als Online-Test im Sommersemester 2017 an vier Universitäten in Essen, Frankfurt, München und Tübingen eingesetzt. Aufgrund der großen Itemanzahl wurde ein Rotationsdesign mit vier Testheften verwendet.

Es liegen insgesamt Daten von 928 Personen vor, wobei die Überprüfung der Bearbeitungszeiten und Antwortmuster nahelegt, dass lediglich Antworten von 763 Personen eine ausreichend hohe Bearbeitungsqualität aufweisen. An der Studie nahmen 607 Frauen (ca. 65%) und 271 Männer teil; das durchschnittliche Alter betrug 22,5 Jahre (SD=3,5; Range: 17-46). Der Fortschritt im bildungswissenschaftlichen Teilstudium weist eine ausreichend große Varianz auf, wobei etwas weniger Personen am Ende des bildungswissenschaftlichen Studiums vertreten sind.

Die Überprüfung der Itemeigenschaften erfolgt derzeit mittels der Partial-Credit-Modellen (Muraki, 1992) im R-Paket TAM (Robitzsch, Kiefer & Wu, 2017), während die faktorielle Struktur des Kurztests in Mplus (Muthén & Muthén, 1998–2012) bzw. R Paket lavaan (Rosseel, 2012) überprüft wird.

**ERGEBNISSE:** Erste Analysen deuten auf eine ausreichende bis zufriedenstellende Reliabilität der Teilskalen im Kurztest hin (Chronbachs Alpha lag zwischen 0,59 und 0,79). IRT-Analysen und Überprüfung faktorieller Struktur sind Gegenstand aktueller Untersuchungen. Für die Validität der Testwertinterpretationen in der optimierten Testfassung spricht, dass Studierende im Grundschullehramt sowie im Lehramt für Sonderpädagogik signifikant höhere Testwerte im Kurztest aufweisen als Studierende im Lehramt an Haupt- und Realstudien. Dieses Ergebnis ist plausibel, da Studierende in den zuerst genannten Studiengängen in der Regel mehr bildungswissenschaftliche Veranstaltungen besuchen als Studierende im Lehramt an Haupt- und Realstudien. Zudem zeigt sich erwartungskonform eine geringe aber signifikante Korrelation zwischen der Wissensleistung im Kurztest und prozentualem Fortschritt im bildungswissenschaftlichen Studium ( $r=.13$ ,  $p=.001$ ). Diese Ergebnisse werden diskutiert in Bezug auf methodische Herausforderungen und Implikationen für Lehrerbildung.



## B05

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Rosshofgasse, S01

### **Sprachförderung professionalisieren - zum Zusammenhang von fachdidaktischem Wissen und Schülerleistung**

**Isabell Deml<sup>1</sup>, Astrid Rank<sup>1</sup>, Anja Wildemann<sup>2</sup>, Anita Schilcher<sup>1</sup>, Stefan Krauss<sup>1</sup>, Linda Lenske<sup>2</sup>, Alexandra Merkert<sup>2</sup>, Magdalena Schulte<sup>1</sup>, Karin Binder<sup>1</sup>, Lena Bien-Miller<sup>2</sup>**

<sup>1</sup>Universität Regensburg, Deutschland; <sup>2</sup>Universität Koblenz-Landau; [isabell.deml@ur.de](mailto:isabell.deml@ur.de), [astrid.rank@ur.de](mailto:astrid.rank@ur.de)

#### Theoretischer Hintergrund

Seit einigen Jahren beschäftigt sich die Mathematikdidaktik zunehmend mit dem Thema „Sprache und Mathematik“, womit der Begriff der Bildungssprache vermehrt in den Fokus rückt (Schilcher et al. 2017). Dass Schülerinnen und Schüler im Allgemeinen fachlich nur gute Leistungen erzielen, wenn sie auch die sprachlichen Anforderungen bewältigen können, ist eine nachvollziehbare Schlussfolgerung (vgl. Gogolin o.J., 6). Für Fachlehrkräfte bedeutet dies, dass sie neben ihren fachspezifischen Kompetenzen auch über Wissen im sprachlichen Bereich verfügen müssen, zumal sie es sind, die maßgeblich zur Schülerleistung beitragen (Kunter et al. 2011; Helmke 2012; Krauss et al. 2017). Darüber hinaus müssen sie die (neue) Aufgabe, nun auch Sprache im Fachunterricht zu fördern, annehmen und umsetzen wollen.

#### Fragestellung

Im Evaluationsprojekt „Eva Prim“ wird untersucht, welches Professionswissen im Sinne von Shulman (1987) Lehrkräfte mitbringen müssen (oder nicht mitbringen müssen), um bildungssprachliche Fähigkeiten im Mathematikunterricht der Grundschule fördern zu können. Es werden außerdem Einstellungen und Haltungen zum Thema Sprachförderung im Fachunterricht erhoben. Bei den getesteten Lehrkräften handelt es sich um eine Gruppe, die im Rahmen der BiSS-Initiative an der Konzeptualisierung und Professionalisierung effektiven Sprachförderhandelns arbeitet. Besonders spannend ist es also, ob sich im Zuge dieser Professionalisierungsmaßnahme das Wissen der Lehrkräfte verändert. Um eventuelle Auswirkungen auf die Leistungen der Schülerinnen und Schüler feststellen zu können, werden deren fachliche und bildungssprachliche Fähigkeiten getestet. Ziel ist es zu zeigen, inwieweit sich das fachdidaktische Wissen der Lehrkräfte im Bereich Sprachförderung und Mathematik verändert und ob diese Kompetenzen tatsächlich prädiktiv für die Schülerleistungen sind.

#### Methode

Da bei der Evaluation eventuelle Veränderungen bei den Lehrkräften (N = 83) und Schülerinnen und Schülern (N = 809) aufgezeigt werden sollen, werden Daten an drei Messzeitpunkten (Schuljahre 2015/2016, 2016/2017, 2017/2018) erhoben. Im Fokus der Untersuchung stehen sowohl die fachdidaktischen Kompetenzen sowie Einstellungen und Haltungen der Lehrkräfte im Bereich Sprachförderung, die mit dem selbstentwickelten Instrument „FaWi-S“ getestet wurden, als auch die fachdidaktischen Fähigkeiten und Haltungen im Bereich Mathematik, die mit dem ebenfalls selbstentwickelten Instrument „FaWi-M“ getestet wurden. Am ersten Messzeitpunkt wurden zunächst die kognitiven Fähigkeiten („CFT 1-R“), die fachlichen Kompetenzen im Bereich Mathematik („DEMAT“) und der rezeptive Wortschatz („PPVT“) abgefragt. Die Lesefertigkeit („SLS 2-9“) und die allgemeinen bildungssprachlichen Fähigkeiten (Hörverstehenstest „BiSpra“) wurden am ersten und zweiten Messzeitpunkt erhoben. Das in einem Promotionsprojekt entstandene Testverfahren „SAMT“ (Sprachliche Ausdrucksfähigkeit in Mathematik) kam am zweiten Messzeitpunkt neu hinzu.

#### Ergebnisse

Die Testverfahren „FaWi-S“ (max. erreichbare Punktzahl: 34) und „FaWi-M“ (max. erreichbare Punktzahl: 8) zeigen ein durchschnittliches Ergebnis der Lehrkräfte in den Bereichen Sprachförderung und Mathematik (FaWi-S: M = 19,46, SD = 3,80; FaWi-M: M = 4,54, SD = 1,54). Bemerkenswert sind die großen Unterschiede innerhalb der Stichprobe (FaWi-S: emp. Min.: 4,3; emp. Max.: 27,5; FaWi-M: emp. Min.: 0,6; emp. Max.: 7,8). Betrachtet man die Einstellungen der Lehrkräfte zu Sprachförderung im Mathematikunterricht, zeigt sich eine durchaus positive Grundhaltung. Anders ist dies, wenn es um den Einbezug von Herkunftssprachen in den Unterricht geht.

Bei den Schülerleistungen fällt auf, dass die bildungssprachlichen Fähigkeiten an Messzeitpunkt eins und zwei in etwa gleich ausfallen, während sich die Leistungen bei der Lesefertigkeit leicht verschlechtern. Kognitive Fähigkeiten, fachliche Leistung in Mathematik und rezeptiver Wortschatz zeigen unauffällige, wenn auch sehr heterogene Werte. Inwieweit der vermutete Zusammenhang zwischen Lehrer- und Schülerkompetenzen in Bezug auf bildungssprachliche Fähigkeiten belastbar zu belegen ist, soll im Vortrag näher erläutert werden.

\*Anmerkung: Hörverstehenstest der Humboldt-Universität zu Berlin, IQB, Otto-Friedrich-Universität Bamberg; entstanden im Rahmen eines BMBF-Projektes von Stanat & Weinert: „Bildungssprachliche Anforderungen: Anforderungen, Sprachverarbeitung und Diagnostik“.

### **Netzeranalysen zur Beschreibung Integrierten Wissens zum Energiekonzept in der Physik**

**Marcus Kubsch**

Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel, Deutschland;  
[kubsch@ipn.uni-kiel.de](mailto:kubsch@ipn.uni-kiel.de)

Ein zentrales Ziel naturwissenschaftlicher Grundbildung ist, dass Schülerinnen und Schüler (SuS) in der Lage sind ihr Wissen in unterschiedlichen, alltagsnahen Kontexten anzuwenden um Phänomene zu interpretieren und zu erklären [1]. Hierfür müssen SuS vor dem Hintergrund des Kontexts auf Basis von Evidenz naturwissenschaftlich relevante Ideen miteinander verknüpfen [2] und artikulieren. Derartiges integrierte Wissen (Integrated Knowledge) [3] fokussiert also darauf wie SuS Wissensnetzwerke beim Lernen entwickeln. Dabei bauen sie Verbindungen zwischen Ideen auf, fügen neue Ideen hinzu und ändern und verfeinern Verbindungen zwischen Ideen. Werden ähnliche Ideen und Verbindungen wiederholt und in verschiedenen Kontexten aktiviert, entwickeln sich Muster die konzeptuelles Verständnis repräsentieren [4]. Erweisen sich in diesem Prozess bestimmte Ideen und Muster als erklärungsmächtig, steigt die Wahrscheinlichkeit, dass diese auch in Zukunft aktiviert werden [5]. SuS mit hohem integrierten Wissen können also auf stark entwickelte Wissensnetzwerke zurückgreifen, um vor diversen Kontexten Phänomene zu erklären, indem sie relevante Ideen verknüpfen.

In der Literatur findet sich allerdings jenseits mittels Construct Mappings [6] entwickelter Testaufgaben (z.B. bei [7]) kaum ein Versuch das Konstrukt des integrierten Wissens direkt zu erfassen der valide Schlüsse auf die Struktur und die Entwicklung der Wissensnetze bei SuS zuließe (siehe z.B. [8], [9]).

Dieser Forschungslücke begegnen wir dadurch, dass wir mit Hilfe Netzwerkanalytischer Methoden Wissensnetze von SuS und damit das Konstrukt des integrierten Wissens erfassen und folgende Fragen beantworten: (1) Welche Einblicke in die Entwicklung von Wissensnetzwerken von SuS können wir mit Netzwerkanalysen gewinnen? und (2) In wie weit entwickeln SuS ein integriertes Wissen von Energie um die Kernidee Energietransfer?

Eine neue Unterrichtseinheit zum Thema Energie wurde an zwei Schulen in den USA mit insgesamt drei Lehrkräften und ca. 300 SuS implementiert. N = 30 zufällig ausgewählte SuS wurden vor und nach der Einheit in halb-strukturierten Interviews fünf verschiedenen Phänomene zu alltagsnahe Situationen präsentiert und jeweils gebeten, diese zu erklären. Egal wie diese erste Erklärung ausfiel, wurde immer nach einer alternativen Erklärung gefragt.

Mit qualitativer Inhaltsanalyse [10] wurde kodiert welche Ideen SuS zur Erklärung nutzten (Interraterreliabilität Kappa = 0.87), sowie welche Ideen SuS pro Phänomen zusammen verwendeten. Mit Hilfe von Netzwerkanalytischen Methoden [11] können nun für alle SuS Wissensnetzwerke zum Prä- und Post-Zeitpunkt erstellt werden und somit sichtbar gemacht werden, welche Ideen SuS verwenden, welche Ideen SuS verbinden, welche Ideen zentral in den Wissensnetzwerken sind und ob sich Subnetzwerke bilden.

Als zentrale Tendenz zeigt sich, dass sich die Ideennetzwerke der SuS im Verlauf der Unterrichtseinheit von relativ unverbundenen kleinen Subnetzwerken in denen basale Ideen wie z.B. Geschwindigkeit dominieren zu hoch verbundenen wenig unterteilten Netzwerken entwickeln in denen komplexere Ideen wie Energietransfer oder Felder an Bedeutung gewinnen. Während z.B. die Idee Energietransfer zum Prä Zeitpunkt kaum verwendet wurde und relativ isoliert war, wurde sie zum Post Zeitpunkt häufig verwendet und ist im gesamten Netzwerk hochverbunden. Auch findet man zum Prä Zeitpunkt oft mehrere kleine Subnetzwerke von Ideen die sich zum Postzeitpunkt in insgesamt stark vernetzte Strukturen auflösen.

Es ist uns also gelungen, mit Hilfe der Netzwerkanalyse Wissensnetzwerke der SuS und damit die Integriertheit des Wissens zu erfassen und die Entwicklung um Kernideen, in diesem Fall im Energiekonzept, zu beschreiben. So wird eine genauere Diagnose des Lernstandes und Lernfortschritts möglich, was Lehrkräften im Rahmen ihres professionellen Handelns die Chance gibt adaptiv auf die Lerngruppe einzugehen. Um dies in größerem Umfang zu ermöglichen, muss allerdings noch, potentiell mit Hilfe der neuen Medien, eine Möglichkeit gefunden werden, die Ideen die SuS verwenden ohne aufwendige Interviews zu erfassen.

### **Diagnose und Entwicklung motorischer Basiskompetenzen**

**Christian Herrmann<sup>1</sup>, Christopher Heim<sup>2</sup>, Erin Gerlach<sup>3</sup>, Harald Seelig<sup>1</sup>**

<sup>1</sup>Universität Basel, Schweiz; <sup>2</sup>Universität Frankfurt, Deutschland; <sup>3</sup>Universität Potsdam, Deutschland; [christian.herrmann@unibas.ch](mailto:christian.herrmann@unibas.ch)

#### Theoretischer Hintergrund

Motorische Basiskompetenzen gewährleisten als erlernbare und funktionale Leistungsdispositionen, dass Kinder qualifiziert an der Sport- und Bewegungskultur teilnehmen können. Sie sichern damit eine basale Handlungsfähigkeit im Sport. Der Erwerb motorischer Basiskompetenzen ist eng an schulische und außerschulische Sozialisationsprozesse (u. a. Sportverein) gebunden und kann daher sehr unterschiedlich ausfallen. Dem Sportunterricht kommt daher die erzieherische Aufgabe zu einen Mindeststandard an motorischen Basiskompetenzen zu sichern (Herrmann, Seiler, Pühse & Gerlach, 2017; Schierz & Thiele, 2013). Mit dem MOBAK-Testinstrument (Herrmann, Gerlach & Seelig, 2015; Herrmann, 2018) können diese motorischen Basiskompetenzen im Sportunterricht jahrgangsstufenspezifisch und curricular valide erhoben werden.

Ziel des Beitrags ist es, die Entwicklung motorischer Basiskompetenzen über ein Schuljahr abzubilden und zu klären, welche Bedeutung endogene und exogene Variablen für die motorische Entwicklung haben.

#### Methode

Im Rahmen der Evaluation des Projekts „Schulkids in Bewegung“ (SKIB) in Frankfurt am Main wurden N=1031 Kinder der ersten und zweiten Jahrgangsstufe (54% Jungen, M<sub>t1</sub>=6.83 Jahre, SD=0.44) im Oktober 2014 (t<sub>1</sub>) und im Juni 2015 (t<sub>2</sub>) längsschnittlich untersucht. Erfasst wurden die motorischen Kompetenzbereiche Sich-Bewegen (Balancieren, Rollen, Springen, Laufen) und Etwas-Bewegen (Werfen, Fangen, Prellen, Dribbeln) mittels des MOBAK-Testinstruments (Herrmann et al., 2015) sowie Geschlecht, Alter, BMI und sportliche Aktivität der Kinder.

Mittels autoregressiver Strukturgleichungsmodelle wurde die Stabilität der motorischen Basiskompetenzen über die Zeit geprüft. Darauf aufbauend wurde der Einfluss endogener (Geschlecht, BMI, Alter) und exogener Variablen (Teilnahme am SKIB-Projekt und Vereinssport) auf die Veränderung der motorischen Basiskompetenzen untersucht (Herrmann, Heim & Seelig, 2017).

#### Ergebnisse

Die faktorielle Invarianz über die Zeit war für die beiden MOBAK-Faktoren Sich-Bewegen und Etwas-Bewegen gegeben, so dass das MOBAK-Testinstrument für Längsschnittanalysen genutzt werden konnte. Die Resultate ließen eine Verbesserung der motorischen Basiskompetenzen im Verlauf eines Schuljahres deutlich erkennen. Die motorische Kompetenzentwicklung war bei den Jungen stärker im Etwas-Bewegen und bei den Mädchen stärker im Sich-Bewegen ausgeprägt. Kinder mit hohem BMI zeigten im Vergleich zu Kindern mit geringerem BMI weniger Kompetenzzuwächse. Das Vereinsengagement im Mannschaftssport besaß einen positiven Einfluss auf die Entwicklung des Etwas-Bewegens, während ein Engagement im Individualsport einen positiven Einfluss auf die Kompetenzentwicklung im Sich-Bewegen hatte. Das SKIB-Programm stand mit seinen zwei zusätzlichen Sportstunden in einem positiven Zusammenhang mit der Entwicklung im Sich-Bewegen (Herrmann, Heim et al., 2017).

#### Diskussion

Das MOBAK-Testinstrument erwies sich als konstruktvalide und messinvariant. Die motorische Kompetenzentwicklung im Grundschulalter konnte damit adäquat abgebildet werden. Es zeigte sich, dass die endogenen Variablen Alter, Geschlecht und BMI sowie die exogene Variable Sportengagement in einem differentiellen Zusammenhang mit den motorischen Basiskompetenzen und deren Entwicklung standen.

### **Das Zusammenspiel von Dekodierfertigkeit und sprachlichen Fähigkeiten bei Grundschulkindern**

**Jan-Henning Ehm<sup>1</sup>, Telse Nagler<sup>1</sup>, Marcus Hasselhorn<sup>1</sup>, Arne Lervåg<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland; <sup>2</sup>University of Oslo; [ehm@dipf.de](mailto:ehm@dipf.de)

## Theoretischer Hintergrund

Am Anfang der Grundschulzeit lesen Kinder Wörter zunächst Buchstabe für Buchstabe. Bei der Verwendung dieser phonologischen Rekodierungsstrategie wird davon ausgegangen, dass vor allem die Kenntnis über die Buchstaben-Laut Zuordnungen, die phonologische Bewusstheit und das phonologische Arbeitsgedächtnis von zentraler Bedeutung sind. Der Abruf phonologischer Repräsentationen aus dem Langzeitgedächtnis, also die Geschwindigkeit, mit der schriftliche Symbole, wie z. B. Buchstaben und Buchstabenfolgen, in die zugehörigen Laute und Lauteinheiten umgewandelt werden können (Benennungsgeschwindigkeit), gilt neben diesen Faktoren als eine weitere wichtige Vorläuferfertigkeit für den Erwerb des Lesens (z. B. Wagner & Torgesen, 1987). Im Verlauf der Grundschule gelingt es Kindern immer besser über den lexikalischen Weg zu lesen, d. h. Wörter anhand bestehender mentaler Wortrepräsentationen auf der Basis der visuellen Wortgestalt direkt aus dem Gedächtnis abzurufen (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001). Für die Frage danach, welche Teilfertigkeiten für das Leseverständnis von Bedeutung sind, gibt die Theorie des Simple View of Reading eine frappierend einfache Antwort: Demnach lassen sich individuelle Unterschiede im Leseverständnis als das Produkt von zwei Teilfähigkeiten beschreiben, nämlich der Dekodierfertigkeit und allgemeiner sprachlicher Fähigkeiten (z. B. Wortschatz und Sprachverständnis; Sénéchal, Ouellette & Rodney, 2006). Der Theorie folgend, können jedoch sprachliche Fertigkeiten erst dann einen Einfluss auf das Leseverständnis ausüben, wenn die Dekodierfertigkeiten ausreichend stark ausgeprägt sind (multiplikative Verknüpfung). Bisherige Studien können diese Annahme weitestgehend bestätigen (siehe z. B. Wagner, Herrera, Spencer & Quinn, 2015). Allerdings sind diese Forschungsergebnisse auf den englischen Sprachraum beschränkt, zentrale theoretische Konstrukte werden in oft nur schlecht operationalisiert und es wird primär die Vorhersage des Leistungsniveaus und nicht des Wachstums im Lesen betrachtet.

## Fragestellung

Die vorliegende Studie geht den Fragen nach, (1) welche Faktoren das Niveau (Intercept) und das Wachstum (Slope) der Dekodiergeschwindigkeit und des Textverständnisses vorhersagen und (2) ob sich ein Interaktionseffekt zwischen der Dekodiergeschwindigkeit und den sprachlichen Fähigkeiten auf das Leseverständnis zeigt.

## Methode

Die Stichprobe bestand aus insgesamt 525 Schülerinnen und Schülern (Alter zum ersten Messzeitpunkt,  $M = 4;9$  Jahre,  $SD = 6$  Monate), deren schriftsprachliche Vorläufer zu drei Messzeitpunkten im Kindergarten und deren Leseleistung zu vier Messzeitpunkten in der Grundschule erfasst wurden (Ende der 1., 2. und 3. sowie Mitte der 4. Klasse). Die Erfassung der Leseleistung erfolgte anhand des WLLP-R (Dekodiergeschwindigkeit; Schneider, Blanke, Faust & Küspert, 2011) und des ELFE 1-6 (Subtest Textverständnis; Lenhard & Schneider, 2006). Die Analysen wurden mit Mplus 8 durchgeführt.

## Ergebnisse

Es zeigten sich bedeutsame Unterschiede im Ausgangsniveau und im Wachstum der Dekodiergeschwindigkeit ( $\sigma^2_{\text{Intercept}} = 179.05$ ,  $\sigma^2_{\text{slope}} = 153.00$ ) und des Textverständnisses ( $\sigma^2_{\text{Intercept}} = 14.69$ ,  $\sigma^2_{\text{slope}} = 2.40$ ). Buchstabenkenntnis, phonologische Bewusstheit, nonverbaler IQ und Benennungsgeschwindigkeit erwiesen sich als bedeutsame Prädiktoren der Dekodiergeschwindigkeit, die sprachlichen Fähigkeiten und die Dekodiergeschwindigkeit als bedeutsame Prädiktoren für das Textverständnis. Die Benennungsgeschwindigkeit erwies sich als bedeutsamer Prädiktor für das Wachstum der Dekodiergeschwindigkeit. Kinder mit einer hohen Benennungsgeschwindigkeit haben also nicht nur ein besseres Ausgangsniveau sondern zeigen auch ein größeres Wachstum im Lesen. Das Wachstum im Textverständnis konnte durch sprachliche Fähigkeiten und die Dekodiergeschwindigkeit vorhergesagt werden, wobei letzterer Effekt nichtlinear war (negativer curvilinear Effekt). Veränderungen im unterem Leistungsspektrum der Dekodierfertigkeiten gehen demnach mit stärkerem Wachstum im Leseverständnis einher als Veränderungen im oberen Leistungsspektrum der Dekodierfähigkeit. In Übereinstimmung mit dem Simple View of Reading Ansatz zeigte sich ein Interaktionseffekt zwischen Dekodiergeschwindigkeit und sprachlichen Fähigkeiten auf das Leseverständnis. Während bei geringer Dekodierfähigkeit eine Zunahme der sprachlichen Fähigkeiten nicht mit einer Verbesserung des Leseverständnisses zusammenhängt, ist dies bei höher ausgeprägten Dekodierfähigkeiten der Fall. Geringe Dekodierfähigkeiten beschränken demnach das Leseverständnis. Sind sie jedoch gut ausgeprägt, gewinnen sprachliche Fähigkeiten (z. B. Wortschatz) an Bedeutung.

## B06

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S212

### Die systematische Reflexion von Schülerrückmeldungen zum eigenen Unterrichtshandeln im Praxissemester

**Katharina Neuber, Kerstin Göbel**

Universität Duisburg-Essen, Deutschland; [katharina.neuber@uni-due.de](mailto:katharina.neuber@uni-due.de)

Die Bereitschaft und Fähigkeit zur Selbstreflexion gilt als wichtiges Merkmal der professionellen Handlungskompetenz von Lehrkräften (Combe & Kolbe, 2008) und soll bereits im Rahmen der Lehrerbildung gefördert werden. Schulpraktische Ausbildungsphasen weisen hierfür ein großes Potenzial auf, da Lehramtsstudierende Unterricht aus der Lehrendenperspektive kennenlernen und die Reflexion und Entwicklung eigenen Unterrichts erstmalig in den Blick nehmen können. Das Praxissemester im Lehramtsstudium bietet die Gelegenheit, durch systematische Maßnahmen zur Anregung von Reflexionsprozessen die Reflexionsbereitschaft und –fähigkeit angehender Lehrpersonen zu fördern und unterschiedliche Reflexionsbedingungen empirisch zu erforschen. In bisherigen Untersuchungen haben sich für die Unterstützung von Reflexionsprozessen insbesondere schriftliche und kollegiale Settings als ertragreich erwiesen (Hatton & Smith, 1995; Kämpfe, 2009). Als unterrichtsbezogene Reflexionsanregung können Studierende in Praxisphasen auch die Unterrichts-rückmeldungen der Lernenden nutzen. Die Schülerschaft hat sich über verschiedene Fächer hinweg als valide Informationsquelle für die Qualität von Unterricht erwiesen (Göbel & Hesse, 2008). Welche Relevanz die Unterrichts-rückmeldungen der Lernenden für die Reflexion des eigenen Unterrichts in den Praxisphasen angehender Lehrpersonen haben können, ist bislang noch kaum untersucht (Lawson, Çakmak, Gündüz & Busher, 2015).

Der vorliegende Beitrag nimmt die Nutzung von Schülerrückmeldungen zum eigenen Unterricht und damit verbundene Veränderungen der Reflexionsbereitschaft und der Einstellung zu Schülerrückmeldungen von Praxissemesterstudierenden in den Blick. Hierzu werden Befunde der Studie „Schülerrückmeldungen zum Unterricht und ihr Beitrag zur Unterrichtsreflexion im Praxissemester (ScRiPS)“ vorgestellt (Göbel & Neuber, 2017). Der Einsatz von Schülerrückmeldungen wird in einem quasi-experimentellen Design mit drei Interventions- und einer Kontrollgruppe untersucht. Die Interventionsgruppen holen zu zwei Zeitpunkten zu zwei Zeitpunkten im Praxissemester systematische Schülerrückmeldungen zu ihrem Unterricht ein und werden bei der Reflexion der Rückmeldeergebnisse auf unterschiedliche Weise unterstützt: IG1) sowohl individuell schriftliche als auch kollegiale Reflexion im Tandem (n=18), IG2) individuelle Reflexion mithilfe eines schriftlichen Reflexionsbogens (n=27) und IG3) keine zusätzliche Reflexionsanregung (n=40). Die Kontrollgruppe (n=30) nutzt kein schülerrückmeldungsgestütztes schriftliches oder kollegiales Reflexionssetting. Alle Untersuchungsteilnehmenden werden zu Beginn und zum Ende des Praxisseesters bzw. vor und nach der Schülerrückmeldungsintervention mithilfe standardisierter Fragebögen zu ihrer Reflexionsbereitschaft sowie bezüglich ihrer Einstellung zu Schülerrückmeldungen befragt; weiterhin wird nach Abschluss der Praxisphase die Akzeptanz der Reflexionsanregungen erfasst. Die Skala Reflexionsbereitschaft wurde im Rahmen des ScRiPS-Projekts entwickelt und umfasst die Dimensionen Relevanz von Unterrichtsreflexion im Lehrerberuf, Einstellung zur systematischen Reflexion, Einstellung zur kollegialen Reflexion und Einstellung zur individuellen Reflexion. Alle Items werden anhand einer vierstufigen Likert-Skala (‘stimme gar nicht zu’ bis ‘stimme voll und ganz zu’) eingeschätzt. Im Zuge erster Analysen haben sich die entwickelten Skalen als zufriedenstellend reliabel erwiesen (Neuber & Göbel, 2016).

Im Hinblick auf die Akzeptanz der Reflexionsanregungen zeigen die deskriptiven Ergebnisse, dass die Schülerrückmeldungsnutzung von den Interventionsgruppen insgesamt positiv beurteilt wird. Darüber hinaus schätzen diejenigen Studierenden, die sowohl individuell schriftlich als auch kollegial reflektierten (EG1), das schriftliche Reflexionsformat als weniger hilfreich für die Reflexion der Schülerrückmeldeergebnisse im Praxissemester ein als das kollegiale Setting. Im Hinblick auf die Veränderung der Reflexionsbereitschaft und der Einstellung zu Schülerrückmeldungen weisen die Mittelwertvergleiche auf Unterschiede zwischen den Untersuchungsgruppen hin. Eine leichte Zunahme der positiven Einstellung gegenüber der Nützlichkeit von Schülerrückmeldungen zum Unterricht zeigt sich nur für die Teilnehmenden der EG1 und der EG2. Zudem schätzen die Studierenden, die die Schülerrückmeldungen sowohl individuell schriftlich als auch kollegial reflektierten (EG1), die allgemeine Nützlichkeit systematischer Reflexionsformate nach der Praxisphase weniger positiv ein als vor der Schülerrückmeldungsintervention.

Die Ergebnisse bekräftigen die Relevanz einer systematischen Reflexionsunterstützung für die Nutzung von Schülerrückmeldungen im Sinne der eigenen Unterrichtsreflexion und -entwicklung. Weiterhin weisen die Befunde auf die Notwendigkeit einer konzeptionellen Weiterentwicklung des schriftlichen Reflexionsformats hin.

### Wieviel Theorie braucht die Praxis? Zum Zusammenhang von bildungswissenschaftlichem Wissen und Elternberatungskompetenz bei Lehrkräften

**Christina Maurer<sup>1</sup>, Sara Ehlers<sup>1</sup>, Kristin Wolf<sup>1</sup>, Martin Gartmeier<sup>2</sup>, Silke Hertel<sup>3</sup>, Mareike Kunter<sup>1</sup>**

<sup>1</sup>Goethe-Universität Frankfurt, Deutschland; <sup>2</sup>Technische Universität München; <sup>3</sup>Ruprecht-Karls-Universität Heidelberg; [maurer@psych.uni-frankfurt.de](mailto:maurer@psych.uni-frankfurt.de)

#### Theoretischer Hintergrund

Neben der zentralen Aufgabe der Unterrichtsgestaltung ist die Zusammenarbeit mit Eltern bzw. deren Beratung ein weiteres wichtiges Handlungsfeld von Lehrkräften (Bruder et al., 2014; Hertel, 2016). Dementsprechend wird Beratungswissen auch als Teil der professionellen Kompetenz und damit als Voraussetzung für qualitativvolles Lehrerhandeln in Beratungssituationen verstanden (Baumert & Kunter, 2006). Dabei wird angenommen, dass sich theoretisches Wissen in der Analyse konkreter beruflicher Situationen manifestiert und so zum Initiieren geeigneter Handlungen beiträgt (Blömeke et al., 2015; Kunter et al., 2017). Der Aufbau einer solchen Wissensgrundlage soll schon in der universitären Phase der Ausbildung von Lehrkräften beginnen: Standards zur Lehrerbildung in den Bildungswissenschaften beschreiben für die universitäre Phase die Behandlung von Inhalten wie Grundlagen der Gesprächsformen und Ansätze professioneller Beratung (KMK, 2014). Die im bildungswissenschaftlichen Studium verfügbare Zeit ist allerdings für die Fülle der als relevant erachteten Themen sehr knapp bemessen (Linninger et al., 2015) und variiert zum Teil stark zwischen Bundesländern und auch einzelnen Universitäten (Bauer et al., 2012; Terhart et al., 2010). Kritiker stellen zudem sowohl die Struktur als auch die Praxisrelevanz der Inhalte der bildungswissenschaftlichen Studien in Frage (z.B. Merz, 2002). Vor dem Hintergrund dieser Herausforderungen ist es ein Anliegen der empirischen Forschung in der Lehrerbildung, die Praxisrelevanz curricularer Inhalte sichtbar zu machen. Für den Bereich des Unterrichtens gibt es inzwischen multiple empirische Hinweise, die für die theoretische Fundierung von

Unterrichtspraxis sprechen (z.B. Depaepe et al., 2013; Voss et al., 2015). Für die unterrichtsfernen Handlungsfelder von Lehrkräften wie der Elternberatung steht dies dagegen noch aus.

#### Fragestellung

Ziel des vorliegenden Beitrags ist es daher, die Relevanz von Grundlagenwissen zu Kommunikation und Beratung für den professionellen Umgang mit Elternberatungssituationen zu überprüfen. Dabei wird erwartet, dass Lehrkräfte mit höherem Grundlagenwissen zu Kommunikation und Beratung:

- eine bessere Analyseleistung beim Ansehen einer Elternberatungssituation, also einen höheren Anteil an bedeutsamen Erklärungsprozessen zeigen (Hypothese 1; vgl. dazu König et al., 2014),
- in Fallbeispielen typischer Elternberatungssituationen häufiger angemessene Verhaltensoptionen wählen (Hypothese 2)
- und sich selbst als kompetenter in Elternberatungssituationen wahrnehmen (Hypothese 3).

#### Methode

Die Untersuchungsdaten stammen aus einer Online-Studie mit 31 Lehrkräften unterschiedlicher Schulformen (Primar- und Sekundarstufe, Förder- und Berufsschule). Die Teilnehmer(innen) waren durchschnittlich 33 Jahre alt (SD=10; Range: 24-65), 65 % waren weiblich und die Berufserfahrung betrug durchschnittlich 6 Jahre (SD = 8; Range: 1-38). Erhoben wurde zum einen bildungswissenschaftliches Grundlagenwissen zu Beratung und Kommunikation mit fünf geschlossenen Aufgaben, die Teil eines inhaltlich umfassenderen Wissenstests zum bildungswissenschaftlichen Wissen sind (Kunter et al., 2017; Linninger et al., 2015). Zum anderen wurde Elternberatungskompetenz multikriterial erfasst. Dies umfasst erstens die professionelle Wahrnehmung in Elternberatungssituationen: Dafür sahen die Lehrkräfte einen Lehrfilm mit zwei Versionen eines Eltern-Lehrer-Beschwerdegesprächs (Gartmeier & Bahr, 2013) und analysierten das professionelle Verhalten der Lehrkraft in der Beratungssituation mit schriftlichen Reflexionen. Analog zur Erfassung der Professionellen Unterrichtswahrnehmung (vgl. König et al. 2014; Stürmer & Seidel, 2015) wurden die Analyseprozesse Beschreiben, Erklären und Handlungsalternativen generieren kodiert und quantifiziert. Zweitens wurde die Kurzversion eines Situational Judgment-Tests zur Erfassung der Elternberatungskompetenz eingesetzt (Klug et al., 2012). Dabei wählten die Lehrkräfte für sechs Fallbeispiele typischer Elternberatungssituationen aus vorgegebenen Antwortmöglichkeiten sinnvolle Verhaltensoptionen aus. Drittens wurde die Selbsteinschätzung der Beratungskompetenz mit einer 6-Item-Skala erfasst (Hertel, 2009).

#### Ergebnisse

In einem Pfadmodell in Mplus zeigt sich, dass sich Lehrkräfte mit höherem Grundlagenwissen zu Beratung und Kommunikation (unter Kontrolle der Berufserfahrung) erwartungsgemäß

- ihre Beobachtungen beim Analysieren von Beratungssituationen häufiger begründen, also eine tiefere Analyse gesehener Elternberatungssituationen vornehmen ( $\beta=1.01$ ),
- im Situational-Judgment Test zum Beratungsverhalten bessere Leistungen zeigen ( $\beta=0.37$ )
- und sich als kompetenter bei der Durchführung von Elternberatungen einschätzen ( $\beta=0.37$ ).

Die Befunde geben einen Hinweis auf die Relevanz theoriebasierter Beratungspraxis. Implikationen für die Lehrerbildung werden im Rahmen des Konferenzbeitrags diskutiert.

### **Prozedurales Fachwissen von angehenden Primarschullehrkräften: Entwicklung eines Instruments zur Erfassung der phonologischen Bewusstheit**

**Karin Berendes<sup>1</sup>, Richard Böhme<sup>2</sup>, Wolfgang Wagner<sup>1</sup>**

<sup>1</sup>Eberhard Karls Universität Tübingen, Deutschland; <sup>2</sup>Universität Regensburg, Deutschland; [karin.berendes@uni-tuebingen.de](mailto:karin.berendes@uni-tuebingen.de)

#### Theoretischer Hintergrund

Eine wesentliche Aufgabe des schriftsprachlichen Anfangsunterrichts ist die Förderung der phonologischen Bewusstheit (Klicpera, Schabmann, Gasteiger-Klicpera & Schmidt, 2017). Daher sollten (angehende) Lehrkräfte über ein umfassendes Fachwissen in diesem Bereich verfügen. Im englischsprachigen Raum deutet jedoch eine Vielzahl an Befunden darauf hin, dass (angehende) Lehrkräfte nicht über das notwendige Professionswissen verfügen, um Übungen zur phonologischen Bewusstheit kompetent anleiten zu können (z.B. Cheesman, McGuire, Shankweiler & Coyne, 2009; Spencer, Schuele, Guillot & Lee, 2008; Washburn, Binks-Cantrell, Joshi, Martin-Chang & Arrow, 2016; Stark, Snow, Eadie & Goldfeld, 2016). Im deutschsprachigen Raum konnte dieser Fragestellung mangels eines entsprechenden Verfahrens bislang noch nicht nachgegangen werden.

#### Ziel der Studie

Das Ziel der vorliegenden Untersuchung ist die Entwicklung eines Instruments zur Erfassung des prozeduralen Fachwissens von angehenden Primarschullehrkräften im Bereich der phonologischen Bewusstheit. Da Kinder im Anfangsunterricht das alphabetische Prinzip, also den regelhaften Zusammenhang von Graphemen und Phonemen erkennen müssen, kommt insbesondere Aufgaben zur phonologischen Bewusstheit auf Phonemebene eine hohe Bedeutung zu. Das entwickelte Verfahren fokussiert daher auf das prozedurale Fachwissen in diesem Bereich.

#### Methode

Es wurden Aufgaben zur Erfassung der phonologischen Bewusstheit auf Phonemebene mit unterschiedlichen Aufgabenformaten (Multiple Choice, halboffenes Antwortformat) auf unterschiedlichen Operationsebenen (z.B. Segmentieren, Analysieren) entwickelt. Diese Aufgaben mit insgesamt 37 Items wurden anschließend anhand einer Stichprobe von 148 Studierenden an vier baden-württembergischen Pädagogischen Hochschulen erprobt. Die Itemanalysen erfolgten auf Basis der Item-Response-Theorie (IRT; de Ayala, 2009). Dabei wurden zunächst die Diskriminationsparameter  $\alpha_i$  basierend auf einem zweiparametrischen logistischen (2PL-) Modell berechnet (ConQuest; Wu, Adams, Wilson & Haldane, 2007). Anschließend wurden Items mit Diskriminationsparametern  $\alpha_i$  unterhalb 0.8 ausgeschlossen (de Ayala, 2009) und die verbleibenden Items in einem einparametrischen (1PL-) Modell überprüft. Weiterhin wurde für diesen Itempool im Rahmen von Differential Item Functioning (DIF) Analysen geprüft, ob sich die Messeigenschaften in verschiedenen Subpopulationen der Stichprobe (relevant) voneinander unterscheiden.

#### Ergebnisse

Die Analysen basierend auf einem 2PL-Modell ergaben Diskriminationsparameter im Bereich von -0.056 bis 4.063, wobei insgesamt 17 Items Diskriminationsparameter unterhalb des Kriteriums von 0.8 aufwiesen. Für die verbleibenden 20 Items

ergaben sich in einem 1PL-Modell akzeptable Infit-Statistiken (weighted mean square, WMNSQ; Wright & Masters, 1982) im Bereich von  $0.88 \leq \text{WMNSQ} \leq 1.11$  und t-Statistiken ( $-1.1 \leq t_i \leq 1.2$ ) bei einer Varianz der latenten Variable von  $\text{Var}(\Theta) = 1.582$ . Die Itemschwierigkeiten deckten den Fähigkeitsbereich gut ab ( $-2.759 \leq \delta_i \leq 3.640$ ), erwiesen sich aber für die Stichprobe als im Durchschnitt tendenziell eher schwierig ( $M(\delta_i) = 0.615$  bei einer auf den Wert Null fixierten durchschnittlichen Personenfähigkeit). Trotz der relativ geringen Itemanzahl ergab sich eine akzeptable (Person-Separation-)Reliabilität (PSR) für die weighted likelihood estimates (WLE; Warm, 1989) von  $\text{WLE PSR} = 0.789$ . DIF-Analysen deuteten auf gruppenspezifische Itemschwierigkeiten hin (z.B. bezüglich der Verwendung von Dialekt in alltäglichen Situationen).

Fazit / Ausblick

Die psychometrischen Kennwerte der Itemauswahl sprechen dafür, dass das entwickelte Verfahren zur Erfassung des prozeduralen Fachwissens von angehenden Lehrkräften geeignet ist, um Leistungsunterschiede in dieser Gruppe mit hinreichender Genauigkeit abzubilden. Darüber hinaus kann vermutet werden, dass sich das Verfahren auch für den Einsatz bei bereits tätigen Lehrkräften eignet. Eine Erhebung mit dieser Zielgruppe steht jedoch noch aus. Mit dem fertig entwickelten Instrument wird es möglich sein, wichtige Informationen zum Fortbildungsbedarf von (angehenden) Lehrkräften zu generieren, um sie möglichst optimal auf die vor ihnen liegende Aufgabe des schriftsprachlichen Anfangsunterrichts vorbereiten zu können. Zudem kann mit dem Verfahren die Bedeutung des prozeduralen Fachwissens von Lehrkräften auf die frühen schriftsprachlichen Leistungen von Kindern erforscht werden.

## **Aufbau von vernetztem Fachwissen im Chemieunterricht der Sekundarstufe I auf der Basis von Learning Progressions**

**Kübra Nur Celik, Prof. Dr. Maik Walpuski**

Universität Duisburg-Essen, Deutschland; [kuebra.celik@uni-due.de](mailto:kuebra.celik@uni-due.de)

Der IQB-Ländervergleich 2012 hat gezeigt, dass 69,7 % der Schülerinnen und Schüler (SuS), die den mittleren Schulabschluss an einer nicht gymnasialen Schule anstreben, im Fach Chemie nicht die Kompetenzstufe 3 erreichen (Pant et al., 2013, S. 216). Vermutlich gelingt es leistungsschwachen SuS kaum vernetztes Wissen aufzubauen. Da das Fach Chemie stark hierarchisch strukturiert ist, könnten diese SuS so früh den Anschluss im Chemieunterricht verlieren. Um diese SuS zukünftig besser unterstützen zu können, ist es wichtig zu untersuchen, wie zentrale chemische Ideen und Konzepte in Abhängigkeit zueinander stehen. Deshalb wurden im Rahmen dieser Studie Learning Progressions zu den drei fachlichen Basiskonzepten der Chemie „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ (MSW, 2011) für die ersten beiden Lernjahre entwickelt. Diese beschreiben mögliche Wege der Kompetenzentwicklung und nehmen eine bestimmte Abfolge von Fähigkeiten und Wissens-elementen an, die die SuS im Laufe eines längeren Zeitraums erwerben (Abbott, 2014; Corcoran, Mosher, & Rogat, 2009; Duncan & Hmelo-Silver, 2009). Dazu wurden in diesem Projekt zunächst sogenannte fachliche „Kernideen“ beschrieben, die - analog zu AAAS (2007) - in einer hierarchisch-logischen Reihenfolge angeordnet und miteinander vernetzt wurden. Insgesamt wurden 57 Kernideen identifiziert, die jeweils durch fachliche Mindestanforderungen weiter ausdifferenziert, von nicht benötigtem Wissen abgegrenzt und durch typische Schülervorstellungen beschrieben wurden. Ziel der Studie ist es, die in den Learning Progressions abgebildeten hypothetischen Abhängigkeiten zwischen den Kernideen empirisch zu überprüfen, um Wissens-elemente, die für die Entwicklung des Verständnisses von chemischem Fachwissen ausschlaggebend sind von solchen zu unterscheiden, die für das weitere Verständnis weniger bedeutsam sind.

Daher fokussiert die Studie auf folgende Fragestellung:

Können die in den Learning Progressions angenommenen hypothetischen Abhängigkeiten empirisch nachgewiesen werden?

Um dieser Forschungsfrage nachzugehen, sollen SuS an Gesamtschulen im ersten und zweiten Unterrichtsjahr in Chemie in NRW getestet werden. Dazu wurde ein Fachwissenstest mit 330 Items im Multiple-Choice single-select Format konstruiert und in einer quasi-längsschnittlichen Pilotierung 787 SuS (50,3 % weiblich) im Multi-Matrix-Design zur Bearbeitung vorgelegt, um die Eignung der Testaufgaben zu untersuchen. In einem über vollständige Itemsätze zu den Kernideen balancierten Incomplete Block Design wurden in der Pilotierung durchschnittlich 32 Antworten pro Item erzielt. Die IRT-Analysen zeigen sehr gute Personen- (.828) und Itemreliabilitäten (.911). Die Testaufgaben sind für das erste Lernjahr signifikant schwieriger als für das zweite Lernjahr,  $t(328) = -6.798$ ,  $p \leq .001$ ,  $d = .751$ . Innerhalb eines Lernjahres sind keine signifikanten Unterschiede zwischen den Basiskonzepten vorzufinden.

In der Hauptstudie wird der überarbeitete Fachwissenstest zu zwei Messzeitpunkten eingesetzt, um die hypothetischen Abhängigkeiten zwischen den einzelnen Kernideen zu untersuchen. Es werden pro Beziehung zwischen zwei Kernideen Testdaten von ca. 100 SuS erhoben. Die Daten der Hauptstudie sollen es ermöglichen, belastbare Aussagen über die angenommenen Abhängigkeiten zu machen, um so auch Defizite diagnostizieren und insbesondere leistungsschwache SuS bei der Aufarbeitung ihrer Defizite unterstützen zu können. Die erstellte Map kann dabei als Orientierungs- und Strukturierungshilfe genutzt werden, um SuS an für den weiteren Wissensaufbau relevanten Schlüsselstellen gezielter fördern zu können.

Für die Auswertung werden die Lösungswahrscheinlichkeiten benachbarter Kernideen im Zusammenhang betrachtet. Geht man davon aus, dass zwischen den Kernideen „A“ und „B“ eine Abhängigkeit besteht (A ist nötig zum Verständnis von B), dann ist zu erwarten, dass idealerweise alle SuS, die die Aufgaben zur Kernidee B richtig lösen, auch die Aufgaben zu Kernidee A korrekt bearbeitet haben. Je größer der Anteil der B-Löser ist, die nicht die Aufgaben zur Kernidee A gelöst haben, desto stärker wird die angenommene Beziehung zwischen den Kernideen widerlegt. Da es kein standardisiertes Verfahren zur Analyse dieser angenommenen Beziehungen gibt, sollen Methoden mit unterschiedlichen Foki, wie der cross-lagged-panel-Analyse, dem McNemar-Test, der Guttman-Skala und den Bayesschen Netzen zur Auswertung herangezogen werden.

## B07

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H120

### Drei Fallstudien zur intensiven ICT-Nutzung in Schulen mit personalisierten Lernkonzepten (perLen)

**Regina Schmid<sup>1</sup>, Dominik Petko<sup>1</sup>, Christine Pauli<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Schwyz, Schweiz; <sup>2</sup>Universität Freiburg/Fribourg, Schweiz; [regina.schmid@phsz.ch](mailto:regina.schmid@phsz.ch)

#### Theoretischer Hintergrund

Personalisierte Ansätze des Lehrens und Lernens werden heute als vielversprechender Ansatz gesehen, um den Herausforderungen der Heterogenität von Schülerinnen und Schülern zu begegnen (Stebler, Pauli & Reusser, 2017). Gleichzeitig eröffnen Digitalisierungstrends neue Möglichkeiten, mit denen sich solche Ansätze unterstützen lassen. Doch trotz langjähriger Bemühung digitale Medien im Unterricht zu integrieren, zeigt sich in der Schweiz und in anderen hochtechnologisierten Ländern eine zögerliche ICT-Nutzung (Frailon et al., 2014; OECD, 2015). Oft wird eine geringe ICT-Nutzung damit erklärt, dass der traditionelle Unterricht und die neuen Lernmöglichkeiten durch digitale Medien zu wenig korrespondieren. Daher wurde in Pilotprojekten oft davon ausgegangen, dass der Einsatz von digitalen Medien einen Wandel in Richtung offene Lehr- und Lernkultur erfordert (Law, Pelgrum & Plomp, 2008; Venezky & Davis, 2002). Diesbezügliche Befunde sind allerdings noch nicht abschliessend. Einige Studien konnten aufzeigen, dass häufige ICT-Nutzung meist auch mit schülerorientierten Unterrichtsmustern einhergeht (Hermans et al., 2008; Petko, 2012). Jedoch der umgekehrte Zusammenhang – ob personalisierter Unterricht mit einer höheren Nutzung digitaler Medien korreliert und in welcher Art und Weise – wurde noch kaum erforscht. In der Praxis zeigt sich hingegen, dass innovative Schulen vermehrt personalisierte Lernkonzepte umsetzen und dabei auch digitale Technologien zur Unterstützung beziehen (z.B. Lernplattformen, Lernsoftwares, etc.). Deshalb untersucht diese explorative mixed-methods-Studie die quantitative und qualitative Rolle der digitalen Medien innerhalb personalisierter Lehr- und Lernformen.

#### Design und Methode

Die Stichprobe setzt sich aus N=31 Sekundarschulen zusammen und ist Teil des längsschnittlichen perLen-Projekts, das Schulen mit betont personalisierten Lernkonzepten untersucht. Im Rahmen dieser Studie wurden N=1017 Schüler/innen der 8. Klasse (486 weiblich, 531 männlich) zu ihrem Unterrichtserleben und ihrer Nutzung digitaler Medien befragt. Auf Basis der Fragebogendaten der Schülerinnen und Schüler wurden drei Fallschulen mit intensiver Computernutzung ausgewählt. Pro Schule wurden halbstrukturierte Interviews mit der Schulleitung, 2-3 Lehrpersonen sowie den ICT-Verantwortlichen durchgeführt. Zusätzlich wurden Beobachtungen offener Unterrichtssequenzen durchgeführt, bei denen die Schülerinnen und Schüler selbstständig an Arbeitsaufträgen arbeiteten.

#### Ergebnisse

Bei der quantitativen Datenauswertung der Schüler(innen)befragung zeigte sich eine deutlich intensivere ICT-Nutzung als im Schweizerischen Durchschnitt. 79% der Schülerschaft mit personalisiertem Unterricht gab an, mindestens einmal pro Woche digitale Medien zu nutzen, wohingegen der Durchschnitt der nationalen Referenzstichprobe von ICILS 2013 bei 34% liegt (Petko, Schmid, Pauli, Stebler & Reusser, im Druck).

In den Fallstudien zeigte sich darüber hinaus, dass die ICT-Nutzung vor allem auf Initiative der Schülerinnen und Schüler und weniger auf die Initiative der Lehrpersonen zurückzuführen war. Insbesondere während den offenen Unterrichtsphasen zeigte sich bei allen drei Fallschulen eine hohe Schüler/innen-Wahlfreiheit hinsichtlich wann, wie und für was den Computer genutzt werden kann. Diese Wahlfreiheit wurde von den Schülerinnen und Schülern entsprechend genutzt. An keiner der drei Schulen war ICT ein ausdrückliches Profilelement des Schulleitbildes. Auch andere Faktoren, wie die Nutzung unterschiedlicher Plattformen schienen keinen Einfluss auf die Häufigkeit der ICT-Nutzung zu haben. Die Ergebnisse deuten darauf hin, dass unter der Bedingung eines offenen Unterrichtskontexts weniger die Lehrpersonen als vielmehr die Schülerinnen und Schüler selbst zum Treiber der digitalen Transformation von Unterrichts- und Lernprozessen werden können.

### Navigation im Hypertext als Problemlöseprozess

**Liene Pucite<sup>1</sup>, Johannes Naumann<sup>1</sup>, Beate Eichmann<sup>2</sup>, Frank Goldhammer<sup>2,3</sup>, Samuel Greiff<sup>4</sup>**

<sup>1</sup>Goethe-Universität Frankfurt am Main, Deutschland; <sup>2</sup>Deutsches Institut für internationale Pädagogische Forschung; <sup>3</sup>Zentrum für internationale Vergleichsstudien; <sup>4</sup>Universität Luxemburg; [pucite@em.uni-frankfurt.de](mailto:pucite@em.uni-frankfurt.de)

Das Lesen und Verstehen von Hypertext zählt zu den Schlüsselkompetenzen, die eine erfolgreiche Teilhabe an Wissensgesellschaften des 21. Jahrhunderts ermöglichen. Die wachsende Bedeutung von Hypertexten ist durch die rapide Entwicklung und Nutzung von computerbasierten Technologien bedingt (vgl. Naumann 2015; OECD 2011). Hypertexte sind durch eine nichtlineare Repräsentation der Information gekennzeichnet. Um auf die benötigte Information zugreifen und diese schließlich bearbeiten zu können, müssen Leser/innen eigenständig die bearbeiteten Informationseinheiten auswählen und deren Reihenfolge festlegen („Navigation“, vgl. Bannert 2007; Lawless und Schrader 2008). Dabei sind die Beziehungen zwischen den Informationseinheiten nicht sichtbar. Diese Aspekte bescheren der Informationssuche bzw. dem Navigationsprozess beim Lesen von Hypertexten einen Problemlösecharakter (Brand-Gruwel et al. 2009), da eine Barriere zwischen dem gegebenen Zustand und dem Zielzustand besteht, die überwunden werden muss (Mayer und Wittrock 2006).

Wenn die Navigation als Problemlöseprozess aufgefasst werden kann, lässt sich vermuten, dass Determinanten des Problemlösens Einfluss darauf haben, wie die Personen bei der Informationssuche im Hypertext navigieren. Zudem belegen diverse Studien den prädiagnostischen Effekt der Navigation auf das Hypertextlesen (Hahnel et al. 2016; Naumann 2015). Somit ist ein indirekter Effekt von Determinanten des Problemlösens auf die Leistung beim Hypertextlesen zu erwarten, welcher von der Navigation vermittelt wird.

Zur Überprüfung dieser Hypothese wurde ein Pfadmodell zur Erklärung der Leistung beim Hypertextlesen geschätzt. Für die Analysen wurden Daten (N = 33907, 50% weiblich) aus der PISA-Studie 2012 herangezogen. Dabei fungierten zwei Determinanten des Problemlösens – problemlösebezogene Selbstwirksamkeit („Openness for problem solving“) und Hartnäckigkeit beim Problemlösen („Perseverance“) (OECD 2014) – als exogene Variablen. Aus den Prozessdaten wurden für das Navigationsverhalten zwei Indikatoren, Explorationsverhalten und Aufgabenbearbeitungszeit, extrahiert, die wiederum als Mediatorvariablen dienten. Der Indikator Explorationsverhalten wurde gebildet, indem die Anzahl besuchter Seiten durch die Anzahl im System verfügbarer Seiten geteilt wurde. So bekommen Personen, die alle verfügbaren Seiten im System besucht hatten, den Wert eins zugeordnet, was die höchste Ausprägung der Exploration widerspiegelt. Als Schätzer für die

Personenfähigkeit im Hypertextlesen wurde der Personenparameter Weighted likelihood estimate (WLE) ermittelt, der in den Analysen als Kriterium modelliert wurde.

Im Pfadmodell wurden lineare Effekte der Variablen Selbstwirksamkeit, Hartnäckigkeit, Explorationsverhalten und Aufgabenbearbeitungszeit auf die Variable Leistung beim Lesen digitaler Texte und Effekte der Selbstwirksamkeit bzw. Hartnäckigkeit auf das Explorationsverhalten und der Aufgabenbearbeitungszeit im Modell zugelassen. Bei der Vorhersage der Mediatoren und des Kriteriums wurden Geschlecht, sozioökonomischer Status sowie häusliche Besitztümer in das Modell als Kontrollvariablen aufgenommen.

Die Analysen zeigten, dass sowohl die Selbstwirksamkeit als auch die Hartnäckigkeit kleine, dennoch statistisch signifikante positive indirekte Effekte auf die Leistung beim Hypertextlesen haben, die durch die Navigation vermittelt werden. So scheinen Leser/-innen mit höherer Selbstwirksamkeit und Hartnäckigkeit mehr zu explorieren und dementsprechend länger mit den Hypertextaufgaben zu interagieren, was sich wiederum positiv auf die Leistung beim Hypertextlesen auswirkt. Zudem zeigte sich ein signifikanter direkter Effekt der Selbstwirksamkeit auf die Leistung beim Hypertextlesen. Der prädiktive Effekt der problemlösebezogenen Selbstwirksamkeit zeigte sich auch in den Ergebnissen von OECD (2014) hinsichtlich der Matheleistung. Da sowohl der Problemlösekompetenz (Greiff et al. 2014; Leutner et al. 2012) als auch der Selbstwirksamkeit (Bandura 2012; Zimmerman 2000) eine bedeutende Rolle bezüglich der Lernleistung fächerübergreifend zugeschrieben wird, liegt die Vermutung nahe, dass auch der Effekt der problemlösebezogenen Selbstwirksamkeit domänenübergreifend bestehen bleibt. Dies bedarf jedoch weiterer Forschung.

## **Zusammenhänge der digitalen häuslichen Lernumwelt mit der Motivation von Schülerinnen und Schülern zur Internetrecherche in Lernkontexten**

**Sabrina Wiescholek, Heike M. Buhl**

Universität Paderborn, Deutschland; [sabrina.wiescholek@upb.de](mailto:sabrina.wiescholek@upb.de)

Die Nutzung des Internets zählt aktuell zu den drei wichtigsten Arten der Mediennutzung von Jugendlichen (MpFS, 2016). Dabei wird die Informationskompetenz von Jugendlichen nicht nur von der Selbstwirksamkeitserwartung (SWE) und der Lernmotivation (Hatlevik & Christopherson, 2013), sondern auch von diese wiederum beeinflussenden familiären Hintergrundvariablen bedingt (Wendt et al., 2014). Neben diesen Einflüssen werden des Weiteren immer wieder Geschlechterunterschiede bei der Selbstwirksamkeitserwartung und der Motivation für die Nutzung von Medien und Internet beschrieben (Lorenz et al., 2014). Ein Blick auf die digitale häusliche Lernumwelt und hier insbesondere auf prozessbezogene Determinanten wie beispielweise die Einstellung der Eltern gegenüber dem Internet oder die Art der gemeinsamen Nutzung des Internets, könnte Aufschluss über relevante veränderbare Wirkfaktoren geben.

Ziel der Studie war daher, den Einfluss der digitalen häuslichen Lernumwelt auf die Selbstwirksamkeitserwartung und Motivation zur Internetrecherche bei Kindern und Jugendlichen zu untersuchen. Auf Seite der familiären Hintergrundvariablen standen prozessbezogene Merkmale, wie häusliche Stimulation, Instruktion, Motivation und die Modellfunktion der Eltern im Fokus (Helmke & Schrader, 2010). Bei der Internetrecherche wurde insbesondere die Recherche für schulische Belange betrachtet. Aufgrund der häufig berichteten Geschlechterunterschiede bei Selbstwirksamkeitserwartung und Motivation wird des Weiteren das Geschlecht als Moderator dieser Zusammenhänge betrachtet.

Zur Beantwortung der Fragestellung wurden 640 Schüler/-innen der fünften und sechsten Jahrgangsstufe zu ihrer Selbstwirksamkeitserwartung (in Anlehnung an Schwarzer & Jerusalem, 1999), zu ihrer Motivation (in Anlehnung an Schiefele & Schaffner, 2016) und zu ihrer Wahrnehmung der digitalen häuslichen Lernumwelt schriftlich befragt. Die Hypothesen wurden mittels Strukturgleichungsanalyse überprüft.

Besonders die gemeinsame Nutzung des Internets zu nützlichen Zwecken stand in positivem Zusammenhang mit der Motivation der SuS zur Internetrecherche. Im Sinne der Modellfunktion der Eltern konnte des Weiteren ein positiver Zusammenhang der von Kindern wahrgenommenen elterlichen Nutzung des Internets für die Arbeit mit der Leistungsmotivation der Kinder gefunden werden.

In Bezug auf die Selbstwirksamkeitserwartung bei der Internetrecherche erwiesen sich ebenfalls die gemeinsame Nutzung sowie die Einstellung der Eltern gegenüber dem Internet als Determinanten.

Zudem zeigte sich das Geschlecht als signifikanter Moderator des Zusammenhangs zwischen der Selbstwirksamkeitserwartung und den Tipps von Eltern bei der Internetrecherche. Während bei Jungen kein Zusammenhang zwischen den Tipps der Eltern und der Selbstwirksamkeitserwartung zu finden war, konnte ein negativer Zusammenhang bei den Mädchen gefunden werden. Je häufiger Mädchen demnach Tipps von ihren Eltern bei der Internetrecherche wahrnahmen, umso geringer war ihre Selbstwirksamkeitserwartung.

Vor dem Hintergrund dieser Ergebnisse zeigt sich der Bedarf einer differenzierten Betrachtung der digitalen häuslichen Lernumwelt insbesondere mit Blick auf das Geschlecht. Werden im Rahmen der zunehmenden Digitalisierung des Bildungssystems immer mehr Angebote und Fördermaßnahmen entwickelt, gilt es im Rahmen der Professionalisierung der digitalen Bildung auch das Elternhaus zu berücksichtigen. Es wird diskutiert, wie das Elternhaus einbezogen werden kann.



## B08

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H103

### Regulation von Motivations- und Wissensproblemen in kooperativen Lernsituationen

**Nadine Melzner, Martin Greisel, Markus Dresel, Ingo Kollar**

Universität Augsburg, Deutschland; [nadine.melzner@phil.uni-augsburg.de](mailto:nadine.melzner@phil.uni-augsburg.de)

#### Theoretischer Hintergrund

Kooperativem Lernen wird erhebliches Potenzial zur Förderung von Lernprozessen und -ergebnissen zugeschrieben (Weinberger, Stegmann, & Fischer, 2010). Allerdings werden diese Potenziale insbesondere in selbstorganisierten Lerngruppen häufig nicht gehoben, etwa wenn innerhalb der Gruppe ungünstige Motivationslagen oder ein geringer Wissensstand gegenüber den Lerninhalten vorliegen (Järvelä & Hadwin, 2013). Bisher ist kaum bekannt, wie Lernende in selbstorganisierten Lerngruppen bei der Regulation derartiger unterschiedlicher Probleme vorgehen. Die zur Minimierung entsprechender Probleme notwendigen Regulationsprozesse können zum einen über eine Analyse der Art der in Gruppen eingesetzten Regulationsstrategien (u.a. kognitive Strategien und Motivationsregulationsstrategien; Boekaerts, 1997) beschrieben werden. Zum anderen kann analysiert werden, auf welchen sozialen Ebenen diese Strategien eingesetzt werden. Järvelä und Hadwin (2013) unterscheiden hierbei zwischen der Self-Ebene, auf der einzelne Gruppenmitglieder ihren eigenen Lernprozess regulieren, der Co-Ebene, auf der einzelne Gruppenmitglieder die Lernprozesse anderer Gruppenmitglieder regulieren, und der Shared-Ebene, auf der Regulationsprozesse innerhalb der Gesamtgruppe ausgehandelt werden.

#### Fragestellungen

Der Beitrag untersucht die Effekte des Vorliegens von Motivations- und/oder Wissensproblemen in selbstorganisierten Lerngruppen (1) auf die Gesamtzahl an Regulationsstrategien, die Studierende berichten, (2) auf das Ausmaß der Anwendung unterschiedlicher Arten von Regulationsstrategien (Motivationsregulationsstrategien und kognitiven Lernstrategien), sowie (3) auf das Ausmaß, zu dem diese Strategien auf unterschiedlichen sozialen Ebenen (Self- vs. Co- vs. Shared-Ebene) gezeigt werden.

#### Methode

Studierende der Erziehungswissenschaft (N=86) sollten sich beim Lesen vierer Fallvignetten vorstellen, sich als Teil einer selbstorganisierten Lerngruppe auf eine Prüfung vorzubereiten. In einem ausbalancierten 2x2-faktoriellen Within-Design wurde variiert, ob ein Motivationsproblem, ein Wissensproblem, beide Probleme, oder kein Problem vorliegt. Für jede Vignette wurden die Versuchspersonen gefragt, welche Regulationsstrategien sie in der jeweiligen Situation zur Regulation des eigenen Lernens („Self“-Ebene), des Lernens anderer Gruppenmitglieder („Co“-Ebene), und des Lernens der Gesamtgruppe („Shared“-Ebene) anwenden würden. Die Kategorisierung der von den Teilnehmer(innen) in einem offenen Format genannten Strategien erfolgte auf Basis eines Schemas zur Klassifizierung von Lernstrategien nach Artelt (2000; Cohens Kappa=.77).

#### Ergebnisse

Mit Blick auf Frage 1 zeigten sich in der Bedingung „mit Motivationsproblem/ohne Wissensproblem“ deutlich weniger Strategienennungen als in den übrigen Bedingungen. Der Interaktionseffekt „Wissensproblem\*Motivationsproblem“ wurde folglich signifikant,  $F(1,85) = 6.904$ ,  $p < .05$ ,  $\eta^2 = .075$ .

Bezüglich Frage 2 ergab sich ein signifikanter Interaktionseffekt „Motivationsproblem\*Strategietyp“,  $F(1,85) = 49.675$ ,  $p < .01$ ,  $\eta^2 = .369$ . Dieser zeigte an, dass beim Vorliegen eines Motivationsproblems deutlich weniger kognitive Lernstrategien berichtet wurden als wenn kein Motivationsproblem vorlag. Gleichzeitig führte das Vorliegen eines Motivationsproblems aber zu einer häufigeren Nennung von Motivationsregulationsstrategien als wenn kein Motivationsproblem vorlag. Der Interaktionseffekt „Wissensproblem\*Strategietyp“ ( $F(1,85) = 1.084$ ,  $p > .05$ ) wurde nicht signifikant.

Bezüglich Frage 3 zeigte sich ein Haupteffekt des Faktors „Soziale Ebene“ auf die Gesamtzahl genannter Strategien,  $F(2,170) = 67.089$ ,  $p < .01$ ,  $\eta^2 = .441$ . Auf Self-Ebene wurde signifikant häufiger als auf Co- ( $p = .00$ ) und auf Shared-Ebene ( $p = .00$ ) reguliert. Auf Shared-Ebene wurden zudem signifikant mehr Strategien genannt als auf Co-Ebene ( $p = .00$ ).

#### Diskussion

Die Ergebnisse zeigen, dass Studierende in selbstorganisierten Lerngruppen auf verschiedene Regulationsprobleme unterschiedlich reagieren. Insbesondere scheinen Motivationsprobleme aus Studierendensicht spezifischer reguliert werden zu müssen als Wissensprobleme. Letztere können aus Studierendensicht offenbar sowohl mit kognitiven als auch mit Motivationsregulationsstrategien angegangen werden, was vor dem Hintergrund von Befunden, dass Wissenserwerb positiv mit motivationalen Lernvoraussetzungen wie Interesse korreliert ist (Schiefele & Schreyer, 1994), durchaus adäquat erscheint. Dass Studierende bei der Regulation insbesondere das eigene Lernen fokussieren und nur selten das Lernen anderer Gruppenmitglieder regulieren bzw. gemeinschaftlich entsprechende Regulationsprozesse aushandeln, könnte darauf hindeuten, dass sie von den Potenzialen des kollaborativen Lernens nicht überzeugt sind (Hunter & Leahey, 2008) und sich – auch angesichts der bevorstehenden Prüfung – zunächst auf den eigenen Lernprozess konzentrieren, um ihren persönlichen Erfolg sicherzustellen.

### With a Little Help From my Peers – Autonomieunterstützung beim Experimentieren im Cross-Age-Tutoring

**Thomas Schlake<sup>1</sup>, Heiko Krabbe<sup>2</sup>, Hendrik Härtig<sup>1</sup>, Maria Opfermann<sup>2</sup>, Hans Fischer<sup>1</sup>**

<sup>1</sup>Universität Duisburg-Essen, Deutschland; <sup>2</sup>Ruhr-Universität Bochum, Deutschland; [thomas.schlake@uni-due.de](mailto:thomas.schlake@uni-due.de)

Experimente mit Lehrer-Schüler-Interaktionen gehören zum naturwissenschaftlichen Unterricht, wobei zunehmend auch andere Lernumgebungen und Organisationsformen an Relevanz gewinnen, wie z.B. das Cross-Age-Tutoring (Korner, 2015; Topping, 2005). Dabei fungieren ältere Schülerinnen und Schüler als Tutoren für jüngere Schülerinnen und Schüler (Tutees) und unterstützen diese z.B. beim eigenständigen Experimentieren.

Die Forschung zum Cross-Age-Tutoring zeigt, dass sich durch solche Lernformen die schulische Leistung sowohl der Tutoren als auch der Tutees verbessern kann (Robinson, Schofield & Steers-Wentzell, 2005) und dass die intrinsische Motivation der Tutoren steigt (Berger, Müller & Hånze, 2017). Ob positive Effekte auftreten, hängt jedoch entscheidend von der Art des Tutorings ab, wobei hier insbesondere das Ausmaß der Autonomieunterstützung ausschlaggebend zu sein scheint (Berger et al., 2017). Autonomieunterstützung beinhaltet das Ermöglichen eines eigenen Arbeitstempos, Handlungs- und Wahlfreiheit,

Anerkennung der Gefühle und Perspektiven der Lernenden sowie ein Vermeiden kontrollierender Sprache durch die Lehrenden (Kaur, Awang-Hashim & Noman, 2016; Reeve, 2009).

Während die positiven Auswirkungen autonomieunterstützenden Verhaltens bislang vorwiegend für Lehrer-Schüler- (Reeve, 2009), Eltern-Kind- (Grolnick & Apostoleris, 2002) oder Vorgesetzter-Angestellter- (Gagné & Deci, 2005) Interaktionen untersucht wurden, gibt es nur wenige Studien zu Cross-Age-Tutoring-Settings bei Schülern. An dieser Lücke setzt unser Projekt an, um professionelles Handeln der Tutoren durch die differenzierte Beschreibung von Lern- und Entwicklungsprozessen der Tutoren besser anleiten zu können. Dabei sollen insbesondere die Fragestellungen fokussiert werden, ob (1) eine Intervention die Neigung zu Autonomieunterstützung von Tutoren steigern kann und inwiefern dies von individuellen Charakteristika der Tutoren abhängt, ob (2) diese individuellen Charakteristika auch das während des Experimentierens tatsächlich gezeigte Tutor-Verhalten vorhersagen können und ob (3) das Kompetenzerleben und das Interesse der Tutees in Abhängigkeit ihrer Wahrnehmung der Autonomieunterstützung durch ihre Tutoren steigen.

Für unsere Studie haben wir eine Lernumgebung entwickelt, in der Acht- und Neuntklässler als Tutoren dafür trainiert werden, Schülerinnen und Schüler der fünften und sechsten Klasse beim physikalischen Experimentieren zu unterstützen. Die Intervention fokussiert auf die Förderung autonomieunterstützenden Verhaltens. Zwei Wochen später betreuen die so trainierten Tutoren jeweils einen Tutee beim Durchführen eines Experiments. Diese Sitzungen werden videobasiert analysiert. Zudem werden vor und nach der Intervention mithilfe von Fragebögen die Neigung der Tutoren zu Autonomieunterstützung, ihre Kausalitätsorientierung, sowie postinterventional zusätzlich die intrinsische Motivation und weitere individuelle Variablen erfasst. Ebenso werden auf Seiten der Tutees nach dem Experimentieren die wahrgenommene Autonomieunterstützung, ihr Kompetenzerleben und ihr Interesse erhoben. Alle verwendeten Skalen zeigen akzeptable bis gute interne Konsistenzen ( $.73 \leq \alpha \leq .88$ ).

Aktuell liegen Daten von 75 Tutoren (41w, 34m) mit einem Durchschnittsalter von 13,6 Jahren und von 74 Tutees (44w, 30m) mit einem Durchschnittsalter von 10,6 Jahren vor. Die Gesamtstichprobe wird jeweils 120 Tutoren und Tutees umfassen.

Für die vorläufige Stichprobe zeigt ein T-Test für verbundene Stichproben einen signifikanten Anstieg der Neigung der Tutoren zu Autonomieunterstützung ( $p < .001$ ;  $d = 1.83$ ). Regressionsanalysen weisen zudem darauf hin, dass dies insbesondere für Tutoren mit einer höheren autonomen Kausalitätsorientierung der Fall ist ( $p < .05$ ;  $\beta = .245$ ). Auf Seiten der Tutees zeigt sich, dass ihre wahrgenommene Autonomieunterstützung sowohl ihr Kompetenzerleben ( $p < .001$ ;  $\beta = .414$ ) als auch ihr Interesse ( $p < .001$ ;  $\beta = .376$ ) vorhersagt. Alter oder Geschlecht der Tutoren und Tutees sagen weder das Kompetenzerleben noch das Interesse der Tutees voraus.

Bis zur Tagung sind die Experimentalsitzungen videobasiert ausgewertet. Es können dann Ergebnisse zu allen drei Forschungsfragen vorgestellt werden. Das erfolgreiche Training kann genutzt werden, um Cross-Age-Tutoring in Schulen neu zu implementieren oder bereits vorhandene Tutor-Trainings zu ergänzen. Die Ergebnisse der Untersuchung können außerdem dazu dienen, Interventionen bezüglich der Autonomieunterstützung in schulischen Lernumgebungen weiterzuentwickeln. Damit einhergehende bessere Leistungen von Schülerinnen und Schülern können ein wertvoller Beitrag zum naturwissenschaftlichen Unterricht sein.

## **Individuelle Lernunterstützung in Schulen mit personalisierten Lernkonzepten - Videogestützte Fallstudien im Fach Deutsch auf der Sekundarstufe I**

**Zoe Dellios**

Universität Zürich, Schweiz; [zdellios@ife.uzh.ch](mailto:zdellios@ife.uzh.ch)

Theoretischer Hintergrund:

Um im Rahmen der vielfältigen Herausforderungen mit heterogenen Lerngruppen sowohl die curricularen Grundanforderungen und gleichzeitig die individuelle Förderung leisten zu können, setzt eine wachsende Zahl von Schulen auf innovative, gemeinschaftsorientierte wie individualisierende Schul- und Unterrichtskonzepte. Die Abkehr von traditionellen Unterrichtsprinzipien (Jahrgangsklassen, Raumorganisation, Studentakt etc.) bewirkt neue Formen der Schul- und Unterrichtsorganisation darunter ‚Personalisiertes Lernen‘. Dieses ist Gegenstand des Forschungsprojekts perLen des Instituts für Erziehungswissenschaft an der Universität Zürich, das personalisierte Lernkonzepte in heterogenen Lerngruppen untersucht.

In personalisierten, schüleradaptierten und individualisierten Lernumgebungen schlägt sich professionelles Handeln sowohl in der Gestaltung der Lernumgebung wie auch in der Interaktion zwischen Lernenden und Lehrpersonen nieder. Während die Lernumgebung mit einem kognitiven Anregungsgehalt die Basis bildet, übt die Qualität der Unterrichtskommunikation für das produktive, verständnisorientierte Lernen im Sinn eines soziokulturellen und konstruktivistischen Verständnisses von Lehr- und Lernprozessen eine zentrale Rolle aus (Stebler, Pauli & Reusser, 2017). Neben den intendierten passgenauen Lernangeboten, kommt eine entscheidende Bedeutung für die Förderung des Verständnisses und den Aufbau flexiblen Wissens der Ko-Konstruktion in der Unterstützungssituation zu (Greeno, 2006).

Fragestellung

Das Erkenntnisinteresse des empirisch-analytischen Forschungsvorhabens richtet sich auf die Lern- und Aufgabenkultur in Schulen mit personalisierten Lernkonzepten im Fachbereich Deutsch und dem professionellen Unterstützungshandeln der Lehrperson.

Leitende Fragestellungen nehmen Bezug auf:

- Die Perspektive der Interaktion: Wie ist die Beziehungs- und Unterstützungskultur in personalisierten Aufgabenstellungen und unterscheiden sich die Interaktionsmodelle in Phasen des geführten und offenen Unterrichts?
- Gibt es einen Zusammenhang zwischen der Qualität der Aufgabenstellung und Adaptivität seitens der Lehrpersonen und deren Lernwirksamkeit seitens der Lernenden?

Methode

Die videogestützten Fallstudien perLA (personalisierte Lern- und Aufgabenkultur) entstanden im Rahmen der perLen-Längsschnittstudie (personalisierte Lernkonzepte in heterogenen Lerngruppen) des Instituts für Erziehungswissenschaften der Universität Zürich. Zur Beantwortung der Fragestellungen dient eine mehrperspektivische Zugangsweise, die unterschiedliche Erhebungsinstrumente kombiniert (Kameraaufnahmen mit vier Perspektiven, Beobachtungsbogen, Aufgabensammlungen und -lösungen, Textproduktion, Fragebogen). Der Fokus richtet sich auf drei Schulen der Sekundarstufe I mit personalisiertem Unterricht im Fach Deutsch. Um eine Varianz zwischen den Schulen, den Unterrichtsarrangements und den Lernenden

generieren zu können, wurden die drei Schulen an drei Tagen während einer Phase von drei Wochen gefilmt. Im Fokus stehen geführte Unterrichtssequenzen, in denen die Lehrenden die zeitliche und inhaltliche Struktur vorgeben und offene Lernarrangements, in denen die Schülerinnen und Schüler in einer Lerngemeinschaft oder für sich allein und selbstbestimmt das Lernen steuern. Die Analyse erfolgte in drei Schritten mittels (1) eines Kriterienrasters zur Analyse des Lernangebots, (2) eines Kodesystems (time-sampling) und (3) eines Raster zur Ereigniserfassung (event-sampling) zur Videoanalyse.

Das multimethodisch-technologieunterstützte Vorgehen kombiniert für die Datenanalyse verschiedene Erhebungs- und Auswertungsverfahren (Dokumentenanalyse, Fragebogen, videogestützte Unterrichtssequenzen, Textproduktionen). Niedrig inferente Beurteilungssysteme zielen auf das präzise Erfassen der Sicht- und Oberflächenstruktur ab, hoch inferente Beurteilungssysteme intendieren das Erfassen komplexer Unterrichtsmerkmale (Rakoczy & Pauli, 2006). Die beiden Verfahren ergänzen sich zur Analyse, Beschreibung und Einschätzung des Datenmaterials (Lotz et al., 2013).

Ergebnisse

Erste Ergebnisse des empirische-analytischen Forschungsvorhabens verweisen auf einen lern- und kulturspezifischen Unterschied im Angebots-Nutzungs-Verhalten zwischen den untersuchten Schulen und interindividuelle Unterschiede zwischen den Lernenden bei der Art und der Qualität der Lernzeitnutzung. Die Befunde unterstützen, dass während der selbständigen Lernerarbeit die Anregung kognitiver Aktivitäten durch eine adaptive Hilfestellung und instruktionale Erklärungen, die Lernenden zu einem nächsten Lösungsschritt befähigen.

## **Der Einfluss von außerschulischem Englischkontakt auf die Sprachkompetenz junger Lernender in Deutschland und der Schweiz**

**Maleika Krüger**

PH/FHNW, Schweiz; [maleika.krueger@fhnw.ch](mailto:maleika.krueger@fhnw.ch)

Obwohl das Prinzip der informellen Lernprozesse in der Fremdsprachenforschung ein anerkannter Bereich ist, fehlt es der deutschsprachigen Forschungslandschaft bisher, gerade im Hinblick auf die neuen Medien und des daraus entstehenden Lernpotentials, an empirischer Forschung. Dies verwundert umso mehr, als Studien aus anderen Ländern bereits die positiven Effekte des außerschulischen medialen Kontaktes auf die Sprachkompetenz nachweisen konnten (u.a. Black, 2005, 2009; Gee, 2007; Kuppens, 2010; Olsson, 2011, 2016; Pearson, 2004; Piirainen-Marsh & Tainio, 2009; Sundqvist, 2009; Sundqvist & Sylvén, 2012a, 2014; Sylvén, 2006; Sylvén & Sundqvist, 2012, 2015; Thorne & Black, 2007; Thorne, Black, & Sykes, 2009). Die Ergebnisse der Studien zeigen, dass Lernende heutzutage zunehmend außerhalb der offiziellen Bildungseinrichtungen auf spielerische und informelle Weise in Kontakt mit der englischen Sprache kommen und ihre sozialen und interkulturellen Fähigkeiten in einem anglophonen Umfeld erproben und ausbauen. Die Freiwilligkeit dieser Tätigkeiten und der hohe Unterhaltungswert stellen dabei einen starken intrinsischen Motivationsfaktor dar.

Das Projekt MEWS (SNF-DFG Lead Agency Projekt, 2016-2019) bietet zum ersten Mal die Möglichkeit das Ausmaß an außerschulischen medialen Englischaktivitäten und deren Einfluss auf die Englischkompetenz im Allgemeinen und die Schreibkompetenz im Speziellen für D und CH im Detail zu betrachten. Die Studie bedient sich dabei des Mixed-Method-Ansatzes, in dem sowohl qualitative, wie auch quantitative Daten eingesetzt werden.

Zu diesem Zweck wurden zu zwei Erhebungszeitpunkten die Lese-, Hör- und Schreibkompetenz von N = 3000 TeilnehmerInnen zwei Jahre vor dem Abitur/ der Matura erfasst sowie ausführliche Informationen zum sozialen Hintergrund und den außerschulischen medialen Englischaktivitäten im Bereich Musik, Lesen, Fernsehen, Computerspiele und Internetnutzung (schriftliche Inhalte und Videos) erhoben. Letztere sind dabei nicht, wie so oft, auf die reine Häufigkeitsbeschreibung reduziert, sondern erfassen detaillierte Informationen zu den genutzten Plattformen, Genres und Verbreitungswegen sowie zu passiven und aktiven Sprachproduktionen der Jugendlichen. Der eingesetzte Fragebogen wurde im Projektteam eigens für diese Umfrage entwickelt und in mehreren Vorstufen quantitativ und qualitativ pilotiert.

In den nächsten Monaten werden zusätzlich qualitative Interviews mit einer ausgewählten Teilstichprobe der Lernenden erfolgen, um die mediale Lebenswelt der Jugendlichen noch genauer zu erfassen.

Erste Analysen der Daten zeigen, dass die Jugendlichen in den beiden Ländern vor allem durch Musik (84%), das Lesen von Inhalten im Internet (51,1%) und durch das Ansehen von Internetvideos (47,3%) täglich oder fast täglich mit Englisch in Kontakt kommen. Im Hinblick auf englische Filme/Serien und englischen Computerspiele geben immerhin noch 18% bzw. 19,5% der Jugendlichen an, dies täglich oder fast täglich zu tun. Deutlich seltener werden dagegen Bücher auf Englisch gelesen (4,8%).

Zum Zeitpunkt der Einreichung standen die Ergebnisse der Leistungstests noch nicht zur Verfügung, es ist aber anzunehmen, dass sich die positiven Effekte internationaler Studien auch in der vorliegenden Studie zeigen. Abzuwarten bleibt, ob alle Medienarten gleich vorteilhaft sind, ob alle Lernenden gleichermaßen von diesen Lerneffekten profitieren und ob das Schreiben argumentativer Texte, auf Grund der spezifischen Anforderungen des Genres, in einem geringeren Maße beeinflusst wird. Der Vortrag wird die ersten Ergebnisse zur Beantwortung dieser Forschungsfragen mit Hilfe eines Strukturgleichungsmodells präsentieren.

Die hier vorgestellte Studie liefert mit ihrer Fragestellung einen wichtigen Beitrag zum Verständnis des Zusammenhangs zwischen schulischen und außerschulischen Faktoren. Kommen Jugendliche auch in Deutschland und der Schweiz verstärkt außerhalb der Schule in Kontakt mit der englischen Sprache, ist es zentral, die daraus resultierenden Lerneffekte zu verstehen und sich eventueller sozialer Einflussfaktoren gewahr zu werden. Nur so ist es möglich die Lernenden auch innerhalb des Klassenzimmers individuell zu fördern und den Unterricht auf die sich verändernden Lebenswelten der Jugendlichen abzustimmen. Die Resultate der Studie stellen daher wichtige Anknüpfungspunkte für das Einbinden authentischer Materialien in den Unterricht dar und können Lehrpersonen als Ansatzpunkt dienen, ihre Lernenden zu motivieren und die Sprache aus ihrem reinen Schulzusammenhang zu lösen.

## B09

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H223

### Unterrichtsqualität – ein Frage der Perspektive

Esther Brunner

Pädagogische Hochschule Thurgau, Schweiz; [esther.brunner@phtg.ch](mailto:esther.brunner@phtg.ch)

#### Theoretischer Hintergrund

Mögliche Antworten auf die Frage, wodurch guter Unterricht gekennzeichnet ist, hängen nicht nur davon ab, mit welchem Ziel Unterrichtsqualität untersucht und wie diese theoretisch konzeptualisiert und operationalisiert wird, sondern auch davon, aus welchem Blickwinkel man Unterricht analysiert. Unterrichtsqualität kann gewissermaßen durch unterschiedliche Linsen (vgl. Givvin et al., 2005) betrachtet werden, beispielsweise aus einer Weitwinkelperspektive mit Blick auf das Ganze oder mit einem Fokus auf Details. Im ersten Fall werden andere Qualitätsmerkmale im Vordergrund stehen als im zweiten Fall. Gleiches gilt, wenn Unterrichtsqualität primär entweder mit Blick auf Sichtstruktur im Gegensatz zur Tiefenstruktur (vgl. Reusser, 2005) betrachtet wird. Im ersten Fall wird auf unmittelbar beobachtbare Merkmale des Unterrichts fokussiert, im zweiten Fall steht hingegen die nicht direkt beobachtbare Lehr-/Lernprozessebene im Zentrum, welche durch die Sichtstruktur bestimmt wird (Kunter & Ewald, 2016). Untersucht man die Qualität von Fachunterricht, stellt sich zudem die Frage der Berücksichtigung genuin fachdidaktischer Merkmale und damit die Frage der Domänenspezifität von Unterrichtsqualität.

Die bislang vorliegenden Analyseinstrumente zur Einschätzung von Unterrichtsqualität (zusammenfassend Schlesinger & Jentsch, 2016) unterschieden sich deshalb nicht nur im gewählten Fokus bzw. in der Gewichtung der beiden Ebenen der Sicht- und Tiefenstruktur, sondern auch bezüglich ihrer Domänenspezifität. Darüber hinaus stellt sich seit der Ausdifferenzierung der sogenannten Basisdimensionen (Klieme, Lipowsky, Rakoczy, & Ratzka, 2006) von Unterrichtsqualität die Frage, inwiefern Qualitätsaspekte einer Hierarchisierung unterliegen.

#### Fragestellung

Auf diesem grob skizzierten Hintergrund interessieren folgende Fragestellungen: 1) Inwiefern wird derselbe Unterricht mit unterschiedlichen Instrumenten ähnlich eingeschätzt? 2) Wie lässt sich Unterrichtsqualität theoretisch konzeptualisieren, wenn sowohl fachspezifische wie auch allgemeine Aspekte in den Blick genommen werden sollen?

#### Methode

Frage 1 wird im Rahmen einer explorativen Studie am Beispiel einer videografierten Mathematikstunde aus dem 6. Schuljahr untersucht. Die Stunde wird von einer Lehramtsstudentin gehalten, der während der Bearbeitung gravierende fachliche Fehler unterlaufen, die nicht korrigiert werden. Die Stunde wird in einem hochinferenten Rating mit drei unterschiedlichen Instrumenten von zwei Rating-Teams konsensbasiert eingeschätzt. Die Instrumente unterscheiden sich nicht nur hinsichtlich ihrer theoretischen Perspektive und der Auswahl der geprüften Qualitätsaspekte, sondern auch hinsichtlich ihres Einbezugs von eher fachspezifischen Merkmalen. Eingesetzt werden ein älteres fachunspezifisches Instrument (Ditton & Merz, 2000), das durch fachspezifische Qualitätsmerkmale ergänzt wurde (Brunner, Kreis, Staub, Schoy-Lutz, & Kosorok Labhart, 2014), ein Instrument, das die drei Basisdimensionen (Klieme et al., 2006) von Unterrichtsqualität fokussiert (Rakoczy & Pauli, 2006) sowie ein genuin fachdidaktisches Instrument (Schoenfeld, Floden, the Algebra Teaching and Mathematics Assessment Project Berkeley, & Lansing, 2014). Die zweite Fragestellung wird auf der Basis der Ergebnisse theoretisch bearbeitet.

#### Ergebnisse

Die Ergebnisse zeigen, dass die videografierte Mathematikstunde entlang der drei Instrumente deutlich unterschiedlich eingeschätzt wird. Mit Instrument 1 wird die Qualität der Mathematikstunde bezüglich allgemeiner Merkmale als sehr gut (97,22 %) eingeschätzt, während die fachspezifischen Merkmale (76,22 %) deutlich niedriger eingeschätzt werden, aber immer noch im obersten Quartil liegen. Die fachliche Korrektheit wird demgegenüber mit dem Wert 1 als „extrem niedrig“ eingeschätzt, kann aber durch die Vielzahl anderer fachspezifischer Merkmale gemittelt kompensiert werden.

Mit Instrument 2 bleibt die extrem niedrige fachliche Korrektheit der Stunde unberücksichtigt. Aspekte der Klassenführung werden als hervorragend eingeschätzt (100 %), die Motivierungsqualität und Schülerorientierung mittel (56,33 %), die Unterstützung beim Wissensaufbau hingegen niedrig (33,33 %).

Mit Instrument 3 werden die fachlichen Schwierigkeiten in den beiden fachspezifischen Dimensionen erfasst (je 25 %). In den drei eher allgemeinen, aber fachlich operationalisierten Dimensionen liegen die Werte im mittleren (je 50 %).

#### Diskussion

Die Ergebnisse verdeutlichen zum einen die Notwendigkeit eines interdisziplinären Dialogs und werfen zum anderen Fragen auf, u. a. bezüglich 1) der Hierarchisierung von Qualitätsdimensionen, 2) dem Stellenwert der fachlichen Korrektheit sowie 3) der grundsätzlichen Kompensierbarkeit von Nichtkompensierbarem. Dies wird im Vortrag anhand eines Modells zu fachlicher Unterrichtsqualität diskutiert.

### Wie viele Klassen sind nötig, um Unterrichtsqualität zu messen? Eine Monte Carlo-Simulationsstudie zur Performanz von frequentistischen und bayesschen Verfahren in kleinen Stichproben.

Christoph Helm

Johannes Kepler Universität Linz, Deutschland; [christoph.helm@jku.at](mailto:christoph.helm@jku.at)

#### Fragestellung

Neben der Kompetenzmessung stellt die Erfassung von Unterrichtsqualität eine zentrale Herausforderung für Bildungsforscher/innen dar. In jüngerer Vergangenheit erfuhr die Praxis statistischer Modellierung von Kontextmerkmalen durch sogenannte Doubly Latent Modelle (DLM, e.g., Lüdtke et al., 2008; Marsh et al., 2009) einen starken Aufschwung. DLM erlauben sowohl für den Stichproben- als auch Messfehler zu kontrollieren und ermöglichen darüber hinaus eine theoriekonforme Modellierung von Unterrichtsmerkmalen auf Klassenebene. Diese hierarchischen, konfirmatorischen Messmodelle setzen allerdings eine hohe Anzahl befragter Schulklassen voraus (z.B. Maas & Hox, 2005). Eine ausreichend große Stichprobe ist deshalb erforderlich, da traditionelle statistische Verfahren (z.B. Maximum Likelihood-Ansatz) auf der Annahme asymptotisch verteilter Statistiken basieren. Ausreichend große Stichproben sind damit eine notwendige Voraussetzung für unverzerrte Ergebnisse (z.B. Stegmüller, 2013). Sind die Stichproben zu klein, können Parameterschätzer, Konfidenzintervalle und damit Signifikanztestungen verzerrt sein. Bayessche Verfahren dagegen beruhen nicht auf der

Annahme asymptotischer Verteilung. Sie verwenden für jeden interessierenden Parameter existierendes Vorwissen (prior distribution), das mit den erhobenen Daten (likelihood distribution) kombiniert wird, um die sogenannte posterior distribution (the posterior) zu bestimmen. Deren Erwartungswerte wiederum stellen die wahrscheinlichsten Parameterschätzungen dar. Daher ist zu erwarten, dass bayessche Verfahren in kleinen Stichproben insbesondere dann verlässliche Ergebnisse produzieren, wenn informative priors vorliegen, für die eine umfangreiche und valide Vorwissensbasis verfügbar ist (Depaoli & Van de Schoot, 2017, p. 240).

Vor diesem methodologischen Hintergrund stellt sich die Frage, ob informative priors in sehr kleinen Stichproben helfen, Probleme, die in traditionellen Verfahren entstehen (z.B. Parameterbias), zu vermeiden.

Zur Frage der Stichprobengröße existieren zwar Simulationsstudien. Allerdings sind diese nur bedingt zur Beantwortung der gestellten Frage geeignet, da sie entweder Modelle mit manifesten Variablen spezifizieren (Stegmueller, 2013; Maas & Hox, 2005) oder keine Mehrebenenstruktur mitmodellieren (Van de Schoot et al., 2015; Lee and Song, 2004). Am ehesten eignet sich die Studie von Zitzmann et al. (2016), allerdings ist dort die kleinste untersuchte Clusterzahl 25 und es wurden vague priors spezifiziert, sodass eine Simulationsstudie mit kleineren Klassenzahlen und informativen priors notwendig und gerechtfertigt erscheint.

Methode

Zur Beantwortung der Forschungsfrage wird eine Simulationsstudie mit folgendem 5 x 4 x 3 x 3 Design (180 Bedingungen) durchgeführt:

- (1) Anzahl der Schulklassen (10, 15, 20, 25, 30)
- (2) Intraklassenkorrelation (ICC(1): .5, .10, .15, .20)
- (3) Reliabilität auf Klassenebene (Anzahl der Schüler/innen in einer Klasse, k: 15, 20, 25)
- (4) Statistisches Verfahren (ML, Bayes non-informative, Bayes informative)

Die Hyperparameter der informativen priors wurden aus den TIMSS-Daten 1999-2005 abgeleitet. Dazu wurden DLM für alle verfügbaren Länder (in Summe 394) geschätzt. Die daraus resultierenden Verteilungen der einzelnen Parameter bilden die priors. Für die L2-Varianzen der latenten Variable sowie den Residuen des DLM wurden Hyperparameters für inverse gamma priors abgeleitet; für die L2-Faktorkladungen und Indikatorenintercepts wurden normal priors spezifiziert. Für die Level 1 Parameter (die aufgrund ausreichend hoher L1-Beobachtungen nicht interessieren) als auch für die non-informativen bayesschen Modelle wurden die default prior-Einstellungen von Mplus 7 (Asparouhov & Muthén, 2010) übernommen.

Die population values für die Datengeneration wurden aus dem TIMSS 2015 Datensatz für die Schweiz abgeleitet.

Ergebnisse

Ergebnisse eines ersten Simulationsdurchgangs zeigen, dass bzgl. der L2-Parameter (insbesondere der Varianzen) die bayesschen Verfahren in Hinblick auf die Convergence Rate (CR) und die Power deutlich besser als die ML-Schätzung „performen“. Was den relativen Parameterbias betrifft, zeigt die ML-Schätzung teilweise Vorteile, insbesondere, wenn der ICC(1) niedrig ist. Dies gilt tendenziell auch für die Confidence Intervall Coverage; nur bei hohem ICC(1) zeigen sich die bayesschen Verfahren vorteilhafter. Der ICC(2), d.h. die Reliabilität der DLM, scheint keinen Einfluss auf die Simulationsergebnisse zu haben.

Die Ergebnisse werden vor dem Hintergrund der Theorie und der Vor- und Nachteile bayesscher Analysen diskutiert.

## **Dimensionen der Unterrichtsqualität - Entwicklung und Validierung eines neuen Ratingsbogens**

**Max Nachbauer<sup>1</sup>, Holger Gärtner<sup>2</sup>, Christopher Kellermann<sup>1</sup>, Felicitas Thiel<sup>1</sup>**

<sup>1</sup>Freie Universität Berlin, Deutschland; <sup>2</sup>Institut für Schulqualität der Länder Berlin und Brandenburg, Deutschland;  
[max.nachbauer@fu-berlin.de](mailto:max.nachbauer@fu-berlin.de)

Theoretischer Hintergrund

Innerhalb der Lernumwelt Schule wird im Allgemeinen der Unterricht als der bedeutendste Einflussfaktor auf den Lernerfolg der Schüler und Schülerinnen angesehen (Muijs & Reynolds, 2001). Von besonderer Relevanz ist hierbei die Unterrichtsqualität (Muijs et al., 2014). Insbesondere in der deutschsprachigen Unterrichtsforschung finden sich mehrere Modelle, welche Unterrichtsqualität anhand von drei Dimensionen beschreiben (z.B. Fauth et al., 2014; Klieme et al., 2006; Klieme, Schümer, Knoll, 2001; Kunter & Voss, 2011). Wenngleich gewisse Unterschiede bezüglich der inhaltlichen Ausgestaltung der drei Dimensionen bestehen, beziehen sich die drei Dimensionen typischerweise auf Klassenführung, kognitive Aktivierung und Unterstützung der Schüler und Schülerinnen.

Unterrichtsqualität kann mithilfe von Schülerbefragungen, Lehrerbefragungen oder Unterrichtsbeobachtungen erfasst werden. Häufig werden Unterrichtsbeobachtungen als „Goldstandard“ angesehen (Göllner et al., 2016). Mit dem „Classroom Assessment Scoring System“ (CLASS) (Pianta, Le Paro, Hamre, 2008) steht ein Beobachtungsinstrument zur Verfügung, welches inhaltlich anschlussfähig an die genannten drei Dimensionen der Unterrichtsqualität ist.

Stuck, Kammermeyer und Roux (2016) überprüften die Faktorstruktur einer deutschen Übersetzung des CLASS und fanden, dass ein dreifaktorielles Messmodell die Daten angemessen beschreibt. Allerdings zeigte sich eine sehr hohe Korrelation von  $r = 0.96$  zwischen den Dimensionen „Classroom Organization“ und „Emotional Support“. Ein zweifaktorielles Messmodell, in dem diese beiden Dimensionen zu einem Faktor zusammengefasst werden, wies einen etwas besseren Modellfit auf als das dreifaktorielle Messmodell. Dieser Befund wirft die Frage nach der theoretischen Trennschärfe der beiden Dimensionen auf.

Vor diesem Hintergrund wird im DFG-geförderten Forschungsprojekt FEED-UB (Thiel & Gärtner, 2016) ein neues Beobachtungsinstrument zur Erfassung von Unterrichtsqualität entwickelt, welches insbesondere als Grundlage für anschließende Feedbackprozesse genutzt werden soll. Die Dimensionen der Unterrichtsqualität werden so konzeptualisiert und operationalisiert, dass sie inhaltlich klar abgrenzbar sind. Die drei Dimensionen werden entsprechend unter Bezugnahme auf drei unterschiedliche Theoriestränge hergeleitet. Für die erste Dimension „Unterstützung des Wissenserwerbs“ wird hierbei auf Theorien zum kognitiven Lernen (Klauer & Leutner, 2012) Bezug genommen. Die zweite Dimension „Motivierung“ basiert auf Theorien zur Lernmotivation (Rheinberg, 2008). Die dritte Dimension „Klassenmanagement“ verweist auf Theorien zur Lehrer-Schüler-Interaktion (Thiel, 2016).

Fragestellungen

In der vorliegenden Studie wird untersucht, welche Faktorstruktur das neu entwickelte Beobachtungsinstrument aufweist. Im Vordergrund stehen die Fragen, ob sich die theoretische Zuordnung der Items zu den drei Dimensionen empirisch bestätigen lässt und wie stark die drei Dimensionen miteinander korrelieren.

#### Methoden

Zunächst wurde eine Pilotierungsstudie mit Lehramtsstudierenden durchgeführt. Die in der Anwendung des Beobachtungsinstruments geschulten Studierenden führten insgesamt 1184 Unterrichtsbeobachtungen in mehreren Schulformen, Jahrgangsstufen und Unterrichtsfächern durch.

Die Daten für die Hauptstudie werden gegenwärtig erhoben. Schulleitungen von Berliner Oberstufenzentren erhalten ein mehrtägiges Training und wenden das Beobachtungsinstrument im Unterricht ihrer Lehrkräfte an.

Als Analyseverfahren wird eine konfirmatorische Faktorenanalyse eingesetzt.

#### Ergebnisse

Bisher konnten die Daten der Pilotierungsstudie analysiert werden. Die Ergebnisse der konfirmatorischen Faktorenanalyse sind insgesamt zufriedenstellend. Die Maße für den Modellfit weisen größtenteils gute Werte auf. Die hohen Faktorladungen sprechen dafür, dass die Items geeignet sind, um die drei Dimensionen zu erfassen. Die latenten Korrelationen zwischen den Dimensionen bewegen sich zwischen  $r = 0.56$  und  $r = 0.63$  und sind damit deutlich niedriger als die Korrelation zwischen den zwei Dimensionen des CLASS. Erste Rückmeldungen von trainierten Schulleitungen bestätigen zudem die inhaltliche Trennschärfe der drei Dimensionen.

Bis zur Konferenz im Februar werden die Daten der Schulleitungskräfte ausgewertet sein. Diese lassen sowohl Rückschlüsse über die Belastbarkeit der Faktorenstruktur im Feld als auch über die Anwendbarkeit des Bogens für ein anschließendes Feedback an Lehrkräfte zu.

## **Muster guten Unterrichts: Latente Profizugehörigkeit angehender Physiklehrkräfte und deren Prädiktion durch professionelle Kompetenz**

**Marvin Krüger, Michael Szogs, Friederike Korneck**

Goethe-Universität Frankfurt, Deutschland; [krueger@physik.uni-frankfurt.de](mailto:krueger@physik.uni-frankfurt.de)

#### Einleitung und theoretischer Hintergrund

Zentrales Ziel bei der Ausbildung zukünftiger Physiklehrkräfte ist die Befähigung zur Gestaltung qualitativ hochwertigen Unterrichts. Vor diesem Hintergrund hat das Frankfurter Projekt PHlactio im physikdidaktischen Diskurs dazu beitragen können, das Professionswissen und die Lehrerüberzeugungen als Determinanten professionellen Unterrichtshandelns von Physiklehramtsstudierenden nachzuweisen (Korneck, Krüger & Szogs, 2017). Unterrichtsqualität wird dabei im Sinne des Modells der Basisdimensionen durch die Merkmale kognitive Aktivierung, konstruktive Unterstützung und Klassenführung konzeptualisiert (Klieme & Rakoczy, 2008; Kunter & Voss, 2011).

Der bisherige variablenzentrierte Ansatz soll mit dieser Studie durch einen personenzentrierten Ansatz ergänzt werden, wie er von Holzberger, Kunter, Praetorius und Seidel (2016) vorgeschlagen wurde. Diese haben anhand einer Reanalyse der COACTIV-Daten von 155 Mathematik-Lehrkräften latente Profile identifizieren können, die sich im Wesentlichen in der Höhe der Merkmalsausprägungen unterscheiden haben und sich als je ein unterdurchschnittliches, durchschnittliches und überdurchschnittliches Profil herausstellten.

Entgegen der Vermutung, dass Lehrkräfte in ihrem Unterricht spezifische Schwerpunkte auf einzelne Merkmale der Unterrichtsqualität legen (Klieme, Schümer & Knoll, 2001), deuten diese Ergebnisse darauf hin, dass es eher gute, normale und schlechte Lehrkräfte gibt. Da diese Befunde jedoch ausschließlich auf der häufig weniger differenzierten Schülersicht (Clausen, 2002) beruhen, gilt es zu klären, inwiefern sie übertragbar sind.

#### Fragestellung

Dieser Beitrag geht der Frage nach, ob sich die von Holzberger et al. (2016) identifizierten Profile mit einer anderen Kohorte, in einer anderen Domäne und bei Beurteilung des Unterrichts durch Videoratings ebenfalls nachweisen lassen. Dazu gilt es zu klären, ob, wie viele und welche Unterrichtsqualitätsmuster sich bei angehenden Physiklehrkräften nachweisen lassen. Darüber hinaus soll untersucht werden, inwiefern Komponenten professioneller Kompetenz zur Vorhersage der Profizugehörigkeit beitragen.

#### Methode

Die Stichprobe umfasst 125 Physik-Lehramtsstudierende. Das Forschungsprojekt ist an ein Microteaching-Seminar gekoppelt, in dem Studierende vorab geplante sog. Unterrichtsminiaturen gestalten. Diese Unterrichtsminiaturen zeichnen sich durch eine Komplexitätsreduktion hinsichtlich der Zeit (12 Minuten), der Schüler(innen) (fremde Klassenhälfte) und des Inhalts (Freihandexperiment der Mechanik) aus. Die zentrale Stellung eines Experimentes ermöglicht einen kohärenten inhaltlichen Abschluss, sodass die Miniaturen vom Einstieg bis zur Ergebnissicherung eine repräsentative Raffung von Regelunterricht darstellen können (Korneck, Oettinghaus, Kunter & Redinger, 2016). Die videographierten Miniaturen dienen der Erfassung der Unterrichtsqualität, deren Einschätzung anhand eines Ratingmanuals erfolgt. Dieses umfasst 177 Items in 20 Subdimensionen der Unterrichtsqualität. Die Subdimensionen (Sd) repräsentieren die kognitive Aktivierung (4 Sd), den Bereich affektiver konstruktiver Unterstützung (6 Sd), struktureller konstruktiver Unterstützung (4 Sd), Klassenführung (4 Sd) und ergänzt Merkmale fachlicher Korrektheit und Transparenz (2 Sd).

Im Vorfeld des Unterrichts erfolgt die Erhebung ausgewählter Komponenten der professionellen Kompetenz. Neben der Erhebung von Fachwissen und fachdidaktischem Wissen auf Basis etablierter Inventare (Riese, 2009), die für die Studie gekürzt wurden, werden außerdem die Lehrerüberzeugungen durch Selbstauskunft erhoben. Dabei werden die drei Dimensionen „Überzeugungen zum transmissiven Lernen“, „Überzeugungen zum selbstständigen Lernen“ und „Wissenschaftsverständnis“ (Seidel, Prenzel & Kobarg, 2005; Riese, 2009; Neuhaus, 2004) berücksichtigt.

Die Subdimensionen gehen in Form aggregierter Mittelwerte von vier Videoratern in latente Profilanalysen (Gibson, 1959) ein, die mit Mplus (Muthén & Muthén, 1998-2015) durchgeführt werden. Als Prädiktoren für die Profizugehörigkeit werden Fachwissen, fachdidaktisches Wissen und Lehrerüberzeugungen sowie soziodemographische Kovariaten berücksichtigt.

#### Ergebnisse und Diskussion

Auf Grundlage unterschiedlicher Modellkriterien und statistischer Tests ist das Modell zu favorisieren, das drei latente Profile berücksichtigt. Diese lassen sich als ein unterdurchschnittliches (etwa 20% der Teilnehmer), ein durchschnittliches (etwa 45%) und ein überdurchschnittliches Profil (etwa 35%) identifizieren. Es konnten damit die Ergebnisse der COACTIV-Reanalyse von Holzberger et al. (2016) repliziert werden. Sie deuten darauf hin, dass die Vermutung von unterschiedlichen Schwerpunktsetzungen auf spezifische Dimensionen der Unterrichtsqualität auch in dieser Studie nicht aufrechterhalten werden kann.

Bei der Berücksichtigung der erhobenen Komponenten professioneller Kompetenz und Kovariaten konnte das Fachwissen der Studierenden als der wesentliche Prädiktor für die unterschiedliche Profilizugehörigkeit identifiziert werden.

## B10

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S107

### Der Umgang mit Fehlern und emotionales Erleben bei Lehrkräften. Differentielle und nicht-lineare Zusammenhänge.

**Anja Böhnke, Felicitas Thiel**

Freie Universität Berlin, Deutschland; [a.boehnke@fu-berlin.de](mailto:a.boehnke@fu-berlin.de)

Schüleremotionen und der Umgang mit Schülerfehlern im Unterricht sind umfangreich beforscht (Pekrun, Goetz, Titz & Perry, 2002; Oser & Spychiger, 2005) und werden in großen Schulleistungsstudien wie PISA mitberücksichtigt (Frey et al. 2009); auf Emotionen von Lehrkräften und ihren Umgang mit Fehlern beim Unterrichten trifft dies nicht in gleicher Weise zu.

Im Rahmen von professionellem Lernen, kann ein lernwirksamer Umgang mit eigenen Fehlern helfen, relevante Wissens- und Kompetenzlücken zu identifizieren (Harteis, Bauer & Gruber, 2008). Dass Lehrkräfte, die im Unterricht rasch und situationsangemessen handeln müssen und Unterrichtsverläufe schwer antizipieren können, von einem lernwirksamen Umgang mit Fehlern besonders profitieren, stellen Böhnke und Thiel (2016) dar. Ein zentraler Aspekt des Fehlerumgangs ist das emotionale Erleben (Oser, 2007). Wie genau die (zumeist negativen) Emotionen, die ein Individuum aufgrund eigener Fehler erlebt, mit einem lernwirksamen Umgang mit diesen Fehlern in Zusammenhang stehen, ist bislang jedoch weitgehend ungeklärt und die wenigen vorliegenden Studien kommen zu uneinheitlichen Befunden (z.B. Zhao, 2011; Harteis & Frost, 2012). Aus theoretischer Sicht stehen sich Argumente für einen hinderlichen und solche für einen förderlichen Einfluss negativer Emotionen auf Lernprozesse gegenüber. Einerseits können negative Emotionen dazu führen, dass Aufmerksamkeit fast ausschließlich auf die Emotion gelenkt und damit von möglichen Lernaktivitäten abgezogen wird (Meinhardt & Pekrun, 2003), Lernmotivation gemindert wird (Zhao, 2011) und Erwartungs- und Werte-beliefs negativ beeinflusst werden (Seo, Feldman Barrett & Bartunek, 2004). Andererseits zeigen negative Emotionen problematische Situationen auf und verweisen auf die Notwendigkeit einer sorgfältigen und detaillierten Analyse (Schwarz, 2002), können eine Alarmfunktion erfüllen (Bauer, 2008) und so ein Gefühl von persönlicher Bedeutsamkeit erzeugen (Oser & Spychiger, 2005).

Hauptanliegen der vorgestellten Studie war es, einen Beitrag zur Klärung des Zusammenhangs vom Umgang mit Fehlern und beteiligten Emotionen zu leisten und damit die Forschung zu Lehreremotionen und Fehlerumgang von Lehrkräften zu erweitern. Unsere Studie nahm zwei Aspekte von Emotionen in den Blick, die zur Erklärung der uneinheitlichen Befundlage beitragen könnten, bislang jedoch nicht ausreichend untersucht sind: Wirkung unterschiedlicher Emotionen und Grad der Intensität. Emotionales Erleben kann nicht als genereller Mechanismus verstanden werden (Hanoch & Vitouch, 2004), vielmehr ist davon auszugehen, dass unterschiedliche Emotionen auch eigene Wirkungsweisen entfalten. Diese Wirkungsweisen können zusätzlich vom Grad der Intensität empfundener Emotionen abhängen (Eysenck, 1982). Es wurden zwei Fragestellungen fokussiert:

1. Weisen einzelne Emotionen (Angst, Ärger, Schuld, Trauer) differentielle Zusammenhänge zu einem lernwirksamen Fehlerumgang auf?
2. Lassen sich dabei hinsichtlich der Intensität der empfundenen Emotion nicht-lineare Zusammenhänge nachweisen?

Um unsere Forschungsfragen zu untersuchen, haben wir N = 324 Referendare zu ihrem Umgang mit eigenen Fehlern während des Unterrichts befragt, außerdem sollten die Befragten einschätzen, in welchem Ausmaß sie einzelne Emotionen im Zusammenhang mit ihren Fehlern erleben (Fragebogenstudie).

Wir nutzten eine schrittweise Schätzprozedur die auf dem Latent Moderated Structural Equations (LMS)-Ansatz (Klein & Moosbrugger, 2000) basierte. Der LMS-Ansatz berücksichtigt eine spezifische Abweichung von der Normalverteilung, die bei der Analyse latenter Interaktionseffekte auftreten kann, impliziert aber auch Limitationen, da keine Model-Fit-Werte oder Indikatoren bzgl. Effektgrößen zur Verfügung gestellt werden.

Entsprechend den Vorschlägen von Maslowsky, Jager & Hemken, (2015) erweiterten wir daher den LMS-Ansatz und spezifizierten drei aufeinander aufbauende Modelle: Model 1 wurde mit der maximum-likelihood-Voreinstellung geschätzt, um Model-Fit-Werte zu erhalten. Da die Schätzung quadratischer Effekte im LMS-Ansatz numerische Integration erfordert, wurde ein Zwischenmodell spezifiziert (nur lineare Effekte, numerische Integration), das via likelihood-ratio-test mit Model 3 (lineare und quadratische Effekte, numerische Integration) verglichen werden konnte.

Es konnten differentielle lineare Effekte aufgezeigt werden (Ärger und Angst hingen negativ, Schuld dagegen positiv mit einem lernwirksamen Fehlerumgang zusammen – für Trauer zeigte sich kein Zusammenhang); zusätzlich zeigten sich für Angst negativ quadratische und für Ärger positiv quadratische Effekte (range  $R^2 = .233 - .355$  ( $ps < .01$ )).

### Die Bedeutung der Selbstwirksamkeit von Lehrpersonen für ihre professionelle Handlungskompetenz in Schulen mit personalisierten Lernkonzepten

**Margaretha Florin, Kurt Reusser**

Institut für Erziehungswissenschaften, Uni ZH, Schweiz; [maflorin@bluewin.ch](mailto:maflorin@bluewin.ch)

Selbstwirksamkeit im Lehrerberuf meint den subjektiven Glauben von Lehrpersonen an ihre eigenen Fähigkeiten und professionellen Handlungskompetenzen, um die anspruchsvollen Aufgaben und Anforderungen im schulischen Alltag erfolgreich zu bewältigen. Beim personalisierten Lernen sind diese Aufgaben unter anderen: Im Unterricht differenzieren, individuelle Lernunterstützung anbieten und selbstgesteuertes Lernen ermöglichen sowie im Team zusammenarbeiten und gemeinsam innovative Lernkonzepte umsetzen.

Das theoretische Konstrukt der Selbstwirksamkeitserwartung (engl. self-efficacy) beruht auf der sozial-kognitiven Theorie von Bandura und wird folgendermassen definiert: „Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the course of action required to produce given attainments“ (Bandura, 1997, p.3). Das Konzept findet auch in der erziehungswissenschaftlichen Forschung breite Resonanz: Selbstwirksamkeit zählt, neben Professionswissen, Überzeugungen und Werthaltungen, intrinsischer motivationaler Orientierung und Selbstregulationsfähigkeit (Kunter et al., 2011), zu den zentralen Lehrermerkmalen professioneller Handlungskompetenz. Holzberger (2013) definiert Selbstwirksamkeitserwartungen von Lehrpersonen (engl. teacher efficacy) als subjektive Erwartung, über die notwendigen Voraussetzungen erfolgreichen Unterrichtens zu verfügen. In zahlreichen Studien konnten positive Zusammenhänge der Selbstwirksamkeit mit anderen Lehrermerkmalen (etwa Enthusiasmus, körperliches Wohlbefinden, Berufszufriedenheit) aber auch mit der Unterrichtsqualität (etwa Schülerleistungen, Schülerurteil zum Unterricht) gefunden werden (Holzberger, 2013; Jerusalem & Hopf, 2002;



Tschannen-Moran & Woolfolk Hoy, 2001; Zee & Koomen, 2016). Allerdings wird die Selbstwirksamkeit in vielen Studien sehr allgemein betrachtet, das heisst, die Dimensionen des Unterrichts, auf welche sich die Selbstwirksamkeitserwartung bezieht, werden nicht detailliert erfasst. Zudem ist die Wirkweise der Lehrerselbstwirksamkeit nicht zufriedenstellend geklärt (Holzberger, 2013). Diese Forschungsdesiderate sind Gegenstand einer laufenden Dissertation. Ziel ist, ein vertieftes Verständnis davon zu erlangen, was das Konstrukt ‚Selbstwirksamkeit‘ im Kontext von ‚learning and instruction‘ beinhaltet und welche Bedeutung ihm als Merkmal professioneller Handlungskompetenz in Bezug auf die Anforderungen, nach personalisierten Lernkonzepten zu unterrichten, zukommt. Schulen mit personalisierten Lernkonzepten stehen vor der herausfordernden Aufgabe, als Kollegium konsensfähige pädagogische und didaktische Reformkonzepte zu entwickeln und umzusetzen, in deren Rahmen die Schülerinnen und Schüler zu autonomen, motivierten und kompetenten Lernenden werden.

Aus der Auseinandersetzung mit Theorien und Forschungsarbeiten zur „Selbstwirksamkeit“ resultiert ein theoretisches Modell zur Genese von Lehrer-Selbstwirksamkeit im Kontext von personalisiertem Lernen, das im Kongressbeitrag vorgestellt wird. Es wird dargelegt, welche personalen und kontextualen Faktoren den kognitiven Prozess der Genese von Selbstwirksamkeitserwartungen beeinflussen.

Dieses Modell dient als Grundlage für die empirischen Analysen mit der Kernstichprobe (53 Schulen) der als Längsschnitt konzipierten Schweizer perLen-Studie (Reusser et al., 2015, [http://www.zora.uzh.ch/id/eprint/119861/1/PerLen\\_Zwischenbereich\\_2015\\_mit\\_Einblicken\\_in\\_Teilprojekte.pdf](http://www.zora.uzh.ch/id/eprint/119861/1/PerLen_Zwischenbereich_2015_mit_Einblicken_in_Teilprojekte.pdf)). Es werden Daten der Onlinebefragungen zu drei Erhebungszeitpunkten analysiert um die wechselseitigen Zusammenhänge von Selbstwirksamkeit mit anderen personalen (z.B. Enthusiasmus) und mit kontextualen Faktoren (z.B. Schul- und Innovationsklima) sowie Unterschiede je nach Personalisierungsgrad aufzuzeigen.

Die bis jetzt vorliegenden Ergebnisse zeigen, dass perLen-Lehrpersonen ihren Beruf mit Enthusiasmus ausüben und sich beim Unterrichten nach personalisierten Lernkonzepten als selbstwirksam einschätzen. Die untersuchten Lehrpersonen beurteilen sowohl das Schulklima als auch die kollektive Selbstwirksamkeit und die Identifikation mit der Schule über alle Messzeitpunkte sehr positiv. In perLen-Schulen mit sehr starker Orientierung an personalisierten Lernkonzepten erleben Lehrpersonen eine signifikant höhere kollektive Selbstwirksamkeit und beurteilen das Schul- und Innovationsklima wesentlich positiver als in Schulen mit moderater Personalisierung. Vertiefende Regressionsanalysen zeigen, dass die individuelle Selbstwirksamkeit am besten durch die Identifikation mit der Schule, die kollektive Selbstwirksamkeit und durch den Enthusiasmus vorausgesagt werden kann. Im Kongress-Beitrag werden diese sowie weitere Ergebnisse aus den (Längsschnitt-)Analysen präsentiert.

### **Direkt oder indirekt – Werden Zusammenhänge zwischen Überzeugungen von Lehrkräften und Lernergebnissen über Adaptivität im Unterricht vermittelt?**

**Julia Dohrmann<sup>1</sup>, Tobias Feldhoff<sup>2</sup>, Brigitte Steinert<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF); <sup>2</sup>Johannes Gutenberg-Universität Mainz; [dohrmann@dipf.de](mailto:dohrmann@dipf.de)

Die Schul- und Unterrichtsforschung hat in den vergangenen Jahren verschiedene Aspekte des Wissens und Könnens von Lehrkräften als zentrale Komponenten des professionellen Handelns herausgearbeitet, die eine Bedeutung für den Unterricht haben (Baumert und Kunter 2006). Ein Teilbereich dieser professionellen Handlungskompetenz sind die Überzeugungen und Werthaltungen von Lehrkräften. Die Forschungslage zu diesem Bereich ist, gerade in Deutschland, eher dürftig – insbesondere in Bezug auf die Fragen, wie Überzeugungen den Unterricht beeinflussen und für welche Aspekte des Unterrichtshandelns und für welche Schüleroutcomes ein solcher Einfluss vorliegt (Reusser et al. 2011).

Um zu klären, welche Zusammenhänge zwischen pädagogischen Überzeugungen von Lehrkräften, ihrem Handeln im Unterricht und den Lernergebnisse von Schülerinnen und Schülern bestehen, liegt mit der Drei-Länder-Studie von Helmut Fend aus den Jahren 1978 und 1979 (Fend 1982) ein geeigneter Datensatz vor.

In der Drei-Länder-Studie werden Überzeugungen von Lehrkräften zu erfolgreichem Lernen von Schülerinnen und Schülern erfasst, z. B. die Einstellung der Lehrkräfte zur Förderung von schwachen Schülerinnen und Schülern und die Einstellung zur Notwendigkeit von Disziplinierung. Darüber hinaus wird die Reformbereitschaft von Lehrkräften in die Analyse mit einbezogen, die sich auf die Offenheit bezieht, mit der die Schule und die Lehrkräfte auf gesellschaftliche Anforderungen reagieren sollten. Im vorliegenden Beitrag wird der Zusammenhang zwischen den Überzeugungen von Lehrkräften, ihrem pädagogischen Handeln im Unterricht und ausgewählten kognitiven und affektiven Schüleroutcomes auf der Grundlage der Daten der Drei-Länder-Studie analysiert. Das Unterrichtshandeln wird anhand der Dimension „Adaptivität des Unterrichts“ untersucht, welche Facetten von Mikroadaptation, beispielsweise individualisierte Rückmeldungen an Schülerinnen und Schüler, und von Makroadaptation, beispielsweise die Anpassung der Aufgaben an die Voraussetzungen der Schülerinnen und Schüler, umfasst (Corno und Snow 1986). Als kognitives Schüleroutcome werden die Englischleistungen der Schülerinnen und Schüler, als affektives Schüleroutcome deren Wohlfühlen im Englischunterricht untersucht.

Die Daten werden mit Hilfe von Zwei-Ebenen-Mediationsanalysen auf den Ebenen Schüler/in und Kurs bzw. Klasse mit Mplus 7.11 analysiert. Bei allen Analysen werden soziobiographische Hintergrundmerkmale auf individueller Ebene (Geschlecht, zu Hause gesprochene Sprache und Sozialstatus des Vaters), ihre Komposition auf Klassenebene sowie die Schulart kontrolliert.

Ein Ergebnis der Analysen ist, dass es im Englischunterricht einen positiven Zusammenhang zwischen der Reformbereitschaft von Lehrkräften und der Adaptivität des Unterrichts gibt sowie zwischen Adaptivität und dem Wohlfühlen der Schülerinnen und Schüler im Unterricht. Somit zeigt sich ein indirekter Zusammenhang zwischen den Überzeugungen der Lehrkräfte, hier Reformbereitschaft, und dem affektiven Schüleroutcome, welcher über die Adaptivität vermittelt wird. Für die Einstellung von Lehrkräften zu Disziplinierung findet sich derselbe indirekte Zusammenhang in negativer Form. Die Einstellung der Lehrkräfte zur Förderung von schwachen Schülerinnen und Schülern hängt sowohl direkt als auch indirekt mit dem Wohlfühlen der Schülerinnen und Schüler zusammen, wiederum vermittelt über die Adaptivität des Unterrichts.

Bei den Englischleistungen der Schülerinnen und Schüler lassen sich für keine der drei Überzeugungen indirekte Zusammenhänge vermittelt über die Adaptivität des Englischunterrichts finden. Allerdings weist die Reformbereitschaft der Lehrkräfte einen direkten negativen Zusammenhang mit den Englischleistungen der Schülerinnen und Schüler auf.

### **Untersuchung der linearen, nichtlinearen, additiven und multiplikativen Effekte unterrichtlicher Faktoren und deren Zusammenspiel auf den Schülerleistungszuwachs im Fach Englisch mithilfe von regularisierter Regression**

**Giang Pham**

## Fragestellung und theoretischer Hintergrund

Die zentrale Fragestellung dieser Arbeit ist die nach den hypothetischen Effekten unterrichtlicher Faktoren und deren Zusammenspiel auf den Schülerleistungszuwachs unter Berücksichtigung von Kontextfaktoren. Dafür wird ein methodischer Schwerpunkt gesetzt, um das Problem des kleinen Stichprobenumfanges bei großer Zahl von Variablen zu berücksichtigen.

Der Zusammenhang zwischen Merkmalen der Unterrichtsqualität und Schülerleistung wurde in der Literatur meist als komplex angenommen. Dabei wurde hypothetisch neben einem linearen Zusammenhang zwischen Unterrichtsvariablen und dem Zielkriterium auch ein nicht-linearer Zusammenhang hypothesiert (Helmke, 2014). Das Zusammenspiel verschiedener Faktoren hat zudem zunehmend an Aufmerksamkeit in der Unterrichts- und Lehr-Lern-Forschung (Holzberger, Kunter, Praetorius, & Seidel, 2016; Seidel & Prenzel, 2004). Diesbezüglich wurde neben additiven Wirkungen mehrerer Unterrichtsmerkmale auch multiplikative und interaktive Wirkungen mit anfänglichen Schülerleistungen (Aptitude-Treatment Effekt, ATI Effekt) angenommen (Helmke, 2014; Oser & Baeriswyl, 2002; Snow, 1989).

## Methode

Die Stichprobe umfasst 2096 Schülerinnen und Schüler von 50 Schulklassen aus 3 Regionen in Vietnam, die Daten wurden im Rahmen eines Forschungsprojekts in Vietnam im Schuljahr 2006/2007 erhoben. Neben einem Messwiederholungsdesign mit zwei Messzeitpunkten am Anfang und am Ende des Schuljahres wurde eine Videostudie in der Mitte des Schuljahres durchgeführt. Bei jedem Messzeitpunkt wurden der adaptierten C-Test und die Fragebögen aus der Studie Deutsch Englisch Schülerleistungen International (DESI) in Deutschland eingesetzt.

Zur Gewinnung verhaltensnaher Indikatoren des Unterrichts wurden die Videoaufzeichnungen von trainierten Experten niedrig-inferent kodiert und hoch-inferent geratet. Die Kodierungs- und Ratingkategorien wurden auf Basis internationaler Befunde der empirischen Unterrichts- und Lehr-Lern-Forschung sowie der Englisch-Fachdidaktik entwickelt.

Die Schülerleistung anhand des C-Tests wurde mithilfe des Raschtestletmodells (Wang & Wilson, 2005) skaliert. Die Leistungsschätzungen von beiden Messzeitpunkten wurden mithilfe eines gemeinsamen Skalierungsmodells mit unterschiedlichen A-Priori-Fähigkeitsverteilungen skaliert und somit miteinander verlinkt. Anschließend wurden die Plausible Values gezogen, indem die fehlenden Werte der Kovariaten und die latenten Schülerleistungswerte simultan behandelt (Robitzsch, 2016).

Die Zusammenhänge zwischen den Unterrichtsfaktoren und dem Zielkonstrukt wurden mittels Regressionsmodelle auf Klassenebene analysiert, dabei wurde die anfängliche Leistung der Schülerinnen und Schüler als Kontrollvariable betrachtet. Zur Untersuchung der nicht-linearen Zusammenhangshypothese wird neben dem Haupteffekt auch der quadratische Term als Modellprädiktor betrachtet (Cohen, Cohen, West, & Aiken, 2003). Die multiplikative Wirkung verschiedener Faktoren und der ATI Effekt werden durch Hinzufügung von Interaktionsterme ins Modell untersucht, die additive Wirkung verschiedenen Faktoren wird mittels multipler Regression analysiert (Eid, Gollwitzer, & Schmitt, 2010). Mit insgesamt 81 Unterrichtsvariablen und 50 Schulklassen kommt das Problem der kleinen Stichprobe bei großer Variablenzahl vor, wenn additive Wirkung verschiedener Faktoren untersucht wird (Bühlmann & van de Geer, 2011). Dabei ist es notwendig, nur die relevanten Variablen zu identifizieren. Um mit diesem Problem umzugehen werden regularisierte Regressionen mit lasso (least absolute shrinkage and selection operators) angewendet. Die Lasso-Regression (Tibshirani, 1996) hat in den letzten Jahrzehnten großer internationaler Aufmerksamkeit erregt, neulich auch in der Bildungsforschung (McNeish, 2015; Pham, Robitzsch, George, & Freunberger, 2016). Durch lasso erhalten alle nichtbedeutsamen Modellprädiktoren einen Regressionskoeffizienten von Null. Zur Untersuchung der nicht-linearen und multiplikativen sowie der ATI Wirkungen der Unterrichtsfaktoren ist es zudem nötig, die hierarchische Struktur der Modellprädiktoren zu behalten (der Regressionskoeffizient eines quadratischen oder Interaktionsterms sollte nur nicht Null sein, wenn die Haupteffekte nicht Null sind). Dafür wird die lasso-Regression für streng hierarchische Interaktionen eingesetzt (Bien, Taylor, & Tibshirani, 2013).

## Ergebnisse

Die Hypothesen zu den Unterrichtseffekten wurden durchweg bestätigt. Die bedeutsamsten Unterrichtsfaktoren des Schülerleistungszuwachses im C-Test sind Aspekte der Motivierungsqualität sowie die Verwendung der gemischten Sprachen im Unterricht. Nicht-lineare Zusammenhänge zwischen Schülerleistungszuwachs und «Schüleräußerung: Ablesen eigenen Texts» sowie «Lehrernutzung der Muttersprache im Übergang» wurden gefunden, bei diesen Aspekten gilt der Ausdruck «je mehr, desto besser» nicht. Die additiven, multiplikativen sowie ATI Wirkungen des Unterrichts wurden ebenso bestätigt und werden diskutiert.

## B12

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H119

### Wünschenswerte Erschwernisse im Mathematikunterricht

Chair(s): **Frank Lipowsky** (Universität Kassel, Deutschland), **Julia Arend** (Universität Kassel, Deutschland)

DiskutantIn(nen): **Aiso Heinze** (IPN Kiel, Deutschland)

Nachhaltige Lernprozesse zu initiieren stellt ein wesentliches Ziel professioneller Handlungskompetenz von Lehrkräften dar. Zur Erreichung dieses Ziels können sogenannte „wünschenswerte Erschwernisse“ beitragen. Unter dem Begriff der „wünschenswerten Erschwernisse“ (desirable difficulties) werden unterschiedliche gedächtnispsychologische Konzepte zusammengefasst, welche den Lernvorgang kurzfristig gezielt erschweren, sich jedoch langfristig förderlich auf den Lern- und Verstehensprozess auswirken sollen (Bjork & Bjork, 2011). Hierzu zählen das verteilte Lernen (distributed learning/ practice), das verschachtelte Lernen (interleaved learning/ practice), die Generierung von Informationen (generation) und Tests als Lerngelegenheiten (testing). Der Fokus des Symposiums liegt auf dem verteilten und dem verschachtelten Lernen.

Die Wirksamkeit wünschenswerter Erschwernisse wurde bereits in zahlreichen Laborstudien mit älteren Schüler/-innen und Erwachsenen belegt (Dunlosky, Rawson, Marsh, Nathan & Willingham, 2013). Inwieweit die Befunde auf den Schulkontext und insbesondere jüngere Lernende übertragbar sind, welche Moderatorvariablen hierbei relevant sind und welche zusätzlichen instruktionalen Unterstützungsmaßnahmen ggf. erforderlich sind, ist dagegen bislang kaum erforscht.

Hier setzt das interdisziplinäre Schwerpunktprogramm der Universität Kassel zu den „Wünschenswerten Erschwernissen“ an, welches im Rahmen der hessischen Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz (LOEWE) gefördert wird. Aus dem Schwerpunktprogramm werden im Rahmen des Symposiums vier Projekte vorgestellt. Die Symposiumsbeiträge umfassen eine Metaanalyse zum verschachtelten Lernen (Brunmair und Richter), eine Unterrichtsstudie zum verteilten Lernen (Barzagar Nazari und Ebersbach) und zwei Unterrichtsstudien zum verschachtelten Lernen (Pede & Borromeo Ferri; Werker, Arend & Lipowsky). Die genannten Unterrichtsstudien wurden im Fach Mathematik durchgeführt.

Als Rahmen für das Symposium fungiert die Metaanalyse von Brunmair und Richter, da sie einen Überblick über die Befundlage und existierende Moderatorvariablen bietet, einen Beitrag zur Einordnung der Ergebnisse der Unterrichtsstudien leistet und Forschungslücken aufzeigt. Konkret zeigt sich bei der Auswertung der verfügbaren Studien zum verschachtelten Lernen, dass verschachteltes Lernen insbesondere dann effektiv ist, wenn sich die Exemplare der zu vergleichenden Kategorien ähneln und diese somit konzeptuell schwer unterscheidbar sind. Darüber hinaus wird eine höhere Wirksamkeit verschachtelten Lernens bei bildhaft dargestellten Inhalten ermittelt. Bezogen auf mathematische Inhalte zeigt sich eine erhebliche Varianz in den berichteten Effekten. Die Metaanalyse verdeutlicht zudem, dass es an unterrichtspraktischen Studien mangelt. Daher ist offen, inwiefern die Befunde aus den vorliegenden Laborstudien auf das Lernen von Mathematik in der Schule übertragbar sind und welche Rolle hierbei Formen instruktionaler Unterstützung der Lernenden spielen. Diese Fragestellungen werden in den weiteren Vorträgen des Symposiums aufgegriffen. Darin wird die Umsetzung des verschachtelten und des verteilten Lernens im Mathematikunterricht der Grundschule und der Sekundarstufe untersucht.

Für beide wünschenswerte Erschwernisse können in der Klassenstufe 3 bedeutsame Effekte nachgewiesen werden, d. h. die teilnehmenden Schüler/-innen, die verteilt bzw. verschachtelt gelernt haben, weisen einen höheren Lernzuwachs in Mathematik auf als diejenigen Schüler/-innen, welche die Lerninhalte massiert bzw. geblockt gelernt haben. In der Sekundarstufe ergeben sich dagegen keine Vorteile der beiden Erschwernisse für die jeweiligen Experimentalgruppen. Darüber hinaus enthalten die Studien Hinweise auf die Bedeutung instruktionaler Unterstützung beim verschachtelten und verteilten Lernen durch Feedback und explizite Anregung zum Strategievergleich.

Die Ergebnisse der Studien sind auch für die Professionalisierung von Lehrpersonen relevant und werden im Hinblick auf ihren Ertrag und ihre Reichweite diskutiert.

*Beiträge des Symposiums*

### Generalisierbarkeit und Wirkmechanismen des verschachtelten Lernens: Eine metaanalytische Untersuchung

**Matthias Brunmair<sup>1</sup>, Tobias Richter<sup>2</sup>**

<sup>1</sup>Universität Kassel, Deutschland, <sup>2</sup>Universität Würzburg, Deutschland

-Theoretischer Hintergrund-

Induktives Lernen anhand von Exemplaren ist eine wichtige Lernform, die z.B. beim Einüben verschiedener mathematischer Prozeduren, beim Erlernen biologischer Taxonomien oder bei der Abstraktion schematischer Kategorien aus Falldarstellungen Anwendung findet. Beim verschachtelten Lernen werden Exemplare unterschiedlicher Kategorien abwechselnd dargeboten, wogegen beim geblockten Lernen die Exemplare einer Kategorie jeweils nacheinander dargeboten werden. Zahlreiche Studien zeigen, dass das verschachtelte Lernen die bessere Form des induktiven Lernens sein könnte. In anderen Studien hatte die verschachtelte Vorgabe dagegen keine oder sogar negative Effekte. Eine mögliche Erklärung dieser inkonsistenten Ergebnisse liegt darin, dass das verschachtelte Lernen nur für bestimmte Arten von Lernmaterialien förderlich sein könnte, für andere dagegen nicht (Generalisierbarkeit). Eine zweite mögliche Erklärung hebt darauf ab, dass durch die verschachtelte Vorgabe Unterschiede zwischen Kategorien hervorgehoben werden, während die geblockte Vorgabe eher die Ableitung von Gemeinsamkeiten erleichtert. Diese Annahme steht im Mittelpunkt der Theorie der Aufmerksamkeitsausrichtung (Carvalho & Goldstone, 2015). Aus der Theorie folgt, dass das verschachtelte Lernen nur dann vorteilhaft sein sollte, wenn die Unterschiede zwischen Kategorien lernrelevant sind, während das geblockte Lernen dann vorteilhaft ist, wenn die Gemeinsamkeiten von Exemplaren derselben Kategorie für den Erwerb von Kategorienstrukturen wichtiger sind.

-Fragestellung-

Die Ziele unserer Metaanalyse bestanden darin, (1) den Gesamteffekt des verschachtelten vs. geblockten Lernens über sämtliche publizierte Untersuchungen hinweg zu ermitteln, (2) die Generalisierbarkeit des Effekts über unterschiedliche Lernmaterialien (und andere Lernbedingungen) hinweg zu prüfen – vor allem mit Blick auf "schulnahe" Lernmaterialien - und (3) die Vorhersagen von Erklärungsmodellen wie der Theorie der Aufmerksamkeitsausrichtung durch Moderatoranalysen zu überprüfen.

-Methode-

In die Metaanalyse wurden sämtliche Studien eingeschlossen, die sowohl eine verschachtelte als auch eine geblockte Lernbedingung untersuchten, wobei alle übrigen Aspekte der Lernbedingungen identisch sein mussten. Von 98 potentiell relevanten wissenschaftliche Arbeiten erfüllten insgesamt 39 Arbeiten mit  $k = 151$  Effekten und  $N = 5307$  Personen die Einschlusskriterien.

Um die Datenstruktur und Abhängigkeiten zwischen Effekten zu berücksichtigen, wurden Mehrebenenmodelle geschätzt, anhand derer ein Gesamteffekt und mittlere Effekte für verschiedene Typen von Lernmaterialien ermittelt wurden. Moderatoranalysen wurden anhand von multiplen Metaregressionen durchgeführt. Die Robustheit der Effekte wurde über Sensitivitätsanalysen abgesichert, und es wurde die Möglichkeit eines Publikationsbias geprüft.

-Ergebnisse-

Insgesamt zeigt sich für das verschachtelte Lernen ein mittlerer Effekt ( $g = 0.44$ ,  $k = 151$ ,  $CI[0.33, 0.54]$ ). Eine verschachtelte Darbietung erwies sich als besser bei naturalistischen Gemälden ( $g = 0.76$ ,  $k = 40$ ,  $CI[0.61, 0.91]$ ), fotografischen Abbildungen ( $g = 0.44$ ,  $k = 13$ ,  $CI[0.22, 0.66]$ ) und künstlich erstellten Bildmaterialien ( $g = 0.31$ ,  $k = 45$ ,  $CI[0.15, 0.48]$ ). Ein knapp signifikanter Effekt zeigte sich auch für expository Texte ( $g = 0.29$ ,  $k = 18$ ,  $CI[0.01, 0.57]$ ). Kein signifikanter Vorteil oder negative Effekte zeigten sich für die verschachtelte Vorgabe mathematischer Aufgaben ( $g = 0.35$ ,  $k = 18$ ,  $CI[-0.05, 0.75]$ ), Geschmacksproben ( $g = 0.57$ ,  $k = 3$ ,  $CI[-0.11, 1.26]$ ) und Wörter ( $g = -0.41$ ,  $k = 13$ ,  $CI[-0.61, -0.20]$ ).

Der positive Effekt des verschachtelten Lernens war umso größer, je ähnlicher sich Exemplare unterschiedlicher Kategorien waren, und umso geringer, je höher die Ähnlichkeit der Elemente innerhalb einer Kategorie war. Es zeigte sich kein Unterschied des Verschachtelungseffekts für simultane vs. sukzessive Vorgabe, für die Testung neuer oder zuvor präsentierter Elemente oder für unterschiedliche Behaltensintervalle. Es gab keine Hinweise auf einen Publikationsbias.

-Diskussion-

Insgesamt stützen die Ergebnisse die Theorie der Aufmerksamkeitsausrichtung von Carvalho und Goldstone (2015). Belege für die Generalisierbarkeit insbesondere auf schulnahe Lernmaterialien und -bedingungen fanden sich nicht. Allerdings wurde das verschachtelte Lernen bislang kaum unter schulnahen Bedingungen und mit längeren Behaltensintervallen, wie sie für schulisches Lernen typisch sind, untersucht. Gleichzeitig war die Heterogenität der Ergebnisse und Studiendesigns hoch. Für eine genaue Bestimmung der Eignung und gewinnbringender Umsetzung verschachtelter Lernanordnungen in schulischen Settings bedarf es daher weiterer Untersuchungen.

### **Besser lernen durch verteiltes Üben? Befunde zur Anwendbarkeit wünschenswerter Erschwernisse im Rahmen des Mathematikunterrichts**

**Katharina Barzaqar Nazari, Mirjam Ebersbach**  
Universität Kassel, Deutschland

-Theoretischer Hintergrund-

Unter wünschenswerten Erschwernissen beim Lernen versteht man Mechanismen, die das Lernen subjektiv erschweren, insbesondere das langfristige Behalten des Gelernten aber verbessern sollen (Bjork, 1994). Das Verteilen der Lern- oder Übungszeit auf mehrere Zeitpunkte statt des Lernens am Stück gehört zu diesen wünschenswerten Erschwernissen und hat in Laborstudien bereits positive Effekte für das Behalten gezeigt (Cepeda et al., 2006). Bislang gibt es jedoch kaum Studien, in denen der Effekt des verteilten Übens in realistischen Unterrichtssituationen mit curriculumsnahen Materialien getestet wurde.

-Fragestellung-

In zwei Studien haben wir den Effekt des verteilten Übens im Rahmen des Mathematikunterrichts in verschiedenen Klassenstufen der Grund- und Sekundarstufe untersucht. Zusätzlich sind wir der Frage nachgegangen, ob der Effekt durch kognitive und motivationale Variablen seitens der Schüler/-innen moderiert wird.

-Methode-

In Studie 1 erhielten 237 Dritt- und Siebtklässler eine Einführung in mathematische Prozeduren (Halbschriftliche Multiplikation bzw. Wahrscheinlichkeitsrechnung), die sie im Anschluss entweder an drei Tagen für jeweils 15 min oder an einem Tag 45 min lang einübten. In einer zweiten Studie erhielten 126 Siebt- und Zehnt-/Elftklässler eine Einführung in die Wahrscheinlichkeitsrechnung und übten diese ebenfalls entweder verteilt oder massiert. Im Vergleich zur ersten Studie erhielten die Schüler/-innen hier nach jeder Übungsaufgabe Feedback, indem ihnen die richtige Lösung angezeigt wurde, um sicherzustellen, dass auch die korrekten Prozeduren eingeübt wurden. Die Prüfung des Haupteffektes der Übungsbedingung erfolgte mittels hierarchischer linearer Regressionen, für die explorative Analyse potentieller Moderatoreffekte wurde zudem auf conditional inference tree-Modelle („bedingte Entscheidungsbäume“) zurückgegriffen.

-Ergebnisse-

In Studie 1 erzielten Drittklässler, die verteilt geübt hatten, sowohl nach einer als auch nach sechs Wochen signifikant höhere Testwerte in Bezug auf diese Prozeduren im Vergleich zu denen, die massiert geübt hatten. Für Siebtklässler ergab sich kein solcher Vorteil. In Studie 2 ergab sich weder für Siebt- noch für Zehnt-/Elftklässler ein Effekt des verteilten Übens. Bezog man allerdings das Vorwissen in Form der Leistung in der ersten Übung mit ein, zeigte sich ein Vorteil des verteilten Übens für Siebtklässler im mittleren Vorwissensbereich, während Siebtklässler/-innen, die die erste Übung sehr schwach oder sehr gut bearbeiteten, nicht stärker vom verteilten Üben profitierten als vom massierten Üben. Weitere Moderatoreffekte fanden sich in keiner der beiden Studien. Zusammenfassend lässt sich festhalten, dass das verteilte Üben von mathematischen Prozeduren im Mathematikunterricht zu besseren Lernergebnissen führen kann; in keinem Fall führt es zu schlechteren Lernergebnissen als das massierte Lernen. Weitere Rahmenbedingungen, unter denen ein positiver Effekt des verteilten Übens auftreten kann, werden diskutiert.

### **Zuordnungen besser verstehen durch Verschachtelung der Inhalte? – Eine Untersuchung zur Lernentwicklung bei Siebtklässlern im Mathematikunterricht**

**Stella Pede, Rita Borromeo Ferri**  
Universität Kassel, Deutschland

-Theoretischer Hintergrund-

Im Rahmen des LOEWE-Schwerpunktes „Wünschenswerte Erschwernisse beim Lernen“ beschäftigt sich das LIMIT-Projekt der Universität Kassel mit dem verschachtelten Lernen im Mathematikunterricht. Das Konzept des verschachtelten Lernens steht im Kontrast zum sogenannten geblockten Lernen (Rohrer & Taylor, 2007). Während die Lerninhalte beim geblockten Lernen

nacheinander behandelt werden, werden sie beim verschachtelten Lernen abwechselnd gelernt und geübt, was eine Erschwernis für die Lernenden darstellt (Bjork & Bjork, 2011). In vielen Lehrbüchern für den Mathematikunterricht werden Inhalte geblockt angeordnet. Das Thema „Zuordnungen“ wird bspw. vielfach so behandelt, dass zuerst Aufgaben zu proportionalen Zuordnungen, anschließend Aufgaben zu antiproportionalen Zuordnungen und zum Schluss Aufgaben zu sonstigen Zuordnungen vorgesehen sind.

In diesem Vortrag wird eine Interventionsstudie aus dem Projekt LIMIT-Sekundarstufe zum Thema „Zuordnungen“ im Fach Mathematik vorgestellt, die in Haupt- und Realschulklassen der Jahrgangsstufe 7 durchgeführt wurde.

-Fragestellung-

Die LIMIT-Studie in der Sekundarstufe I widmet sich u.a. der Frage, ob sich Unterschiede in der Entwicklung des prozeduralen und des konzeptuellen Wissens zwischen Schüler/-innen zeigen, die geblockt oder verschachtelt lernen.

-Methode-

Insgesamt 124 Probanden haben an einer achtstündigen Unterrichtseinheit zum Thema „Zuordnungen“ teilgenommen. Hierbei wurde eine randomisierte Zuweisung zu den beiden Bedingungen vorgenommen, sodass jeweils eine Klassenhälfte geblockt und die andere Klassenhälfte verschachtelt lernte.

Als abhängige Variablen wurden die Leistungen der Schüler/-innen bei der Bearbeitung von Aufgaben, welche prozedurales Wissen erforderten, und die Leistungen bei der Bearbeitung von Aufgaben, welche konzeptuelles Wissen verlangten, erfasst. Die Leistungen wurden vor der Unterrichtseinheit, unmittelbar danach (Posttest) und zusätzlich drei Wochen (Follow-Up-Test 1) sowie zehn Wochen (Follow-Up-Test 2) später erhoben. Die Lernentwicklung über die Messzeitpunkte wurde mittels Varianzanalyse mit Messwiederholungsdesign ermittelt.

-Ergebnisse-

In den Untersuchungsergebnissen zeigt sich zwischen den geblockt und verschachtelt Lernenden über alle Messzeitpunkte hinweg kein Unterschied in der Entwicklung des prozeduralen Wissens (Zeit x Lernbedingung:  $F(3,270) = 0.568$ ,  $p = 0.636$ ,  $\eta^2 = 0.006$ ). Dagegen lässt sich jedoch ein signifikanter Unterschied in der Entwicklung des konzeptuellen Wissens zugunsten der verschachtelt Lernenden zwischen Posttest und Follow-Up Test 2 nachweisen. Die Leistungen der Schüler/-innen in der verschachtelten Bedingung entwickeln sich günstiger als in der geblockten Bedingung (Zeit x Lernbedingung:  $F(2,178) = 3.348$ ,  $p = 0.037$ ,  $\eta^2 = 0.036$ ).

## **Verschachteltes Lernen von Subtraktionsstrategien im Mathematikunterricht der Grundschule**

**Katharina Werker, Julia Arend, Frank Lipowsky**

Universität Kassel, Deutschland

-Theoretischer Hintergrund-

Verschachteltes Lernen meint die abwechselnde Behandlung von Inhalten (ABCBACACB), während diese beim geblockten Lernen nacheinander eingeführt und geübt werden (AAABBBCCC). Verschachteltes Lernen stellt ein wünschenswertes Erschwernis für die Lernenden dar, da durch den ständigen Wechsel der Inhalte kognitiv anspruchsvolle Aktivitäten evoziert werden (Bjork & Bjork, 2011; Rohrer & Taylor, 2007).

Bisherige Laborstudien zum verschachtelten Lernen zeigen positive Effekte auf den Lernerfolg und das Behalten (Dunlosky et al., 2013; Rohrer & Taylor, 2007). Als mögliche Erklärung für die positiven Effekte des verschachtelten Lernens wird die Diskriminations-/Kontrasthypothese diskutiert (Birnbauer et al., 2013; Kang & Pashler, 2012). Durch die abwechselnde Behandlung unterschiedlicher Inhalte werden die Lernenden beim verschachtelten Lernen angeregt, zwischen Kategorien zu vergleichen und zu diskriminieren.

Bezogen auf das Verschachteln von Subtraktionsstrategien ist anzunehmen, dass Lernende hierbei nicht nur prozedurales, sondern auch konditionales Wissen erwerben, indem sie die Anwendungsbedingungen der Strategien kennen und unterscheiden lernen (Dunlosky et al., 2013). Dies ist relevant, da Drittklässler nach der Einführung des schriftlichen Normalverfahrens beim Rechnen selten ihr Wissen über Zahlbeziehungen nutzen, sondern dieses Verfahren ungeachtet der Aufgabencharakteristika häufig als Universalstrategie anwenden. Der Strategieeinsatz ist damit wenig flexibel und adaptiv (Selter, 2001).

Insbesondere Aufgabenstellungen, welche die Schüler/-innen explizit zum Vergleichen und zur Abgrenzung von Strategien sowie zur Reflexion über die Anwendungsbedingungen anregen, scheinen unterstützend auf Diskriminations- und Vergleichsprozesse zu wirken (Ziegler & Stern, 2014, 2016).

-Fragestellung-

Die Grundschulstudie des Projekts LIMIT (Verschachteltes Lernen im Mathematikunterricht) untersucht, ob sich verschachteltes Lernen positiv auf die Korrektheit von Aufgabenlösungen, auf die adaptive Wahl von Subtraktionsstrategien und auf die Strategieflexibilität von Drittklässlern auswirkt.

-Methode-

Diese Fragestellung wird in einem experimentellen Studiendesign untersucht. Hierzu wurden eine verschachtelte und eine geblockte Lernumgebung entwickelt und im Mathematikunterricht von 12 Grundschulklassen des 3. Schuljahres (N=236) implementiert. Hierbei wurde jede Klasse geteilt, sodass eine Klassenhälfte die ausgewählten fünf Subtraktionsstrategien (Stellenweise, Schrittweise, Ergänzen, Hilfsaufgabe, schriftliches Normalverfahren) geblockt erlernte, die andere Klassenhälfte verschachtelt. Die Zuteilung der Schüler/-innen zu den beiden Bedingungen erfolgte randomisiert.

Die Intervention umfasste 14 Unterrichtsstunden und wurde von geschulten Mitarbeiter/-innen durchgeführt, die jeweils beide Bedingungen unterrichteten. Die Schüler/-innen der verschachtelten Bedingung wurden fortlaufend explizit angeregt, die Strategien miteinander zu vergleichen und hinsichtlich ihrer Adaptivität zu reflektieren. In der geblockten Bedingung wurden die Subtraktionsstrategien nacheinander behandelt, wobei keine Vergleiche zwischen den Strategien stattfanden.

Als abhängige Variablen werden die Korrektheit der Aufgabenlösungen, die Strategieflexibilität (d.h. das Repertoire an eingesetzten Strategien) und die Adaptivität (d.h. die Fähigkeit, eine Aufgabe mit einer effizienten Strategie zu lösen) untersucht. Diese wurden mit einem Rechenstrategietest zu vier Messzeitpunkten erfasst (eine Woche vor der Intervention, unmittelbar sowie eine und vier Wochen nach der Intervention).

Für die Auswertung der Strategietests wurde ein Kodiermanual entwickelt, anhand dessen geschulte Kodierer/-innen unabhängig voneinander die verwendeten Strategien kodierten ( $\kappa \geq .90$ ). Weiterhin wurde für jede Aufgabe die Adaptivität der

Strategien unter Berücksichtigung der Fehleranfälligkeit, der Anzahl der Lösungsschritte und des kognitiven Aufwands eingeschätzt ( $\kappa \geq .70$ ).

Um die Lernentwicklung im gesamten Untersuchungszeitraum für die drei abhängigen Variablen zu prüfen, werden Varianzanalysen mit Messwiederholung gerechnet.

-Ergebnisse-

Für die Entwicklung der Korrektheit ergibt sich kein signifikanter Unterschied zwischen beiden Bedingungen. Betrachtet man die Entwicklung der Strategieflexibilität, zeigt der Vergleich beider Gruppen, dass die verschachtelt lernenden Schüler/-innen nach der Intervention mehr Strategien nutzen als die geblockt lernenden. Der Unterschied ist mit einer kleinen Effektstärke signifikant. Ein deutlicher Unterschied (starker Effekt) zwischen beiden Bedingungen lässt sich für die Entwicklung der Adaptivität nachweisen. Die Schüler/-innen der verschachtelten Gruppe weisen nach der Intervention eine signifikant höhere Adaptivität beim Lösen der Aufgaben auf als die Vergleichsgruppe.

Die Ergebnisse der Studie zeigen damit, dass verschachteltes Lernen mit instruktionaler Anregung von Vergleichsprozessen ein hohes Potential hat, um die Strategieflexibilität und –adaptivität von Drittklässlern zu fördern.

## B13

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H118

### Das Verhältnis von Intelligenz, Vorwissen und Problemlösen in unterschiedlichen Bildungskontexten

Chair(s): **Felix Walker** (Technische Universität Kaiserslautern, Deutschland)

DiskutantIn(nen): **Detlev Leutner** (Universität Duisburg-Essen)

#### \*Problemstellung/Ausgangslage\*

Problemlösefähigkeit ist eine zentrale Zielgröße schulischer, akademischer und beruflicher Bildung. Die konzeptionelle Abgrenzung zu Intelligenz und die Bedeutung von Vorwissen für erfolgreiches Problemlösen sind dabei Bestandteil anhaltender Kontroversen (z. B. Süß, 1999). Die Operationalisierung von Problemlösefähigkeit und Intelligenz sowie deren empirische Zusammenhänge wurden in der Fachöffentlichkeit zuletzt wieder intensiv diskutiert (z.B. Funke & Dörner, 2017; Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2017). Während einige Forscher hohe Zusammenhänge postulierten und fanden, ergaben andere Studien nur geringe Zusammenhänge (für einen Überblick siehe Stadler, Becker, Gödker, Leutner, & Greiff, 2015). Leutner (2002) formulierte in Anlehnung an Elshout (1987) und Raaheim (1988) die „Elshout-Raaheim-Hypothese“, wonach Intelligenz immer dann bedeutsam beim Lösen von Problemen wird, wenn eine Person ein mittleres Vorwissensniveau bezüglich des Problemkontexts aufweist (vgl. Leutner, 2002, S. 686). Dieser mit steigendem Vorwissen umgekehrt U-förmige Zusammenhang zwischen Intelligenz und Problemlösen ist durchaus plausibel: Ist das Vorwissen sehr hoch, stellt sich die Problemsituation ggf. gar nicht als Problem, sondern lediglich als Aufgabe dar. Ist das Vorwissen jedoch so gering, dass bspw. schon ein Grundverständnis für die Problemsituation fehlt, hilft auch Intelligenz kaum weiter. Wenngleich diese Argumentation sehr plausibel ist, zeigen sich in empirischen Studien zum Teil deutlich abweichende Befunde (z.B. Goode & Beckmann, 2010).

#### \*Zielsetzung/Zuschnitt des Symposiums\*

Die Beiträge des Symposiums widmen sich der Frage, inwiefern sich der in der „Elshout-Raaheim-Hypothese“ postulierte Zusammenhang zwischen Intelligenz und Problemlösen in empirischen Studien in unterschiedlichen Bildungskontexten wiederfindet. In den Beiträgen werden diesbezügliche Forschungsergebnisse zum Verhältnis von Intelligenz und fachübergreifendem/domänenfreiem oder fachbezogenen/domänenspezifischen Problemlösen aus Studien mit Schülern/innen, Studierenden und Auszubildenden vorgestellt und diskutiert.

Matthias Stadler und Samuel Greif (Universität Luxemburg) wenden sich dem domänenübergreifendem komplexen Problemlösen zu, das auf Basis des MicroDYN Ansatzes computerbasiert erfasst wurde, und betrachten in einer Stichprobe von 1.196 finnischen Schüler/innen das Verhältnis von Vorwissen (operationalisiert als das Wissen über die Vary-One-Thing-At-A-Time-Strategie), Intelligenz und Problemlösen.

Jennifer Stemmann und Martin Lang (Universität Duisburg- Essen) untersuchen aus einer technikkdidaktischen Perspektive das Problemlösen von 147 Technik-Studierenden. In der Computersimulation geht es um den Umgang mit 15 technischen Alltagsgeräten. Untersucht wird auch hier der Zusammenhang zwischen Vorwissen (operationalisiert als eingeschätzte Erfahrung mit dem Alltagsgerät), Intelligenz und dem Lösen von Problemen mit technischen Alltagsgeräten.

Felix Walker (TU Kaiserslautern) geht auf Basis von 209 Auszubildenden des Berufes zum/zur Elektroniker/in für Automatisierungstechnik der Fragen nach, inwieweit der Zusammenhang zwischen Intelligenz und domänenspezifischen Problemlösen (operationalisiert als Fehlersuche in automatisierten Systemen) in Abhängigkeit vom Vorwissen (operationalisiert als Fachwissen über automatisierte Systeme) dem Verlauf der Elshout-Raaheim-Hypothese entspricht.

Andreas Rausch (Universität Mannheim) und Kristina Kögler (Goethe-Universität Frankfurt) analysieren Daten von 780 kaufmännischen Auszubildenden hinsichtlich des Zusammenhangs von fluider Intelligenz, betriebswirtschaftlichem Vorwissen (erfasst anhand eines Multiple-Choice-Tests) und domänenspezifischer Problemlösekompetenz, die anhand der Bearbeitung dreier komplexer Problemszenarien in einer computerbasierten Bürosimulation erfasst wurde. Bei der Prüfung der Elshout-Raaheim-Hypothese werden ergänzend auch die Einflüsse von Numeracy und Literacy untersucht.

Detlev Leutner (Universität Duisburg-Essen) wird die vier Beiträge des Symposiums zusammenfassend diskutieren.

#### *Beiträge des Symposiums*

### **Wissen ist gut, mehr wissen ist nicht immer besser: Der non-lineare Zusammenhang zwischen komplexem Problemlösen, Intelligenz und Vorwissen**

**Matthias Stadler, Samuel Greiff**  
Universität Luxemburg

Der Zusammenhang zwischen der Leistung in klassischen Intelligenztests und der Fähigkeit komplexe Probleme zu lösen bleibt ein heiß diskutierter Forschungsgegenstand (z.B. Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2017). Während einige Forscher hohe Zusammenhänge postulierten und fanden, ergaben andere Studien nur geringe Zusammenhänge (für einen Überblick siehe Stadler, Becker, Gödker, Leutner, & Greiff, 2015). Um diese widersprüchlichen Befunde zum Zusammenhang zwischen den Leistungen in den beiden Aufgabenformen zu erklären, postulierte Leutner (2002) die Elshout-Raaheim-Hypothese (vgl., Elshout, 1987; Raaheim, 1988). Nach der Elshout-Raaheim-Hypothese hängt der Zusammenhang zwischen Intelligenz und Problemlösefähigkeit umgekehrt u-förmig vom verfügbaren Vorwissen ab, dass für die Problemlösung eingesetzt werden kann. Dabei kann die individuelle Intelligenz optimal eingesetzt werden wenn mittelviel Vorwissen zur Verfügung steht. Bei zu wenig Vorwissen kann die individuelle Intelligenz nicht genutzt werden, bei zu viel Vorwissen muss sie nicht genutzt werden.

In dieser Studie wurde der moderierende Einfluss von Vorwissen über die Vary-One-Thing-At-A-Time (VOTAT) Strategie auf den Zusammenhang zwischen Intelligenz und der Leistung in komplexen Problemlöseaufgaben untersucht. Basierend auf der Elshout-Raaheim-Hypothese wurde eine non-lineare Moderation erwartet. Eine bevölkerungsrepräsentative Stichprobe von N = 1196 finnischen Schüler\*innen bearbeitete eine Reihe von komplexen Problemen (nach dem MicroDYN Ansatz; Greiff, Fischer, Stadler, & Wüstenberg, 2015), eine Aufgabe zur Erfassung der numerischen Verarbeitungskapazität und eine Aufgabe zur Erfassung des Wissens über VOTAT. Die Moderationshypothese wurde mittels manifester hierarchischer Regressionsanalysen überprüft.

Die Ergebnisse der Analysen bestätigen die Annahmen der Elshout-Raaheim-Hypothese. Das Wissen der Schüler\*innen über VOTAT erwies sich als non-linearer Moderator des Zusammenhangs zwischen komplexem Problemlösen und Intelligenz (R<sup>2</sup> =

.26). Dabei war der Zusammenhang für die Schüler\*innen mit mittlerem Vorwissen ( $r = .32$ ) am stärksten und am schwächsten für die Schüler\*innen mit sehr geringem ( $r = .23$ ) und sehr hohem Vorwissen ( $r = .24$ ).

Diese Studie trägt weiter dazu bei die scheinbar widersprüchlichen Befunde zum Zusammenhang zwischen Intelligenz und komplexem Problemlösen zu erklären. Besonders die Rolle von Vorwissen über VOTAT war bisher nicht ausreichend gut verstanden und wurde bisher vorwiegend theoretisch diskutiert (z.B. Kretschmar, 2017). Daher bereichert diese empirische Untersuchung die aktuelle Diskussion um weitere wichtige Erkenntnisse.

## **Intelligenzeinfluss auf den Umgang mit technischen Alltagsgeräten in Abhängigkeit vom Vorwissen**

**Jennifer Stemann, Martin Lang**

Universität Duisburg-Essen

Die Implementierung von Mikroelektronik und Digitaltechnik in Alltagsgegenstände setzt vermehrt die Interaktion mit solchen Geräten voraus. Wenn technische Geräte neu sind oder nicht erwartungsgemäß funktionieren, wird ihre Nutzung zu einem Problem. Der Erwerb fehlenden Bedienungswissens ist durch den Bedeutungsverlust von Bedienungsanleitungen vielfach nur durch zielgerichtetes Ausprobieren möglich (Tully, 2003). Im Gegensatz zum Instruktionslernen setzt die selbstständige Aneignung von Eingriffswissen vermehrt globale, kognitive Fähigkeiten voraus. Vor allem fluide Intelligenz (Cattell, 1971) determiniert den Wissenserwerbsprozess beim Problemlösen. In dem mehrstufigem Wissenserwerbsprozess, bestimmt Intelligenz Prozesse des Bewusstwerdens von Informationen und des Haltens unterschiedlicher Informationsaspekte im Arbeitsgedächtnis (Rost, 2009). Je lückenhafter die Wissensstruktur, desto mehr Informationen müssen verarbeitet werden; durch die Begrenzung der Arbeitsgedächtniskapazität, die eng mit fluider Intelligenz zusammenhängt (Kyllonen & Christal 1990), ist die Leistungsfähigkeit allerdings limitiert. Je weniger Informationen aufgrund vorhandenen Wissens verarbeitet werden müssen, desto geringer ist die Arbeitsbelastung und damit der Intelligenzeinfluss auf die Problemlöseleistung (Spada & Mandl, 1988). Zwischen dem Intelligenzeinfluss und dem Vorwissen kann jedoch kein linear abfallender Zusammenhang angenommen werden, denn wenn kein problemlösungsrelevantes Wissen existiert, fehlen Anknüpfungspunkte für neues Wissen (Beckmann & Goode, 2014). Ist das benötigte Wissen umfangreich und stark fachspezifisch, lässt es sich bei einer limitierten Bearbeitungszeit auch unter Intelligenzeinfluss schwer aneignen. Bei mittlerem Vorwissen kann dieses für den weiteren Lernprozess genutzt werden, so dass hier der Zusammenhang zwischen Intelligenz und Problemlöseleistung maximal wird (Elshout, 1987; Raaheim, 1988). Dieser als Elshout-Raaheim-Hypothese bezeichnete umgekehrt U-förmige Zusammenhang wurde von Leutner (2002) in einem fachübergreifenden Problemkontext bestätigt. Ob der Zusammenhang zwischen Intelligenz und Problemlöseleistung auch im Umgang mit technischen Alltagsgeräten abhängig vom Vorwissen variiert, ist die Fragestellung dieses Beitrags. Angenommen wird, dass der Intelligenzeinfluss bei mittlerem Vorwissen maximal wird.

Die Problemlösekompetenz im Umgang mit technischen Alltagsgeräten wurde mit einem Computertest bestehend aus 15 simulierten Geräten erfasst. Jedes Testitem beinhaltet zunächst eine Explorationsphase, in der das entsprechende Gerät ausprobiert und Bedienungswissen erlangt werden kann. Es folgt eine Steuerungsphase, in der das Gerät nach vorgegebener Problemstellung in einen eindeutigen Systemzustand gebracht werden muss. Das gerätespezifische Vorwissen vor der Exploration lässt sich nicht vollständig erfassen. Relevant für den Umgang mit Alltagsgeräten ist vor allem das sogenannte Eingriffswissen, das als prozedurales Wissen repräsentiert ist. Da für den Erwerb solcher Prozeduren Übung nötig ist, die im Alltag durch Erfahrungen im Umgang mit den Geräten erworben wird, lässt sich diese Erfahrung durch Single-Items erfassen. Die Probanden schätzen dabei auf einer dreistufigen Ratingskala ihre Erfahrungen für jedes der simulierten Geräte ein. Fluide Intelligenz wurde mit der Kurzform des CFT-20R (Weiß, 2006) erfasst. Die analysierten Daten stammen aus einer Untersuchung mit 147 Technik-Studierenden ( $w=52, m=95$ ), deren Durchschnittsalter  $MW=26.6$  Jahre ( $SD=5.89$  Jahre) beträgt.

Der Zusammenhang zwischen Intelligenz und der Steuerungsleistung beträgt in der Gesamtstichprobe  $r=.49$  ( $p<.001$ ). Die Selbsteinschätzung der Probanden bezüglich ihrer Vorerfahrungen wurde genutzt, um die Stichprobe auf Itemebene in drei Gruppen (wenig, mittleres und viel Vorwissen) einzuteilen. Für jede Gruppe und Item wurde die Korrelation zwischen Intelligenz und Steuerungsleistung berechnet und anschließend geprüft, ob die Gruppen aus einer Grundgesamtheit stammen. Auf Itemebene konnte kein signifikanter Unterschied zwischen den Korrelationskoeffizienten festgestellt werden. Auch bei Bildung eines Gesamtscores (aus den Einzelitems zum Vorwissen), der in drei gleich große Abschnitte eingeteilt wurde, zeigte ein Vergleich der Korrelationen, dass sich der Zusammenhang zwischen Intelligenz und der Steuerungsleistung nicht für Personen mit unterschiedlichen Vorerfahrungen unterscheidet. Der Zusammenhang in der Gruppe mit wenig Vorerfahrung beträgt  $r=.37$  ( $p=.03$ ,  $n=35$ ), in der Gruppe mit mittlerer Vorerfahrung  $r=.47$  ( $p<.005$ ,  $n=88$ ) und in der Gruppe mit viel Vorerfahrung  $r=.57$  ( $p=.01$ ,  $n=18$ ). Im letzten Schritt wurde die Explorationsvollständigkeit als Vorwissensvariable genutzt und die Probanden ebenfalls in drei Gruppen eingeteilt. Auch hier unterscheiden sich die Korrelationskoeffizienten nicht signifikant voneinander, so dass die Elshout-Raaheim-Hypothese hier nicht bestätigt werden kann.

## **Das Verhältnis von Intelligenz, Fachwissen und Problemlösen bei Elektronikern/innen für Automatisierungstechnik**

**Felix Walker**

Technische Universität Kaiserslautern, Deutschland

\*Theoretischer Hintergrund\*

Das Lösen domänenspezifischer bzw. fachspezifischer Probleme ist eine zentrales Ziel technischer beruflicher Erstausbildung (vgl. KMK, 2003). Aktuelle Befunde legen insbesondere nahe, dass für das Lösen fachspezifischer Probleme maßgeblich das vorhandene Fachwissen über das technische (z.T. defekte) System entscheidend für den Problemlöseerfolg ist (z.B. im Überblick Walker, Link & Nickolaus, 2016). Unklar ist der Einfluss bzw. die Rolle von Intelligenz im Rahmen des fachspezifischen Problemlösens in der beruflichen Erstausbildung. In den Berufen des kfz-Mechatronikers, des Elektroniklers für Energie- und Gebäudetechnik (vgl. Abele et al., 2012) als auch in dem Beruf des Elektroniklers für Automatisierungstechnik (vgl. Walker et al., 2015; Walker, Link, van Waveren et al., 2016) konnte die Intelligenz keine signifikanten Beitrag zum Lösen fachspezifischer Probleme leisten, nachdem das Fachwissen in das Regressionsmodell integriert wurde. Dem gegenüber stehen Befunde aus der bautechnischen Erstausbildung (vgl. Petsch, Norwig & Nickolaus, 2015) in dem Intelligenz auch nach Einbezug des Fachwissens das Lösen bautechnischer Probleme signifikant erklären konnte.

Ein mögliche Erklärung dieser inhomogenen Befundlage könnte die von Leutner (2002) formulierte „fuzzy-relationship“ bzw. die Elshout-Raaheim-Hypothese liefern. Gegenwärtig wurde kein Versuch unternommen die Übertragbarkeit bzw. den Geltungsbereich der Hypothese auf die berufliche technische Erstausbildung zu prüfen.

\*Fragestellung\*



Vor diesem Hintergrund prüft der Beitrag in wie fern der invertierte U-förmige Zusammenhang von Intelligenz und fachspezifischen Problemlösen in Abhängigkeit des Vorwissens bei Elektronikern für Automatisierungstechnik zutrifft. Unter Problemlösen wird die kognitive Voraussetzung dafür verstanden, dass Auszubildenden die Fehlerursache für einen Defekt an einem steuerungstechnischen System identifizieren und Behebungsvorschläge benennen können. Konkret wird der Frage nachgegangen wie der Zusammenhang zwischen Intelligenz und fachspezifischen Problemlösen bei Elektronikern für Automatisierungstechnik mit niedrigem, mittlerem und hohem Fachwissen (im Bereich Steuerungstechnik) ausfällt.

**\*Methode\***

Die Hypothesenprüfung erfolgt anhand von n=209 Daten von Elektronikern für Automatisierungstechnik, welche überwiegend männlich (89.1%), im Mittel M= 20.63, SD= 2,67 Jahre alt sind und sich im 3. oder im 4. Ausbildungsjahr (64.2%/35.8%) befanden.

64.2% besitzen als höchsten Schulabschluss einen mittleren Bildungsabschluss (Hauptschule 8.3%; Gymnasium 30.2%). 71.7% absolvieren Ihre Ausbildung als Industrieverberuf. Als Maß für die Intelligenz kam der CFT-20-R (vgl. Weiß, 2006), für das Fachwissen ein Papier und Bleistift-Test (vgl. van Waveren & Nickolaus, 2015) und für das Problemlösen eine authentische Computersimulation (vgl. Nickolaus & Walker, 2016) zum Einsatz. Für die Hypothesenprüfung wurden die Auszubildenden, um eine gleich große Gruppenstärke zu gewährleisten, in Perzentile des Fachwissens im Bereich Steuerungstechnik aufgeteilt.

**\*Ergebnisse\***

Der gegenwärtige Stand der Analysen spricht gegen den invertierten U-förmigen Zusammenhang von Intelligenz und Problemlösen. In den beiden unteren Perzentilen des (0-25% & 25%-50%), d.h. bei Auszubildenden mit geringem Fachwissen im Bereich Steuerungstechnik sinkt die Korrelation von  $r = .3$  auf  $r = .2$  ab. Den geringsten Zusammenhang findet sich im mittleren Perzentil (50%-75%) mit  $r = .01$ . Bei Elektronikern für Automatisierungstechnik mit dem höchsten Fachwissen (75%-100%) korreliert Intelligenz und Problemlösen ähnlich hoch wie in der Gruppe mit dem geringsten Fachwissen.

Der Beitrag liefert erstmals Ergebnisse zur Übertragbarkeit der Elshout-Raaheim-Hypothese im Bereich der technischen Berufsausbildung und bietet die Gelegenheit mögliche Ursachen zu diskutieren.

**Zum Einfluss von Fachwissen, Intelligenz, Literacy und Numeracy auf die domänenspezifische Problemlösekompetenz bei Kaufleuten**

**Andreas Rausch<sup>1</sup>, Kristina Kögler<sup>2</sup>**

<sup>1</sup>Universität Mannheim, <sup>2</sup>Goethe-Universität Frankfurt

**Theoretischer Hintergrund**

Problemlösekompetenz ist wichtig im Arbeitsleben, da Routinetätigkeiten zunehmend ausgelagert oder automatisiert werden (Frey & Osborne, 2013). Probleme zeichnen sich dadurch aus, dass das verfügbare Handlungswissen zunächst nicht ausreicht, um einen Ausgangszustand in einen erwünschten Zielzustand zu überführen (z. B. Duncker, 1945). Rausch und Wuttke (2016) definieren vier kognitive Facetten domänenspezifischen Problemlösens: (1) Identifikation von Handlungsbedarfen und Informationsquellen, (2) Verarbeitung domänenspezifischer Informationen, (3) Treffen einer begründeten Entscheidung und (4) angemessenes Kommunizieren der Entscheidung. Dieser Ansatz entspricht in hohem Maße Weinerts (2001) Kompetenzverständnis, der ähnliche Kompetenzstrukturen über verschiedene Domänen hinweg postuliert, auch wenn sich die Handlungen in verschiedenen Domänen stark unterscheiden. Individuelle Unterschiede zur Erklärung erfolgreichen Problemlösens zu finden, ist ein wichtiges Ziel der Problemlöseforschung (Dörner & Funke, 2017).

**Fragestellungen**

Im vorliegenden Beitrag werden Fachwissen ('knowledge-is-power-hypothesis'; Hambrick & Engle, 2002), Intelligenz ('why g matters'; Gottfredson, 1997), Numeracy und Literacy als mögliche Einflussfaktoren untersucht. Es wird erwartet, dass Fachwissen ein mittelstarker Prädiktor der domänenspezifischen Problemlösekompetenz ist, während Intelligenz einen nur geringen Einfluss hat (vgl. Rausch, 2017). Ferner geht die so genannte „Elshout-Raaheim-Hypothese“ (Leutner, 2002) davon aus, dass Intelligenz einen höheren Zusammenhang mit Problemlösen aufweist, wenn ein mittleres Vorwissensniveau gegeben ist, aber schwächere Zusammenhänge bei niedrigem und hohem Vorwissen.

**Methode**

An der Querschnittstudie beteiligten sich 780 kaufmännische Auszubildenden im zweiten oder dritten Jahr ihrer Ausbildung, die zuvor ihr schriftliches Einverständnis zur freiwilligen Teilnahme erteilt haben. Domänenspezifische Problemlösekompetenz wurde anhand dreier komplexer kaufmännischer Szenarien ermittelt, die in einer computerbasierten Bürosimulation in jeweils 30 Minuten zu bearbeiten waren. Die Auswertung des Testverhaltens erfolgte in einem dreischrittigen Prozess ähnlich wie bei Bennett et al. (2003) und resultierte in zufriedenstellenden EAP/PV Reliabilitäten für die vier o. g. Kompetenzfacetten und einen die Facetten umfassenden Gesamt-Score (Rausch et al, 2016). Das für die Szenarien relevante Fachwissen wurde mit einem 14 Items umfassenden Multiple-Choice-Test erfasst. Für die Messung fluider Intelligenz wurde die deutsche Version des Culture Fair Test (CFT) von Weiss (2006) verwendet. Ferner wurden Literacy und Numeracy mit Hilfe des adaptiven Tests ‚MAK-adapt‘ von Ziegler et al. (2016) erhoben. Den letztgenannten Test konnte aus technischen und administrativen Gründen nur knapp die Hälfte einer Teilnehmenden bearbeiten.

**Ergebnisse**

Tabelle 1 zeigt die Ergebnisse einer Regressionsanalyse zur Erklärung der Problemlösekompetenz.

Tabelle 1. Ergebnisse der multiplen linearen Regression zur Erklärung der Problemlösekompetenz (Gesamt-Score)

Variable B SEB  $\beta$  t p

(Constant) 292.52 36.48 8.018 .000

Fachwissen 11.96 1.76 .344\*\* 6.784 .000

Fluide Intelligenz 1.01 .34 .153\*\* 2.993 .003

Numeracy 11.56 5.14 .120\* 2.248 .025

Literacy 24.58 5.32 .244\*\* 4.622 .000

Note. R2 = .381; adj. R2 = .373 (df = 289, F = 43.91, p < .001); \* < .05; \*\* < .01

Wie schon bei Rausch (2017) erweist sich das Fachwissen als stärkster Prädiktor. Insgesamt erklärt das Modell immerhin 37 % der Varianz der Problemlösekompetenz.

Tabelle 2 gibt die Ergebnisse zur Überprüfung der „Elshout-Raaheim-Hypothese“ wieder, indem die Korrelation zwischen Intelligenz und Problemlösen für drei Gruppen unterschiedlichen Vorwissens dargestellt werden. Explorativ werden auch die Zusammenhänge für Numeracy und Literacy überprüft.

Tabelle 2. Korrelationskoeffizienten des Zusammenhangs zwischen Prädiktoren und Problemlösekompetenz gruppiert nach Fachwissensniveau

Teilstichprobe der Industriekaufleute

Intelligenz x Problemlösen (insgesamt (n = 487) .286\*\*)

Niedriges Fachwissen Mittleres Fachwissen Hohes Fachwissen

(n = 169) .226\*\* (n = 144) .176\*\* (n = 174) .265\*\*

Numeracy x Problemlösen (insgesamt (n = 194) .299\*\*)

(n = 68) .145 (n = 55) .119 (n = 71) .371\*\*

Literacy x Problemlösen (insgesamt (n = 191) .313\*\*)

(n = 66) .320\*\* (n = 55) .241 (n = 70) .228

Note. Pearson's correlation coefficient; \*\* p < .01; \* p < .05

Die Ergebnisse widersprechen der Elshout-Raaheim-Hypothese, weisen aber zugleich interessante Einflüsse von Numeracy und Literacy auf, die weiter zu analysieren und zu diskutieren sind.

## B14

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H117

### Von einer Bilanz- zur Prozessperspektive in der Hochschulforschung: Die Bedeutsamkeit psychologischer Faktoren für Studienerfolg und -misserfolg

Chair(s): **Stefan Janke** (Universität Mannheim, Deutschland), **Martin Daumiller** (Universität Augsburg)

DiskutantIn(nen): **Klaus-Peter Wild** (Universität Regensburg)

In der Vergangenheit wurde der Erfolg von Hochschullehre häufig unter einer Input-Output-Perspektive beurteilt. So wurde bewertet, in welchem Ausmaß Studienerfolg im Verhältnis zu den investierten Ressourcen steht (u.a. Tam, 2001). Diese Perspektive ermöglicht es, Hochschullehre an ihrem tatsächlichen Erfolg zu messen, lässt allerdings offen, was zu diesem Erfolg führt. Um solche Wirkmechanismen zu verstehen und neue Maßnahmen zur Gestaltung der Lehre und des Lernkontextes entwickeln zu können, ist entsprechend die Aufklärung der zu Grunde liegenden Prozesse von zentraler Bedeutung (Bendermacher, et al., 2017). Im Rahmen des Symposiums werden Forschungsarbeiten vorgestellt, die verschiedene psychologische Prozessvariablen untersuchen, die zum Studienerfolg (beispielsweise operationalisiert als Studienleistung, aber auch Studienzufriedenheit) oder -misserfolg (beispielsweise operationalisiert als Studienabbruchintention) beitragen können. Im ersten Beitrag wird dazu grundlegend ein Prozessmodell der Herausbildung der Intention zum Studienabbruch vorgestellt. Das vorgeschlagene Phasenmodell verdeutlicht, dass Studienabbruch als Prozess verstanden werden kann und eine Abbruchentscheidung mit dispositionalen und motivationalen Variablen zusammenhängt. Darauf aufbauend wird im zweiten Beitrag die Bedeutung dispositiver Variablen für das Erreichen von Studienerfolg genauer untersucht. Dabei wird aufgezeigt, dass sich der relativ neu eingeführte und kontrovers diskutierte Persönlichkeitsfaktor „Grit“ (grob übersetzt: Zielstrebigkeit), welcher vorgeblich Studienerfolg beeinflusst (Duckworth, et al., 2007), schwer von dem etablierten psychologischen Konstrukt Selbstkontrolle abgrenzen lässt und auch nicht mit objektiven Studienleistungen oder subjektiv wahrgenommenem Studienerfolg in Zusammenhang steht. In dem dritten und vierten Beitrag werden anschließend motivationale Variablen fokussiert betrachtet. Dabei baut die Betrachtung dieser Variablen auf der Selbstbestimmungstheorie der Motivation auf, die postuliert, dass insbesondere ein Fokus auf persönliches Wachstum und eine interessen geleitete Studienorientierung zu langanhaltender Studienmotivation und -zufriedenheit beiträgt (Deci & Ryan, 2000; Schmuck, et al., 2000). Diese Annahme wird im dritten Beitrag auf den Prüfstand gestellt, indem die Bedeutung intrinsischer Studienwahlmotivation (Studienaufnahme auf Basis persönlicher Interesse) im Vergleich zu extrinsischer Studienwahlmotivation (Studienaufnahme auf Basis beruflicher und finanzieller Perspektive) in einer Längsschnittstudie untersucht wird. Die Ergebnisse zeigen, dass eine Studienaufnahme aus intrinsischer Motivation positiv im Zusammenhang zur Lernfreude nach dem ersten akademischen Jahr steht. Extrinsische Studienwahlmotivation scheint die Studienzufriedenheit hingegen zu untergraben und steht positiv im Zusammenhang mit der Studienabbruchintention. Im abschließenden, vierten Beitrag wird schließlich im Detail beleuchtet, wie das Erleben von Autonomie, Kompetenz und sozialer Eingebundenheit an der Universität im Kontext der persönlichen Bedeutsamkeit dieser Grundbedürfnisse im Zusammenhang mit Studienerfolg steht. Dabei zeigt sich, dass insbesondere das Kompetenzerleben Studienleistung positiv vorhersagt. Weiterhin erwiesen sich Kompetenz- und Autonomieerleben als relevant für die Stärke der individuellen Studienabbruchintention. Zusammengefasst verdeutlichen die Beiträge, dass psychologische Prozessmodelle vielversprechend zur Erklärung von Studienerfolg und -misserfolg sind. Darüber hinaus verweisen sie auf erste praktische Möglichkeiten zur Beeinflussung des Studienerfolgs. So geben beispielhaft die Befunde des letzten Beitrags erste Hinweise darauf, welche Umweltfaktoren sich möglicherweise präventiv beeinflussen lassen könnten (Unterstützung des Autonomie- und Kompetenzerlebens an der Universität), um zu verhindern, dass Studierende die erste Phase (Bildung einer Abbruchintention aufgrund der Wahrnehmung einer geringen Passung) des im ersten Beitrag vorgestellten Phasenmodells des Studienabbruchs erreichen. Weitere Implikationen der vorgestellten Beiträge werden für ein umfassendes Verständnis von Studienerfolg und -misserfolg aus einer prozeduralen Perspektive diskutiert sowie Empfehlungen für zukünftige Forschungsarbeiten abgeleitet.

*Beiträge des Symposiums*

### „Ich schmeiß´ alles hin und werde Prinzessin!“ – oder: Ist Studienabbruch ein mehrstufiger motivationaler Prozess?

**Lisa Bäumle, Markus Dresel**  
Universität Augsburg

In Deutschland brechen ungefähr ein Drittel aller Bachelorstudierenden ihr Studium ab (Heublein, 2014). Um aufzuklären, wie Studierende zu dem Entschluss kommen ihr Studium abzubrechen, oder ihren Studiengang zu wechseln, untersucht die vorliegende Arbeit den dafür verantwortlichen Prozess. Dieser lässt sich theoretisch als motivationaler Entscheidungsprozess konzipieren, der verschiedene Phasen hat (Mashburn, 2000). Diese reichen von der ersten Wahrnehmung, dass der Studiengang nicht zur eigenen Person passt, über die Informationssuche zu Alternativen und die Abbruch-/Wechselentscheidung bis zur Exmatrikulation. Basierend auf entscheidungs- und handlungspsychologischen Modellen (z. B. Achtziger & Gollwitzer, 2010; Betsch, 2005) kann angenommen werden, dass jede Phase durch spezifische Orientierungen und Bewusstseinslagen gekennzeichnet ist. Aufbauend auf diesen Modellen postulieren wir, dass vor der eigentlichen Exmatrikulation folgende Phasen durchlaufen werden: (I) Wahrnehmung einer Nicht-Passung zwischen Studienangebot und Person, (II) Gedanken an Studienabbruch/Studiengangwechsel, (III) Gezielte Information über außeruniversitäre Tätigkeiten oder andere Studiengänge, (IV) Abwägen (Wert, Erwartung) über Verbleib im Studiengang im Vergleich zu Handlungsalternativen, (V) Entscheidung, das Studium abzubrechen/den Studiengang zu wechseln, getroffen, aber noch nicht vollzogen. Der vorliegende Beitrag untersucht, inwiefern sich diese Phasen faktoriell trennen lassen und ob die darin thematisierten Prozesse in ihrer Abfolge dem postulierten Phasenmodell entsprechen. Weiterhin wird geprüft, inwiefern die einzelnen Phasen des Studienabbruchs bzw. Studiengangwechsels mit Persönlichkeitsaspekten, sowie motivationalen und affektiven Variablen zusammenhängen (in Übereinstimmung mit vorangehender Forschung, z. B. Respondek, Seufert, Stupnisky & Nett, 2017; Vallerand & Fortier, 1997). Dazu wurde eine Studie mit n = 1.005 Studierenden aller Studiengänge einer Universität im süddeutschen Raum durchgeführt (61% weiblich; Durchschnittsalter = 23.0 Jahre). Die Studierenden wurden per E-Mail rekrutiert und bearbeiteten einen Online-Fragebogen. Zu jeder postulierten Phase wurden fünf Items entwickelt (getrennt für Studienabbruch und Studiengangwechsel), die mit Likert-Antwortskalen präsentiert wurden. Darüber hinaus wurden das Fähigkeitsselbstkonzept (SESSKO, Schöne et al., 2002) das subjektive Wohlbefinden (Single-Item-Skala L1, Beierlein et al., 2014), leistungsbezogene Emotionen (AEQ, Pekrun et al., 2005) und Persönlichkeitsmerkmale (BFI-K, Rammstedt & John, 2005) mit Hilfe von etablierten Fragebögen erfasst. Die Ergebnisse der Datenanalyse bestätigen die Abgrenzbarkeit der postulierten Phasen: Konfirmatorische 5-Faktorenmodelle passten sehr gut zu den Daten, sowohl für

Studienabbruch (RMSEA = .07; CFI = .96; TLI = .95), als auch für Studiengangwechsel (RMSEA = .06; CFI = .97; TLI = .96). Sie waren allen möglichen 4-Faktorenmodellen, die aus der Zusammenfassung jeweils zweier Faktoren resultieren, überlegen (alle  $p < .001$ ). Weiterhin sprachen die Korrelationen zwischen Studienabbruch- und Studiengangwechselfaktoren für deren separate Betrachtung ( $.33 < r < .59$ ). Die Mittelwerte der daraufhin gebildeten Skalen folgten dem angenommenen Verlauf der Auftretenswahrscheinlichkeit. Eine Ausnahme stellte dabei Phase IV dar, die eine höhere Auftretenswahrscheinlichkeit aufwies, als die beiden vorangehenden Phasen. Die Ausprägung der verschiedenen Faktoren (Studienabbruch- und Studienwechselphasen) korrelierte substanzial negativ mit Fähigkeitsselbstkonzept, Wohlbefinden, Freude und Gewissenhaftigkeit ( $-.59 < r < -.10$ ), sowie positiv mit Angst, Scham und Neurotizismus ( $.10 < r < .41$ ). Dabei war eine Abnahme in der Korrelationsstärke mit steigender Phase zu erkennen. Insgesamt bestätigen die Ergebnisse die Annahmen zu den Phasen des motivationalen Prozesses des Studienabbruchs bzw. Studiengangwechsels. Eine Ausnahme stellt lediglich das Abwägen über den Verbleib im Studiengang dar (Phase IV). Diese Abwägung erfolgte möglicherweise zumindest bei einem Teil der Studierenden vor der (systematischen) Suche nach Alternativen. Weiterhin deutet das Zusammenhangsmuster daraufhin, dass der Einfluss motivationaler und dispositionaler Variablen auf Studienabbruch/Studiengangwechsel besonders zu Beginn des Prozesses bedeutsam ist und sich der Prozess zu späteren Zeitpunkten verselbstständigt. Die Befunde zeigen auf, dass die postulierten Phasenmodelle zu Studienabbruch bzw. Studiengangwechsel helfen, die beteiligten Regulationsprozesse zu verstehen und damit zur Früherkennung von potenziellen Studienabbrücker(inne)n beitragen.

### **Grit: Prädiktor von Erfolg im Studium? Der Zusammenhang zwischen Grit und Studiennoten vs. subjektiv wahrgenommenem Studienerfolg**

**Lisa Bardach<sup>1</sup>, Manuel Knoos<sup>2</sup>, Marko Lüftenegger<sup>1</sup>**

<sup>1</sup>Universität Wien, <sup>2</sup>Leibniz-Institut für Wissensmedien Tübingen

Im Studium sehen sich Studierende mit einer Vielzahl an Anforderungen konfrontiert. Manchen Studierenden gelingt es dabei besser als anderen, diese Anforderungen zu bewältigen und in ihrem Studium erfolgreich zu sein. Wodurch zeichnen sich diese erfolgreichen Studierenden aus? Ein relativ neues Konstrukt, das im Hinblick auf den Erfolg von Lernenden diskutiert wird, ist Grit. „Grit“ beschreibt die Fähigkeit von Lernenden, Langzeitziele mit Beharrlichkeit und beständigem Interesse zu verfolgen (Duckworth & Gross, 2014; Duckworth, Peterson, Matthews, & Kelly, 2007). Während Entscheidungsträger/innen im Bildungsbereich das Konstrukt Grit durchaus positiv aufnehmen (z.B. U.S. Department of Education, 2013), wird Grit und dessen Beitrag zur Erklärung von Lernerfolg auch kritisch gesehen. So bemängeln Forscher/innen eine unzureichende Abgrenzung zwischen Grit und der Persönlichkeitseigenschaft Gewissenhaftigkeit - bzw. insbesondere deren Facette Selbstkontrolle - und stellen die Frage, inwieweit es sich bei Grit statt um ein neues Konstrukt eher um „alten Wein in neuen Schläuchen“ handeln könnte (Credé, Tynan, & Harms, 2016). Zudem ist festzuhalten, dass die Forschung zum Zusammenhang zwischen Grit und dem Erfolg von Lernenden inkonsistente Befunde liefert (z.B. Duckworth et al., 2007; Ivcevic & Brackett, 2014). Eine mögliche Ursache hierfür könnte allerdings in den untersuchten Maßen für Studienerfolg liegen. Da Grit auf individuell bedeutsame Langzeitziele fokussiert, besitzt Grit für das vielmals hinzugezogene Erfolgskriterium Noten unter Umständen eine geringere Vorhersagekraft als für ein individuell definiertes Kriterium, das den subjektiv wahrgenommenen Erfolg in Lern- und Leistungskontexten widerspiegelt. In der vorliegenden Querschnittsstudie soll unter Kontrolle der Variable Selbstkontrolle der Zusammenhang zwischen Grit und zwei Maßen von Studienerfolg – Studienerfolg anhand von Noten sowie Studienerfolg anhand eines selbst spezifizierten Erfolgskriteriums - untersucht werden. An der im Rahmen dieser Studie durchgeführten Onlinebefragung nahmen 220 Studierende der Rechtswissenschaften einer österreichischen Universität teil (M = 22.30 Jahre, SD = 2.75 Jahre; 80% weiblich, Studienzeit: M = 6.41 Semester, SD = 3.92). Die zwei Grit-Komponenten Beharrlichkeit und beständiges Interesse wurden mit der deutschen Übersetzung der 12-Items-Gritskala von Duckworth und Kollge/innen (2007) erfasst (Fleckenstein, Schmidt, & Möller, 2014). Zur Messung von Selbstkontrolle wurde die adaptierte deutsche Version der Kurzform der Self-Control Scale (SCS-K-D) verwendet (Bertrams & Dickhäuser, 2009). Als Studienerfolgsmaß wurde neben der Studienleistung in Form von Noten abgefragt, was für Studierende persönlich Erfolg im Studium bedeutet. Studierende wurden im Anschluss daran gebeten, die Erreichung dieses individuell festgelegten Erfolgskriteriums zu bewerten. Die Analysen der Studie wurden unter Verwendung von Mplus und dem MLR-Schätzverfahren durchgeführt. Zum Umgang mit fehlenden Daten wurde das in Mplus implementierte Full-Information-Maximum-Likelihood-Verfahren (FIML) hinzugezogen, welches zur Schätzung von Modellen auf alle verfügbaren Informationen in den Daten zurückgreift. Die Effekte der zwei Grit-Komponenten Beharrlichkeit und beständiges Interesse auf Noten und das subjektive Studienerfolgskriterium wurden in einem Strukturgleichungsmodell unter Kontrolle des Effekts von Selbstkontrolle auf beide Studienerfolgsmaße modelliert. Es konnte kein signifikanter Zusammenhang zwischen der Grit-Komponente Beharrlichkeit und Studienerfolg anhand eines selbst festgelegten Erfolgskriteriums ( $\beta = .143, p = .471$ ) sowie Studienerfolg anhand von Noten ( $\beta = -.107, p = .421$ ) festgestellt werden. Auch für die Grit-Komponente beständiges Interesse zeigte sich weder ein signifikanter Zusammenhang zu Studienerfolg anhand eines selbst festgelegten Erfolgskriteriums ( $\beta = -.077, p = .539$ ), noch zu Studienerfolg anhand von Noten ( $\beta = -.017, p = .882$ ). Die Kontrollvariable Selbstkontrolle hing mit keinem der beiden Studienerfolgsmaße zusammen ( $\beta = .115, p = .585$  für Studienerfolg anhand eines selbst festgelegten Erfolgskriteriums, und  $\beta = .077, p = .539$  für Studienerfolg anhand von Noten). Die Ergebnisse dieser Studie stehen im Einklang mit Befunden, die Grit als Prädiktor von Noten ebenfalls nicht bestätigen konnten (z.B. Ivcevic & Brackett, 2014; Rimfeld, Kovas, Dale, & Plomin, 2016) und verweisen darauf, dass „grittiness“ auch für subjektiv wahrgenommenen Studienerfolg keine Rolle zu spielen scheint.

### **Gekommen um zu bleiben: Zur Bedeutsamkeit von Studienwahlmotivation für Studienzufriedenheit, Studienmotivation und Abbruchgedanken im Längsschnitt**

**Stefan Janke, Oliver Dickhäuser**

Universität Mannheim

Studienanfänger/-innen entscheiden sich aus verschiedenen Gründen heraus dazu, ein Studium aufzunehmen. Dabei lässt sich insbesondere zwischen intrinsischer (interessensgeleitete wachstumsorientierte Studienwahl) und extrinsischer Studienwahlmotivation (einkommensorientierte statusgeleitete Studienwahl) unterscheiden (siehe insbesondere Künsting & Lipowsky, 2011). In Einklang mit der Selbstbestimmungstheorie kann davon ausgegangen werden, eine intrinsisch motivierte Studienwahl eine Voraussetzung für die Entwicklung von intrinsischer Studienmotivation (Freude und Eigeninteresse am Studieren) und einer Lernzielorientierung (Streben nach Kompetenzerweiterung) darstellt (Vansteenkiste, et al., 2006). Dagegen sollte eine extrinsisch motivierte Studienaufnahme die Lernfreude und Studienzufriedenheit langfristig negativ beeinflussen (Künsting & Lipowsky, 2011, Schmuck, et al., 2000). Personen-Umwelt-Fit-Modelle legen nahe, dass intrinsische Studienwahl auch zu einer höheren Sicherheit führen, das richtige Studienfach gewählt zu haben, was sich in einem nächsten Schritt positiv auf die Studienzufriedenheit und negativ auf Abbruchgedanken auswirken (Le, et al., 2014). Empirische Befunde

zeigen allerdings bisher ein uneinheitliches Muster dahingehend ob Studienwahlmotivation überhaupt bedeutsame Prädiktoren für Studienzufriedenheit sind (Wach, et al., 2016). Darüber hinaus gibt es wenige Befunde zur Bedeutsamkeit von Studienwahlmotivation für die weitere Entwicklung der Studienmotivation. Bestehende Befunde basieren zusätzlich häufig auf retrospektiven Erfassungen (u.a. Blüthmann, et al., 2008) oder zeichnen sich durch eher kleine Stichprobengrößen aus (u.a. Schmidt-Atzert, 2005). Letztlich zeigt neuere Forschung, dass intrinsische Aspirationen in Lern- und Leistungskontexten auch positiv mit der Stärke der Lernzielorientierung assoziiert sind, während extrinsische Aspirationen wiederum positiv prädiktiv für Leistungszielorientierungen (Streben nach Kompetenzdemonstration) sind (Janke & Dickhäuser, 2016). Entsprechend ist nicht auszuschließen, dass bisherige Effekte der Studienwahlmotivation eigentlich nur Unterschiede in der individuellen Zielorientierung widerspiegeln. In der vorliegenden Studie sollen deshalb längsschnittliche Effekte der Studienwahlmotivation auf Studienzufriedenheit, Studienbelastung, Studienmotivation und Abbruchintention unter Kontrolle von Zielorientierungen untersucht werden. Im Rahmen der hier berichteten Längsschnittstudie wurden 536 Erstsemesterstudierende (M = 20.0 Jahre alt; SD = 2.6 Jahre; 65.3% weiblich) einer deutschen Universität zunächst wenige Wochen nach Beginn ihres Studiums und dann erneut am Ende ihres ersten akademischen Jahres befragt. Zum ersten Messzeitpunkt wurde die intrinsische und extrinsische Studienwahlmotivation mit etablierten Items aus groß angelegten Studierenden surveys erfasst (Ramm, et al., 2014); Lebenszufriedenheit wurde mit einer deutschen Fassung der Satisfaction with Life Scale erhoben (Schumacher, 2003). Ein Jahr später wurden etablierte Instrumente zur Messung der Studienzufriedenheit (Westermann, et al., 1996), der Studienbelastung (Gusy, et al., 2010) und intrinsische Studienmotivation (Deci & Ryan, 2016) eingesetzt. Abbruchgedanken und die subjektive Sicherheit das richtige Studienfach gewählt zu haben wurden mit Einzelitems gemessen. Zu beiden Messzeitpunkten wurden Zielorientierungen mit Hilfe der Skalen zur Messung von Lern- und Leistungsmotivation (Spinath, et al., 2002) erhoben. Die Hypothesen wurden pfadanalytisch mit Hilfe von Mplus und dem MLR-Schätzverfahren geprüft. Fehlende Daten wurden mit dem Full-Information-Maximum-Likelihood-Verfahren imputiert. Für den Einfluss von Abiturnote, Lebenszufriedenheit und Zielorientierungen zu Studienbeginn wurde kontrolliert. Intrinsische Studienwahlmotivation erwies sich als positiv prädiktiv für die Sicherheit das richtige Studienfach gewählt zu haben ( $\beta = .46, p < .001$ ) und die intrinsische Studienmotivation am Ende des ersten akademischen Jahres ( $\beta = .21, p = .004$ ). Es zeigte sich außerdem, dass die intrinsische Studienwahlmotivation nicht nur zu Studienbeginn positiv mit der Stärke der Lernzielorientierung zusammenhängt ( $r = .51, p < .001$ ), sondern sich auch über die Zeit positiv auf die Entwicklung der Lernzielorientierung auswirkt ( $\beta = .22, p < .001$ ). Extrinsische Studienwahlmotivation war negativ prädiktiv für das Ausmaß der intrinsischen Studienmotivation ( $\beta = -.17, p = .002$ ) und die Studienzufriedenheit ( $\beta = -.20, p = .001$ ). Sie erwies sich als positiv prädiktiv für die Stärke von Studienabbruchgedanken ( $\beta = .14, p = .020$ ) und die Studienbelastung ( $\beta = .13, p = .032$ ). Das Ergebnismuster zeigt, dass die Aufnahme eines Studiums aus intrinsischer Motivation langfristig positive Effekte auf die Entwicklung von Lernfreude hat, während extrinsische Studienwahlmotivation eher zu Belastungserleben und Abbruchgedanken führen.

### **Die Erfüllung der Bedürfnisse soziale Eingebundenheit, Kompetenzerleben und Autonomie und deren Zusammenhang mit Studienerfolgskriterien**

**Carla Bohndick<sup>1</sup>, Eva Lohmüller<sup>2</sup>**

<sup>1</sup>Universität Hamburg, <sup>2</sup>Universität Koblenz-Landau

#### Theoretischer Hintergrund

Laut Selbstbestimmungstheorie ist die Erfüllung der Bedürfnisse nach sozialer Eingebundenheit, Kompetenzerleben und Autonomie essentiell für die Entstehung von intrinsischer Motivation, Wohlbefinden und persönlichem Wachstum (Deci & Ryan, 2000). Dementsprechend wichtig ist die Bedürfniserfüllung auch für das Lehren und Lernen an der Hochschule (z.B. Faye & Sharpe, 2008; Reason et al., 2006; Tinto, 2010). Der Einfluss der Bedürfnisbefriedigung nicht nur auf die Motivation, sondern auch auf den Studienerfolg wurde in den letzten Jahren insbesondere im Zusammenhang mit der Person-Environment-Fit (P-E Fit) Theorie untersucht. Die P-E Fit Theorie besagt, dass eine Passung zwischen personalen (z. B. Bedürfnissen) und situationalen Bedingungen (z. B. Erfüllung der Bedürfnisse) im Arbeitskontext zu Outcomes wie Zufriedenheit, Leistung, Commitment oder Wohlbefinden führt (Edwards et al., 2006). Im Hochschulkontext wurden insbesondere Zusammenhänge zwischen Passung und den Studienerfolgsindikatoren Studienleistung (Etzel & Nagy, 2016; Li et al., 2012) und Studienabbruchintention (Li et al., 2015) gefunden.

#### Fragestellung

In der bisherigen Forschung zu P-E Fit an der Hochschule sind vor allem zwei Aspekte offen geblieben:

- 1) Zum einen wurde die Bedürfniserfüllung hier eher global erfasst und nicht in die drei Bedürfnisse nach sozialer Eingebundenheit, Kompetenzerleben und Autonomie unterteilt. Es blieb bisher also offen, ob alle drei Bedürfnisse gleichermaßen wichtig für den Studienerfolg sind oder welche differentiellen Effekte hier auftreten. Wenn bekannt ist, welche Bedürfnisse besonders stark mit Studienerfolg zusammenhängen, können Fördermaßnahmen gezielt eingesetzt werden.
- 2) Zum anderen muss – analog zu aktueller Forschung zum Verhältnis von Anforderung und Fähigkeiten (Bohndick et al., 2017) – auch der Frage nachgegangen werden, ob alle Studierenden gleichermaßen das Bedürfnis nach sozialer Eingebundenheit, Kompetenzerleben und Autonomie aufweisen und inwiefern diese Ausprägung der Bedürfnisse auch für den Zusammenhang zwischen Bedürfniserfüllung und Studienerfolgsindikatoren relevant ist. Sollte der Zusammenhang abhängig von der Bedürfnisausprägung sein, müsste beim Einsatz struktureller Fördermaßnahmen stärker auf die individuellen Bedürfnisausprägungen Rücksicht genommen werden.

#### Methode

Um die Fragestellungen zu beantworten, wurden in einer längsschnittlich angelegten Onlinebefragung 115 Studierende (73 % weiblich, M = 23.4 [SD = 2.76] Jahre alt, M = 7.28 [SD = 3.93] Hochschulsesemester) deutscher Hochschulen zu zwei Messzeitpunkten im Abstand von drei Monaten befragt. Die Onlinebefragung enthielt zu beiden Messzeitpunkten dieselben Instrumente. Zur Erfassung der Stärke der Bedürfnisse und der Bedürfniserfüllung wurde eine Adaption der Psychological Needs Scale (Neubauer & Voss, 2016) eingesetzt. Die interne Konsistenz war zufriedenstellend bis gut ( $.68 < \alpha < .90$ ). Zur Messung des Studienerfolgs wurden die Abbruchintention mit drei eigens entwickelte Items mit einer internen Konsistenz von  $\alpha = .74$  eingesetzt und für die Leistung wurde der Mittelwert der letzten vier Noten gebildet.

#### Ergebnisse

Zur Überprüfung der Fragestellungen wurden verschiedene Pfadmodelle spezifiziert. In allen Modellen wurde die Erfüllung der drei Bedürfnisse zum ersten Messzeitpunkt als exogene Variablen und beide Studienerfolgskriterien zum zweiten Messzeitpunkt als endogene Variablen eingesetzt. Zur Beantwortung der zweiten Fragestellung wurde darüber hinaus in der Bildung der Bedürfniserfüllung variiert: In einem Modell wurde die einfache Bedürfniserfüllung genutzt und in einem zweiten Modell wurde die Bedürfniserfüllung an die individuelle Bedürfnisausprägung angepasst. Konkret wurde dafür die Differenz

zwischen Bedürfnissen und Bedürfniserfüllung berechnet, sodass negative Werte auf unerfüllte Bedürfnisse hinweisen und die perfekte Passung zwischen Bedürfnissen und der Erfüllung der Bedürfnisse durch die Zahl Null repräsentiert ist.

Die Analysen zeigen, dass die Zusammenhänge zwischen Bedürfniserfüllung und den Studienerfolgsmerkmalen mit einer Varianzaufklärung von 5 % bei der Note und 12 % bei der Abbruchintention unerwartet gering sind. Dabei wird durch die einfache Bedürfniserfüllung deutlich mehr Varianz vorhergesagt als durch die an die individuellen Bedürfnisse angepasste Erfüllung. Das Kompetenzerleben sagt sowohl die Noten als auch die Abbruchintention vorher. Darüber hinaus wird die Abbruchintention auch von der erlebten Autonomie vorhergesagt. Gegen die Erwartung hat die soziale Eingebundenheit keinen Einfluss auf die Studienerfolgsindikatoren.

## B15

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H116

### Unterrichtsvideos – vom Instrument der Diagnose zum Mittel der Förderung unterrichtlicher Kompetenzen

Chair(s): **Matthias Baer** (Pädagogische Hochschule / Universität Zürich)

DiskutantIn(nen): **Staub Fritz** (Institut für Erziehungswissenschaft, Universität Zürich)

Das Interesse an Videos nicht nur zur Erfassung von Unterrichtskompetenzen von (angehenden) Lehrpersonen, sondern zunehmend auch um solche Kompetenzen zu fördern ist gross. Ausgehend von ihrer ersten Verwendung in TIMSS sind Unterrichtsvideos zu einem unentbehrlichen Forschungsinstrument geworden, das zur Fokussierung auf das unterrichtliche Handeln beigetragen hat (Pauli & Reusser, 2006). Mittlerweile werden in der empirischen Unterrichtsforschung multiperspektivische, multikriteriale, längsschnittliche Befragungs- und Teststudien mit Videoanalysen verknüpft, die eine Vielzahl instruktionspsychologischer und fachdidaktischer Prozessmerkmale zugänglich machen (Klieme, 2012). Neben der Erfassung eignen sich Unterrichtsvideos in hervorragender Weise für die Förderung von unterrichtlichen Kompetenzen. Sie kommen der Komplexität des Unterrichts nahe, ermöglichen, ihn zeit- und raumunabhängig, wiederholt und unter verschiedenen Gesichtspunkten zu analysieren sowie wissenschaftliches Wissen mit unterrichtlichem Handeln zu verbinden. Schlüsse auf vorhandene und zu fördernde Kompetenzen als Voraussetzung für adaptives Handeln im Unterricht werden möglich (Gaudin & Chaliès, 2015).

Systematisch verwendet werden Unterrichtsvideos in der Lehreraus-/weiterbildung allerdings immer noch weniger als ihr Potential es zulässt. Dabei ist über Unterrichtsvideos die Ausrichtung auf die individuellen Kompetenzen und ihre adaptive Weiterentwicklung möglich, was besonders auf ihre längsschnittliche Verwendung im Rahmen der Lehreraus- und -weiterbildung zutrifft.

Das Kodieren von videografierten Unterrichtsstunden bezieht sich auf verhaltensnahe, gut beobachtbare, abgrenzbare Merkmale und damit die niedrig-inferente Analyse von unterrichtlichen Sichtstrukturen. Dies allein ist jedoch eingeschränkt, weil nicht auf die Qualität der Lernprozesse der Schüler/innen geschlossen werden kann. Dafür sind tiefer liegende, sich auf die Qualität des Unterrichts beziehende Merkmale hoch-inferent zu beurteilen. Das Beurteilen der Unterrichtsqualität und das Kodieren der Art des Unterrichts haben je unterschiedliche Stärken und Schwächen, weshalb sie komplementär eingesetzt werden sollten.

Es bestehen unterschiedliche Möglichkeiten, wie Unterrichtsvideos in der Lehreraus- und -weiterbildung im Sinne des fallbasierten Lernens zwecks Förderung der Analysekompetenz von unterrichtlichen Qualitätsmerkmalen bei (angehenden) Lehrpersonen und auch zur Diagnose von durchgeführtem Unterricht eingesetzt werden können. Einerseits lassen sich anhand von Unterrichtsvideos fallbezogene Diskussionen über Unterricht sowie dessen Qualität anregen, andererseits können Unterrichtsvideos fokussiert mit niedrig-inferenten Analysen bzw. hoch-inferenten Ratings betrachtet werden. Die beiden Verfahren können parallel eingesetzt werden, um möglichst detailliert Auskunft über (a) Art und (b) Qualität des Unterrichts sowie (c) von eingesetzten Lernaufgaben zu erhalten. Im Sinne von problembasiertem Lernen zielt der Einsatz von Unterrichtsvideos in lehrkompetenzfördernder Absicht darauf ab, handlungsleitende Kognitionen zu vergegenwärtigen und ko-konstruktiv berufsrelevantes Wissen und Können aufzubauen bzw. weiterzuentwickeln.

Das Symposium thematisiert aktuelle Forschung, welche sich zum einen fallbasiert, reflektierend, zum anderen niedrig-inferent mit der Analyse der oberflächenstrukturellen Merkmale und parallel dazu hoch-inferent mit der tiefenstrukturellen Beurteilung der Qualität von Unterricht sowie mit der Beurteilung der Qualität von Lernaufgaben und der Reflexion von eigenem durchgeführtem Unterricht durch die betreffende Lehrperson befasst, sei dies in diagnostizierend-feststellender oder intervenierend-fördernder Weise und mit oder ohne längsschnittliche Perspektive.

*Beiträge des Symposiums*

### Lernen mit Unterrichtsvideos in der Lehrerbildung: Gelingensbedingungen aus der Sicht der Interventionsteilnehmenden

**Isabelle Hugener, Krammer Kathrin, Frommelt Manuela**  
Pädagogische Hochschule Luzern

Unterrichtsvideos sind eine geeignete Grundlage für das fallbasierte Nachdenken über Unterrichtsqualität und damit zur Verbindung von Theorie und Praxis der Unterrichtsgestaltung (Gaudin & Chaliès, 2015). Die Analyse von Unterrichtsvideos fördert das Anwenden von theoretischem Wissen auf Unterrichtssituationen und damit die Entwicklung der unterrichtsbezogenen Analysekompetenz. Die Fähigkeit zur differenzierten Analyse von Unterrichtssituationen gilt als bedeutsam für das erfolgreiche Unterrichten und umfasst die Fähigkeiten, lernwirksame Situationen zu erkennen und diese hinsichtlich ihrer Bedeutung wissensbasiert zu interpretieren (u.a. Seidel & Stürmer, 2014). Interventionsstudien zeigen einen positiven Effekt auf deren Entwicklung durch die Analyse von Videos (Hellermann et al., 2015; Krammer et al., 2016). Mit dem vermehrten Einsatz von Unterrichtsvideos in der Lehrerbildung gewinnt auch die Frage nach der Gestaltung von Settings zum Lernen mit Unterrichtsvideos an Bedeutung. Zentral für die Wirksamkeit des Lernens mit Unterrichtsvideos ist die Passung von Video und inhaltlichem Fokus sowie Begleitmaterialien (Blomberg, Renkl, Sherin, Borko & Seidel, 2013; Brophy, 2004; Santagata, 2014). Neben einem klaren Analysefokus kommt beim Lernen mit Videos der Strukturierung durch die Dozierenden eine wichtige Rolle zu, damit angehende Lehrpersonen ihre Aufmerksamkeit auf das Potenzial von Aufgabenstellungen und tiefenstrukturelle Aspekte lenken und eine vertiefte und konstruktive Auseinandersetzung mit den Videos stattfinden kann (Borko et al., 2011; Tripp & Rich, 2012).

Im Interventionsprojekt VideA wurden die Effekte der fallbasierten Arbeit mit eigenen bzw. fremden Videos zur Förderung der unterrichtsbezogenen Analysekompetenz untersucht. Die teilnehmenden Studierenden der Studiengänge Kindergarten-, Primarstufe oder Sekundarstufe I absolvierten das zweite Semester. Sie analysierten Unterrichtsvideos von fremden Lehrpersonen oder Ausschnitte aus dem eigenen Unterricht in Hinblick auf die Umsetzung pädagogisch-psychologischer Merkmale der Unterrichtsqualität (Zielklarheit, Lernbegleitung, Lernatmosphäre). In einer Kontrollgruppe analysierten Studierende schriftliche Lehr- und Lernmaterialien. Die ko-konstruktive Analyse wurde durch Dozierende entlang von Fragen moderiert, welche das Erkennen und die theoriebasierte Interpretation der Unterrichtssituationen unterstützten. Es konnte gezeigt werden, dass durch die Analyse von Unterrichtsvideos die unterrichtsbezogene Analysekompetenz der Studierenden bereits im ersten Ausbildungsjahr gefördert werden kann (Krammer et al., 2016). Um über die Gestaltung von Lernsettings zur produktiven Arbeit mit Videos mehr zu erfahren, wurde mit Hilfe von Interviews der Frage nachgegangen, welche Elemente der

Intervention das Lernen mit Unterrichtsvideos unterstützt haben. Hierfür wurden am Ende der Intervention mit 12 Studierenden und 8 Dozierenden aller drei Interventionsgruppen Leitfadeninterviews durchgeführt. Der Leitfaden enthielt Fragen zur Eignung der Unterrichtsvideos, dem vierschrittigen Vorgehen gemäss „lesson analysis framework“ von Santagata und Guarino (2011) sowie zur Rolle der Dozierenden bei der Moderation und Strukturierung der ko-konstruktiven Analyse. Die Interviews wurden transkribiert und mittels qualitativer Inhaltsanalyse (Kuckartz, 2014) unterstützt durch die Software und MAXQDA ausgewertet.

Die Ergebnisse der Interviews zeigen, dass die Studierenden und Dozierenden die Unterrichtsvideos als geeignet zur theoriebasierten und tiefenstrukturellen Analyse von Lehr-Lernprozessen einschätzen. Als besonders hilfreich nennen die Dozierende und die Studierende die Ergänzung der Unterrichtsvideos durch Begleitmaterialien wie beispielsweise Unterrichtsplanungen und Arbeitsblätter mit Lösungen der Schüler/innen, um den Blick auf tiefenstrukturelle Merkmale richten zu können. Das gewählte Analyseverfahren nach dem „lesson analysis framework“ wurde von den Dozierenden als sehr geeignet zur Erarbeitung von theoretischen Begründungen eingeschätzt und sie beschreiben, dass dadurch die Qualität der Diskussionen bezüglich Verarbeitungstiefe und Differenziertheit erhöht wurde, auch wenn es für die Studierenden herausfordernd war zuerst den Fokus auf die Aufgaben und das Lernen der Schüler/innen zu richten und erst in einem nächsten Schritt das Lehrerhandeln zu analysieren. Die Studierenden beurteilen das Vorgehen ebenfalls als geeignet, erlebten es teilweise aber auch als Einschränkung in der Diskussion. Der Begleitung durch die Dozierenden wird für das Gelingen des Lernens mit Videos von den Studierenden und Dozierenden eine grosse Bedeutung beigemessen. In der Gruppe der eigenen Videos wird zudem die Notwendigkeit zur Unterstützung bei der Auswahl von geeigneten Unterrichtsausschnitten deutlich.

### **Art und Qualität von Aufgabenstellungen im Deutschstunden: Ergebnisse einer videobasierten Untersuchung auf der Primarstufe**

**Moser Adrian<sup>1</sup>, Baer Matthias<sup>2</sup>**

<sup>1</sup>Universität / Pädagogische Hochschule Zürich, <sup>2</sup>Pädagogische Hochschule / Universität Zürich

Lernwirksamer Unterricht muss als Kerngeschäft der Schule hohe Qualitätsansprüche erfüllen. Im Rahmen der empirischen Unterrichtsforschung wurden in den letzten Jahren einzelne Unterrichtsqualitätsmerkmale erarbeitet und in wenigen Basisdimensionen zusammengefasst (Klieme, Schümer & Knoll, 2001). Die entsprechenden Studien sind allerdings vorwiegend in mathematisch-naturwissenschaftlichen Fächern verankert. Grössere Beiträge zur empirischen Unterrichtsqualitätsforschung für das Fach Deutsch existieren noch kaum (Stahns, 2013). Dies gilt auch für den Bereich der Aufgabenqualität, der einen wichtigen Bestandteil der Unterrichtsqualität darstellt (Reusser, 2014). Die ausserordentlich dünne Forschungslage ist mitverantwortlich, dass sich die Fachdidaktik Deutsch kaum an empirisch gesicherten Qualitätsmerkmalen orientiert, sondern stark praxeologisch ausgerichtet ist.

Im Beitrag wird auf eine explorative Pilotstudie eingegangen, welche die unzureichend geklärte Frage nach der Qualität von Aufgabenstellungen im Deutschunterricht aufnimmt. Ziel der Studie war es, den Ist-Zustand formaler Merkmale sowie der Qualität von Aufgaben im Deutschunterricht zu erfassen, zu beschreiben und Hinweise auf Entwicklungsfelder zu generieren. Untersucht wurden die Videoaufzeichnungen von 18 Deutschlektionen, die im Rahmen der berufspraktischen Ausbildung von Studierenden der Pädagogischen Hochschule Zürich entstanden sind. Zur Datenerhebung wurden ein Kategoriensystem und ein Rating-Inventar eingesetzt. Während das niedrig bis mittel inferente Kategoriensystem vorwiegend Aussagen über die formalen Merkmale der beobachteten Aufgabenstellungen zulies, diente das hoch inferente Rating-Inventar der Einschätzung der Qualität derselben.

Die Ergebnisse zeigten, dass aufgabenbezogene Aktivitäten mit einem mittleren zeitlichen Anteil von über 80% einen grossen Stellenwert in den untersuchten Lektionen einnahmen. Dennoch werden qualitätsrelevante Ansprüche an Aufgabenstellungen grösstenteils zu wenig berücksichtigt. So nutzten beispielsweise fast 90% der untersuchten Aufgabenstellungen kein reflexives Sprachhandeln, obwohl dies von Fachdidaktikern (Budde, Riegler & Wiprächtiger-Geppert, 2012) als besonders wichtig erachtet wird. Auch die übrigen untersuchten Aspekte von Aufgabenqualität wiesen einen markanten Verbesserungsbedarf auf. Insgesamt wird deutlich, dass bezüglich der Aufgabenqualität im Deutschunterricht viel Handlungsbedarf besteht. Entsprechend soll der Beitrag Anstoss für weitere Forschungsarbeiten geben, die zur Klärung der Qualitätsfrage für das Fach Deutsch beitragen.

### **Forschungsbasiertes Kodieren und Rating videographierter Unterrichtsstunden zur Förderung (ko-)konstruktivistischer und kompetenzorientierter Lehrkompetenz bei angehenden und erfahrenen Lehrpersonen – eine doppelte Intervention**

**Baer Matthias<sup>1</sup>, Kocher Mirjam<sup>2</sup>, Locher Anna Carina<sup>2</sup>, Villiger Christine<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule / Universität Zürich, <sup>2</sup>Pädagogische Hochschule Zürich

Unterrichtsvideos kommen nicht nur der Komplexität des Unterrichts nahe und werden für empirische Forschungsinteressen verwendet, sie eignen sich ebenso für die Förderung unterrichtlicher Kompetenzen. Dem Beitrag liegt die Feststellung zugrunde, dass ungeachtet des wissenschaftlichen Paradigmenwechsels nach wie vor hauptsächlich mithilfe fragend-entwickelnder Unterrichtsgespräche und der damit einhergehenden stark vorgeplanten, am Unterrichtsverlauf orientierten Inszenierung unterrichtet wird. Dieser unterliegt oft der Realisierung von problematischen, institutionell bedingten, (fach-)didaktischen Handlungsmustern. Für die Lehrerbildung erweist sich dies als ein zentrales Problem (Seidel, 2003, 2011; Baer et al. 2011, 2014, 2015; Mackowiak et al. 2013; Kocher et al., 2013, 2017; Wyss et al., 2013; Wuttke, 2009).

Durch eine forschungsbasierte Intervention mit Videos, in der die Analyse und Qualitätsbeurteilung eigener Unterrichtsvideos im Zentrum standen, wurde ein Beitrag zum notwendigen Paradigmenwechsel hin zu (ko-)konstruktivistischem kompetenzorientiertem Unterrichten unter Berücksichtigung der Heterogenität der Lernenden und ihrer kognitiven Aktivierung zu leisten versucht.

Proband/innen sind Studierende im sechsemestrigen Studiengang Primarstufe sowie erfahrene Lehrpersonen, die als Praxislehrpersonen mit der Betreuung von Praktika wesentlich zur unterrichtspraktischen Ausbildung der Studierenden beitragen. Der Fokus der Intervention liegt auf der adaptiven Förderung (1) angehender und parallel dazu (2) erfahrener Lehrpersonen, ausgehend von videographierten Unterrichtsstunden, die sie je im Fach Deutsch selber erteilt haben. Die Intervention mit den Studierenden besteht aus drei Teilen im dritten (innerhalb eines zweistündigen Semesterwochenmoduls), vierten (im Quartalspraktikum) und fünften (wiederum innerhalb eines zweistündigen Semesterwochenmoduls) Semester. Da sich die Intervention auf Studierende (Interventionsgruppen I) und erfahrene Lehrpersonen (Interventionsgruppe II) bezieht, die als Praxislehrpersonen tätig sind, und sie sowohl vom Forschungsteam als auch im zweiten Interventionsteil von den darauf vorbereiteten Praxislehrpersonen (Interventionsgruppe II) durchgeführt wurde, wird sie als doppelte Intervention bezeichnet. Ein zentrales Element der Intervention besteht darin, dass die Studierenden und die Praxislehrpersonen sowie die Mitglieder



des Forschungsteams im Sinne von «Third Space» (Gut et al. 2014) während des Quartalspraktikums zusammenwirkten. Beteiligt sind auch die Kontrollgruppen I (Studierende der bestehenden berufswissenschaftlichen Ausbildung) und II (Praxislehrpersonen ohne Interventionsteilnahme).

Die Intervention wurde multiperspektivisch mit acht Datenerhebungsinstrumenten überprüft, eingesetzt bei den angehenden Lehrpersonen zu vier Messzeitpunkten t1 bis t4 und zu zwei Messzeitpunkten t1 und t2 bei den Praxislehrpersonen. Es interessierten (a) die Ergebnisse zur längsschnittlichen Entwicklung über die vier bzw. zwei Messzeitpunkte der Kompetenz für (ko-)konstruktivistisches, kompetenzorientiertes Unterrichten im schulischen Alltag im Fach Deutsch bei den Interventionsgruppen I und II und (b) die Ergebnisse der querschnittlichen Vergleiche der drei Interventionsgruppen mit den Kontrollgruppen I und II zu jedem der vier bzw. zwei Messzeitpunkte.

Präsentiert werden Ergebnisse zu den längs- und den querschnittlichen Vergleichen der verschiedenen Probandengruppen. Sie weisen darauf hin, dass die doppelte Intervention, in welcher das längsschnittlich (t1 bis t4) wiederholt erfolgende Kodieren und die längsschnittlich ebenso wiederholte Qualitätsbeurteilung eigener auf Video aufgenommener Unterrichtsstunden durch die Probanden der Interventionsgruppen sowie die systematisch veranlasste Zusammenarbeit von Student/in, Praxislehrperson und Forschungsteam in der Funktion von Dozierenden der Lehrer/innenbildung, die auch auf spezifisch dafür entwickelten Tools beruhte, zentral waren und in welcher es darum ging, im Fach Deutsch kompetenzorientiertes, konstruktivistisches Unterrichten zu fördern, wirksam war. Ebenso zeigt sich, dass auch bei den erfahrenen Lehrpersonen in der Funktion von Praxislehrpersonen, die in die Intervention einbezogen waren (doppelte Intervention), von t1 zu t2 Zuwächse im Sinne der Ziele der Interventionsstudie zu verzeichnen sind.

## **Wie reflektieren Lehrpersonen für Wirtschaft und Gesellschaft eigene videografierte Unterrichtssequenzen?**

**Doreen Holtsch**

Institut für Erziehungswissenschaft, Universität Zürich

Das fachdidaktische Wissen der Lehrpersonen unterstützt die Gestaltung von kognitiv aktivierendem Unterricht und die Kompetenzentwicklung der Lernenden (Baumert & Kunter, 2011). Die kognitive Aktivierung von Lernenden kann beispielsweise mit der Aktivierung von Vorwissen oder der Initiierung von kognitiven Konflikten erreicht werden (Helmke, 2011, S. 639; Kunter & Trautwein, 2013, S. 68; Seidel, 2011, S. 616). Das Wissen ist jedoch nur ein Teil des fachdidaktischen Handlungsrepertoires zur kognitiven Aktivierung und erklärt nur ansatzweise, warum Lehrpersonen kognitiv aktivierenden Unterricht mehr oder weniger realisieren. Retrospektive Berichte von Lehrpersonen können über den Zusammenhang zwischen ihren kognitiven Prozessen und ihrem unterrichtlichen Können Aufschluss geben (Leuchter, 2009, S. 74; Peterson & Clark, 1978, p. 560).

Im Beitrag wird der Forschungsfrage nachgegangen, wie Lehrpersonen für «Wirtschaft und Gesellschaft» (W&G) eigene videografierte Unterrichtssequenzen reflektieren, in denen sie ihre Lernenden kognitiv aktivierten. Die von Lehrpersonen verbalisierten Begründungen und Alternativen für ihr Vorgehen können auf Ursachen für (mehr oder weniger) kognitiv aktivierende Unterrichtssituationen hinweisen.

Im Herbst 2014 wurden in der Deutschschweiz in Klassen des Ausbildungsberufs Kaufmann/Kauffrau jeweils zwei bis drei W&G-Lektionen auf Video aufgezeichnet. Anschließend wurde mit den sieben in die Untersuchung einbezogenen Lehrpersonen ein Stimulated Recall-Interview zu ihrem Unterricht durchgeführt, die zwischen 90 bis 120 Minuten dauerten. In den Stimulated Recall-Interviews wurden als Daten Verbalisierungen genutzt, die auf kognitiven Prozessen der Lehrperson beruhen (Shavelson & Stern, 1981, p. 458). Vor den Interviews wurden für jede Lektion jene Sequenzen (Teaching Events) identifiziert, in denen die Lehrpersonen Lernende kognitiv aktivierten. Entlang von Interviewleit-Fragen zur Begründung und Alternativen ihres Vorgehens reflektierten die Lehrpersonen diese Teaching Events.

Das Kodiermanual für die transkribierten Interviews wurde in Anlehnung an den von Borko, Roberts & Shavelson (2008) adaptierten „Teacher decision making framework“ (Shavelson & Stern, 1981) sowie die Überlegungen von Borko, Virmani, Khachatryan & Mangram (2014) sowie Es van (2011) als Matrix konzipiert. Das Kodiermanual berücksichtigte als Gegenstände der Reflexion vertikal Lernende, Lehrpersonen, fachdidaktische Aspekte und Bedingungen. Bezüglich der Levels wurden horizontal Deskription, Einschätzung, Begründung und Alternativen kodiert.

Die Analysen zeigten, dass Lehrpersonen die Teaching Events zur kognitiven Aktivierung vorrangig hinsichtlich fachdidaktischer Aspekte reflektierten. Obwohl in den Stimulated Recall-Interviews explizit nach Begründungen und Alternativen zu den beobachteten Teaching Events gefragt wurde, beschrieben die Lehrpersonen ihr Vorgehen in diesen Teaching Events fast genauso oft wie sie es begründeten. Die Lehrpersonen erklärten vergleichsweise selten fachdidaktische Alternativen zum videografierten Vorgehen in den Teaching Events.

Die in der Summe dargestellten Ergebnisse zeigten darüber hinaus sehr individuelle Reflexionsmuster. Beispielsweise reflektierte eine Lehrperson über alle Levels der fachdidaktischen Aspekte, während eine andere Lehrperson zwar verschiedene Aspekte thematisierte, allerdings nur auf einem beschreibenden Niveau. Abschliessend werden Schlussfolgerungen für die Aus- und Weiterbildung von Lehrpersonen für W&G gezogen.

## B16

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H115

### **Bedingungen und erste Ergebnissen der Bildungs- und Arbeitsmarktintegration von Geflüchteten**

Chair(s): **Aileen Edele** (Humboldt-Universität zu Berlin, Deutschland), **Petra Stanat** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Humboldt-Universität zu Berlin)

DiskutantIn(nen): **Cornelia Kristen** (Universität Bamberg)

In den Jahren 2013 bis 2016 haben etwa 1,5 Millionen Menschen einen Erstantrag auf Asyl in Deutschland gestellt (BAMF, 2017). Die Integration der neuzugewanderten Geflüchteten in das hiesige Bildungssystem und den Arbeitsmarkt stellt nicht nur eine enorme gesellschaftliche Herausforderung dar, sondern ist auch aus wissenschaftlicher Perspektive hoch interessant. So ist weitgehend ungeklärt, inwiefern sich theoretische Annahmen der Integrations- und Bildungsforschung auf schutzsuchende Neuzuwanderer übertragen lassen, deren Situation unter anderem im Hinblick auf Migrationsmotive, Aufnahmebedingungen und Bleibeperspektiven spezifisch ist. Zentral ist in diesem Zusammenhang die Frage, welche Bildungsvoraussetzungen die Geflüchteten mitbringen und durch welche Maßnahmen sich ihr Bildungspotenzial am besten unterstützen lässt. Die Integrationsforschung zeigt, dass Kenntnisse der Instruktionssprache bzw. Sprache des Aufnahmelandes von hoher Bedeutung sind für den Erfolg im Bildungssystem (z.B. Kempert et al., 2016) und Arbeitsmarkt (z.B. Chiswick & Miller, 2007). Als besondere Ressource von Zuwandern und ihren Nachkommen gelten ihre besonders hohen Bildungsaspirationen (Kao & Tienda, 1995; Gresch et al., 2012; Stanat et al., 2010). Zu den sprachlichen Voraussetzungen der Geflüchteten gibt es bislang jedoch kaum belastbare Befunde, insbesondere mangelt es an objektiven Kompetenzdaten. Und zur Schulbildung der Geflüchteten und ihren Bildungsaspirationen liegen nur erste Anhaltspunkte vor (z.B. Schupp et al., 2017). Gesicherte Erkenntnisse über den Erfolg von Integrationsmaßnahmen fehlen bislang ebenfalls.

Das Symposium beschäftigt sich mit der Bildungs- und Arbeitsmarktintegration von Schutzsuchenden, die in den letzten Jahren nach Deutschland gekommen sind, und analysiert, welche Faktoren diese begünstigen bzw. behindern. Ein besonderes Augenmerk liegt auf den Bildungsvoraussetzungen, die die Geflüchteten mitbringen und auf spezifischen Integrationsmaßnahmen (Vorbereitungsklassen und der frühe Besuch eines Integrationskurses). Um die Integration der Geflüchteten möglichst umfassend zu untersuchen, werden verschiedene Bildungskontexte in den Blick genommen, nämlich die allgemeinbildende Schule, die berufliche Bildung, die tertiäre Bildung, sowie der Arbeitsmarkt.

Der erste Beitrag analysiert auf Basis der IAB-BAMF-SOEP-Befragung von Geflüchteten die schulische Integration von schutzsuchenden Kindern und Jugendlichen. Er beschäftigt sich unter anderem mit Bedingungen, die den Besuch einer Vorbereitungsklasse begünstigen – einer umstrittenen Integrationsmaßnahme (z.B. Karakayali et al., 2017), die als Tradition der Ausländerpädagogik kritisiert wird (Diehm & Radtke, 1999).

Der zweite Beitrag untersucht Bildungsvoraussetzungen und bildungsbezogene Aspirationen von Geflüchteten, die eine berufsvorbereitende Klasse besuchen. Der beruflichen Bildung kommt für die Integration Geflüchteter eine Schlüsselrolle zu, da viele von ihnen als junge Erwachsene nach Deutschland gekommen sind und eine Berufsausbildung die Chance auf eine dauerhaft erfolgreiche Arbeitsmarktintegration eröffnet.

Im dritten Beitrag geht es um Geflüchtete, die sich für ein Hochschulstudium im Rahmen eines Online-Studienangebots interessieren. Es wird geprüft, welche studienrelevante Vorbildung und welche Kompetenzen in der Instruktionssprache Englisch die Studieninteressierten besitzen.

Basierend auf der IAB-BAMF-SOEP-Befragung analysiert der vierte Beitrag, inwiefern die frühe Teilnahme an einem Integrationskurs die (selbsteingeschätzte) Deutschkompetenz und Arbeitsmarktintegration begünstigt. Er liefert Hinweise auf Wirkungen der im Jahr 2015 eingeführten Integrationsmaßnahme (Asylbeschleunigungsgesetz).

Das Symposium bietet erste Einblicke in die Bildungsvoraussetzungen und Bildungsintegration von Geflüchteten. Zudem informiert es über zentrale Datensätze, die zur Bildungs- und Arbeitsmarktintegration von neuzugewanderten Schutzsuchenden in Deutschland zur Verfügung stehen, und illustriert deren Analysepotenzial. Das Thema wird aus psychologischer, soziologischer und wirtschaftspädagogischer Perspektive analysiert und aus einem soziologischen Blickwinkel diskutiert.

*Beiträge des Symposiums*

### **Bildungsbeteiligung von geflüchteten Kindern und Jugendlichen**

**Lisa Pagel**

Deutsches Institut für Wirtschaftsforschung (DIW Berlin)

Viele der Schutzsuchenden, die seit 2013 nach Deutschland gekommen sind, sind jünger als 18 Jahre. Eine erfolgreiche Eingliederung in das deutsche Schulsystem kann ihre Integration auf kultureller, struktureller, sozialer und emotionaler Ebene (Esser, 2006) ermöglichen. Obwohl unterschiedliche Regulationen in den einzelnen Bundesländern existieren, beginnt generell auch für geflüchtete Kinder und Jugendliche spätestens sechs Monaten nach ihrer Ankunft in Deutschland die Schulpflicht (Massumi et al., 2015). Erste Befunde aus Brandenburg deuten darauf hin, dass Jugendliche mit Asylbewerber- oder Flüchtlingsstatus häufiger niedrigere Schulformen und seltener ein Gymnasium besuchen als Schülerinnen und Schüler der Allgemeinbevölkerung mit und ohne Migrationshintergrund (Kemper, 2016). Je nach regionalen, strukturellen und schulspezifischen Vorgaben wurden vielerorts Vorbereitungsklassen für Kinder und Jugendliche mit geringen Deutschkenntnissen eingerichtet. Diese werden jedoch vielfach kritisch beurteilt (z.B. Karakayali et al., 2017). Internationale Studien deuten darauf hin, dass separierte Programme von Geflüchteten als Bestrafung empfunden werden (Miller, Mitchell, & Brown, 2005). Zudem scheint die separate Beschulung die Entwicklung eines Zugehörigkeitsgefühls zur Schule zu erschweren, welches positiv mit der schulischen, verhaltensbezogenen und psychologischen Anpassung assoziiert ist (z.B. Hastings, 2012; Kia-Keating & Ellis, 2007).

Bislang gibt es kaum gesicherte Erkenntnisse dazu, welche Schulformen geflüchtete Kinder und Jugendliche besuchen, die seit 2013 nach Deutschland gekommen sind, und von welchen Faktoren ihre Bildungsbeteiligung abhängt. Welchen Einfluss der Besuch von Vorbereitungsklassen auf ihre Eingliederung in das Schulsystem ausübt und welche Faktoren den Besuch separater Klassen begünstigen ist bisher ebenfalls unklar. Der vorliegende Beitrag betrachtet daher in einem ersten Schritt anhand der repräsentativen IAB-BAMF-SOEP-Befragung von Geflüchteten 2016 (Brücker, Roth, & Schupp, 2017) die Bildungsbeteiligung von geflüchteten Kindern und Jugendlichen und untersucht, welche regionalen und familiären Faktoren mit der Bildungsbeteiligung sowie dem Besuch einer Vorbereitungsklasse im Zusammenhang stehen.

Die Analysen beruhen auf Angaben der Erziehungsberechtigten zu ihren im Jahr 2010 oder früher geborenen minderjährigen Kindern (N = 3546). Es werden deskriptive Statistiken zum Schulbesuch und der Bildungsbeteiligung nach Schulform berichtet und mit der von 12-16-Jährigen ohne und mit Migrationshintergrund der Allgemeinbevölkerung (Sozio-Oekonomisches Panel, 2016) verglichen. Zudem werden zwei logistische Regressionen geschätzt, die die Fragen klären sollen, ob (1) regionale und familiäre Faktoren die Wahrscheinlichkeit eines Gymnasialbesuchs erhöhen und ob (2) regionale und familiäre Faktoren sowie die besuchte Schulform in einem Zusammenhang mit dem Besuch einer Vorbereitungsstufe stehen.

Die Ergebnisse zeigen, dass 89 Prozent der 6-17-jährigen Geflüchteten in Deutschland zur Schule gehen. 12-16-jährige Geflüchtete besuchen häufiger eine Grundschule oder eine Hauptschule und seltener ein Gymnasium als Jugendliche gleichen Alters mit und ohne Migrationshintergrund. Die logistischen Regressionsanalysen zeigen, dass die Unterbringung in einer Gemeinschaftsunterkunft sowie ein geringer oder unbekannter Bildungsstand signifikant die Wahrscheinlichkeit verringern, auf ein Gymnasium zu gehen. Leben in einer Kleinstadt hingegen erhöht die Wahrscheinlichkeit eines Gymnasialbesuchs. Zudem lernen Heranwachsende, die in einer Gemeinschaftsunterkunft oder in einer Großstadt leben, eher in einer Vorbereitungsstufe, wohingegen die Wahrscheinlichkeit dafür bei einer längeren Aufenthaltsdauer sinkt.

Ein Großteil der geflüchteten Kinder und Jugendlichen hat also bereits Zugang zu Schulbildung. Beim Besuch von weiterführenden Schulen lassen sich jedoch deutliche Unterschiede zur Allgemeinbevölkerung feststellen, welche letztendlich in niedrigeren Schulabschlüssen und einer erschwerten Integration in den Arbeitsmarkt münden könnten. Insbesondere deutet sich eine Benachteiligung von Kindern und Jugendlichen an, die in Gemeinschaftsunterkünften untergebracht sind. Sie besuchen besonders selten ein Gymnasium und besonders häufig eine Vorbereitungsstufe. Die erschwerten Lern- und Erholungsbedingungen in den Unterkünften (Behrens & Westphal, 2009), die Nachteile einer niedrigeren Schulform und einer Separierung im Schulkontext könnten im Zusammenspiel die Integration dieser Kinder und Jugendlichen erschweren. Empirisch bestätigt sich, dass Vorbereitungsstufen ein Übergangsphänomen darstellen und bei längerem Aufenthalt üblicherweise ein Wechsel in Regelklassen erfolgt. Die vorliegenden Ergebnisse bilden die Grundlage für zukünftige Analysen, welche die Rolle von Vorbereitungsstufen auf verschiedenen Aspekten der schulischen Integration untersuchen werden.

### **Junge Geflüchtete in der Berufsvorbereitung: Ergebnisse einer Primärerhebung**

**Claudia Diehl, Elisabeth Maué, Stephan Schumann**  
Universität Konstanz

Angesichts der Altersstruktur der jüngst zugewanderten Flüchtlinge (30% sind 16-25 Jahre; BAMF, 2017) hat das System der beruflichen Bildung eine entscheidende Bedeutung für ihre Integration in Deutschland. Ein hoher Anteil der Zugewanderten unterliegt der Berufsschulpflicht. Die Bundesländer reagieren auf diese – in ihrem Ausmaß neue – Herausforderung in unterschiedlicher Weise. Dies geschieht derzeit naturgemäß weitgehend ohne eine entsprechende Informationsgrundlage bezüglich Merkmalen und Ausgangsvoraussetzungen der jungen Flüchtlinge, ihren frühen Erfahrungen in Deutschland und ihren Zukunftsplänen.

Eine zentrale Herausforderung besteht darin, die geflüchteten Jugendlichen soweit in den Arbeitsmarkt zu integrieren, dass ihre gesellschaftliche Teilhabe in Deutschland möglich ist (Esser, 2001). Herausfordernd ist dies, weil ihre individuellen Voraussetzungen für die Arbeitsmarktintegration dadurch gekennzeichnet sind, dass sie in völlig anderen Gesellschafts- und Bildungssystemen sozialisiert wurden und gleichzeitig in einem Alter sind, in dem Integrationsprozesse langsamer verlaufen als bei jüngeren Kindern.

Das System der beruflichen Bildung ist für den Integrationsprozess Geflüchteter von enormer Bedeutung. Erstens steht es auch Personen mit geringer Vorbildung offen und ermöglicht ihnen den Erwerb eines Schulabschlusses, zweitens stellt es eine respektable Alternative zum tertiären Bildungssystem dar, das vielen Geflüchteten aufgrund ihrer bisherigen Bildungsbiographie (zunächst) verschlossen ist, und drittens ist es insofern aufnahmefähig, als derzeit viele Lehrstellen insbesondere im Handwerk unbesetzt bleiben.

Fraglich ist, inwieweit das System der beruflichen Bildung für die Betroffenen eine attraktive Option darstellt. Sein Besuch stellt eine Investition in Humankapital dar, die nicht nur mit der universitären Bildung konkurriert – die für viele Betroffene vermutlich an den fehlenden Bildungsvoraussetzungen scheitert. Auch der direkte Einstieg auf dem Arbeitsmarkt könnte sich angesichts der derzeit geringen Arbeitslosigkeit als Alternative zur Berufsausbildung erweisen, gerade vor dem Hintergrund, dass auf vielen jungen Migranten der Druck lastet, die Familie im Herkunftsland finanziell zu unterstützen. Eine Ausbildung – inklusive der entsprechenden Vorbereitung in berufsvorbereitenden Klassen, in denen zunächst vorrangig Deutsch gelernt wird – wird möglicherweise als langfristige und mühsame Alternative zum schnellen Arbeitsmarkteinstieg wahrgenommen, der sich langfristig in vielen Fällen mangels entsprechender Qualifikationen als Mobilitätsfalle (Wiley, 1970) entpuppen könnte.

Bezüglich der Merkmale und Ausgangsvoraussetzungen der jungen Flüchtlinge, ihrer frühen Erfahrungen in Deutschland und ihrer Zukunftspläne besteht ein Informationsdefizit. Diesem begegnet das interdisziplinäre Projekt RISE mit der Verbindung einer soziologischen und einer bildungswissenschaftlichen Perspektive. Es untersucht die Kompetenzen, Vorkenntnisse und sprachlichen Fähigkeiten der jungen Flüchtlinge in berufsvorbereitenden VABO-Klassen, ihre bisherigen Erfahrungen in Deutschland sowie ihre Zukunftspläne.

Vor dem theoretischen Hintergrund der Bedeutung von realistischen und idealistischen Bildungsaspirationen für Bildungsentscheidungen (Becker, 2017; Boudon, 1974; Stocké, 2014) und empirischen Befunden höherer Bildungsaspirationen von SchülerInnen mit Migrationshintergrund (Aspiration-Achievement-Paradox; Dollmann, 2017) geht der Beitrag der Frage nach, welche Merkmale die Befragten kennzeichnen (z.B. ihre Bildungsbiographie), und welche schulischen und beruflichen Aspirationen sie haben.

Die Datengrundlage bilden 675 SchülerInnen (79% männlich) in 60 berufsvorbereitenden VABO-Klassen (Vorqualifizierungsjahr Arbeit/Beruf mit Schwerpunkt Erwerb von Deutschkenntnissen) an 22 beruflichen Schulen im Regierungsbezirk Freiburg i.Br. Sie durchliefen einen kognitiven Fähigkeitstest (CFT20; Weiß, 1978) sowie einen Deutschttest (Eigenentwicklung) und füllten einen Online-Fragebogen aus (t1: Sommer 2017; t2: Frühjahr 2018).

Erste deskriptive Auswertungen von t1 zeigen, dass die geflüchteten Jugendlichen im Durchschnitt 7.6 Jahre im Herkunftsland zur Schule gegangen sind. 65% mussten den Schulbesuch vor dem Schuleintritt in Deutschland länger unterbrechen. Fast alle (93%) möchten nach dem Besuch der VABO-Klasse einen Schulabschluss in Deutschland erwerben, vorrangig einen Hauptschulabschluss (44%).

Weitere Auswertungen, die bei der GEBF präsentiert werden, beziehen sich auf den geplanten weiteren Bildungsweg bzw. direkten Einstieg in den Arbeitsmarkt, die Erklärung von Unterschieden in den Zukunftsplänen und die Frage, ob die Bildungsaspirationen mit zunehmender Aufenthaltsdauer – in Deutschland und innerhalb des Systems der beruflichen Bildung

– realistischer werden und von der sozialen Integration in Deutschland sowie Praktika und ersten Erfahrungen auf dem deutschen Arbeitsmarkt beeinflusst sind.

### **Vorbildung und sprachliche Eingangsvoraussetzungen Studieninteressierter mit Fluchthintergrund**

**Olga Zlatkin-Troitschanskaia<sup>1</sup>, Roland Happ<sup>2</sup>, Tobias Deribo<sup>2</sup>, Renata Suter<sup>3</sup>, Steffen Brandt<sup>3</sup>, Sarah Nell-Müller<sup>1</sup>, Franziska Reinhardt<sup>1</sup>**

<sup>1</sup>Johannes Gutenberg-Universität Mainz, <sup>2</sup>Johannes-Gutenberg Universität Mainz, <sup>3</sup>Kiron Open Higher Education

#### Theoretischer Hintergrund

Die Anzahl der Menschen, welche weltweit zum Verlassen ihrer Heimat gezwungen werden, ist aktuell so hoch wie nie zuvor (UNHCR, 2017). Innerhalb der Gruppe der Geflüchteten befinden sich auch eine Vielzahl von Personen, welche zu einem Hochschulstudium berechtigt sind oder bereits ein Studium aufgenommen hatten (UNESCO Institute for Statistics, 2017). Die aufnehmenden Gesellschaften (wie Deutschland) und die Geflüchteten selbst stehen jedoch vor großen Herausforderungen bei der Integration in den tertiären Bildungssektor. Zur erfolgreichen Teilhabe Geflüchteter an der höheren Bildung spielen dabei Vorbildung und Sprachkompetenzen eine Schlüsselrolle (z.B. Esser, 2006).

Das vom Bundesministerium für Bildung und Forschung geförderte Projekt Studierenerfolg und Studienchancen für Geflüchtete (SUCCESS) erforscht die Effektivität von Kiron Open Higher Education (Kiron), einer Online-Lernplattform, die Studieninteressierten Geflüchteten einen direkten und unbürokratischen Zugang zu höherer Bildung ermöglicht. SUCCESS fokussiert auf die Effektivität des Online-Studienangebotes von Kiron, das auf digitalen Bildungsangeboten und Massive Open Online Courses basiert. Dabei wird in SUCCESS eine differenzierte Eingangsdagnostik durchgeführt, welche zur Optimierung von Kiron's Lehr-Lern-Angeboten und damit zum Lernerfolg der Studieninteressierten beitragen soll. Es wurden u.a. Daten zur Studienmotivation und zur intellektuellen Leistungsfähigkeit (gemessen über eine Skala aus einem Intelligenztest) erfasst (Autoren, in press). Den theoretischen Rahmen für SUCCESS bilden Forschungsarbeiten zum Studierenerfolg (z.B. Richardson et al. 2012), welche basierend auf der aktuellen Forschung zu Studierenden mit Migrationshintergrund (z.B. Duong et al., 2016; Tjaden & Hunkler, 2017) sowie zur bildungsorientierten Integration von Geflüchteten (z.B. Schroeder & Seukwa, 2017) und deren akademischem Erfolg (z.B. van Herpen et al., 2017) ausdifferenziert und erweitert wurden (Autoren, in press).

#### Fragestellung

In diesem Beitrag werden erste Ergebnisse der Eingangsdagnostik beim Eintritt der Studieninteressierten in Kiron präsentiert und kritisch diskutiert. Der Fokus liegt auf der Vorbildung und der Englischkompetenz als notwendige Eingangsvoraussetzungen der Studierenden und umfasst folgende Forschungsfragen:

1. Über welche (studienrelevante) Vorbildung in formalisierten Bildungsinstitutionen verfügen die Studieninteressierten?
2. Wie sind die Sprachkompetenzen der Geflüchteten in Englisch?
3. Lässt sich ein Zusammenhang zwischen der selbstberichteten Vorbildung und den Sprachtestergebnissen sowie mit weiteren soziobiografischen Merkmalen (wie Alter, Geschlecht, Herkunftsland) der Studieninteressierten feststellen?

#### Methode

1268 Studieninteressierte mit Fluchthintergrund wurden im Zeitraum Juni-September 2017 bei der Registrierung auf der Kiron-Lernplattform befragt. Dabei wurden u.a. soziodemographische Merkmale, Vorbildung als Selbstangabe sowie Sprachkompetenzen in Englisch mittels eines C-Tests erfasst (Grotjahn, 2002; Norris, 2006). Dieser Test besteht aus vier Texten mit 77 bis 93 Wörtern und umfasst pro Text 23 bis 39 Items (Lückenwörter). Bei der Registrierung auf der Kiron-Plattform war der Sprachtest nicht verpflichtend, es handelt sich somit um ein low-stake-testing, bei dem die Testmotivation der Probanden erheblich variieren und z.T. die Validität der Testwertinterpretationen einschränken kann (z.B., Haertel, 1999). 1075 Personen bearbeiteten den Test. Die nachfolgenden Analysen zu den o.g. Fragestellungen basieren damit auf einem Teildatensatz mit 1075 Personen.

#### Ergebnisse

Die Ergebnisse zeigen extreme Unterschiede in den Eingangsvoraussetzungen der Kiron-Studieninteressierten: bei der Vorbildung reichen die Angaben von der nicht-abgeschlossenen Primarstufe bis hin zur abgeschlossenen Promotion. Der Einteilung des CEFR (North, 2014; Verhelst et al., 2009) folgend erreichen 18.9 % (N = 201) der Probanden ein Sprachniveau von B2 und höher. Von der überwiegenden Mehrheit der Studieninteressierten wird das für ein reguläres Hochschulstudium erforderliche B2-C1 Sprachniveau noch nicht erreicht, so dass spezielle Fördermaßnahmen erforderlich sind. Bemerkenswert ist die deutliche Diskrepanz zu der Selbsteinschätzung der Kiron-Studierenden, wobei sich 62.6 % der Probanden selbst auf einem B2-Level und höher einschätzen. Eine höhere (selbstberichtete) Vorbildung steht in keinem signifikanten Zusammenhang mit höheren Sprachtestergebnissen. Dieser Befund verdeutlicht, dass für die Entwicklung von effektiven Lehr-Lern-Maßnahmen zur erfolgreichen Integration von Studieninteressierten mit Fluchthintergrund ins reguläre Hochschulsystem objektive, valide und reliable Informationen über die individuellen Studieneingangsvoraussetzungen unverzichtbar sind.

### **Die Auswirkungen des frühzeitigen Zugangs zu Sprachkursen auf die Sprachkenntnisse und die Beschäftigungschancen von Geflüchteten: Erkenntnisse aus dem Asylverfahrensbeschleunigungsgesetz 2015**

**Tim Müller<sup>1</sup>, Zerrin Salikutluk<sup>1</sup>, Martin Kroh<sup>2</sup>**

<sup>1</sup>Berliner Institut für empirische Integrations- und Migrationsforschung (BIM), Humboldt-Universität zu Berlin (HU), <sup>2</sup>Deutsches Institut für Wirtschaftsforschung (DIW Berlin)

#### Theoretischer Rahmen

Die Beherrschung der Sprache des Aufnahmelandes (L2) ist positiv mit dem Arbeitsmarkterfolg und den Bildungschancen von Migranten korreliert (vgl. Chiswick and Miller 2007, Dustmann und Fabbri 2003, Esser 2006). Der Spracherwerb kann als Funktion aus dem Kontakt mit der L2-Sprache (Exposure), der Motivation und der Effizienz der Lernenden modelliert werden (vgl. van Tubergen 2010). Geflüchtete sind im Vergleich zu anderen Migrantengruppen speziellen Problemlagen ausgesetzt, da sie sich längere Zeit in Erstaufnahmeeinrichtungen aufhalten und geringeren Kontakt mit Einheimischen haben, Unsicherheiten über ihren Aufenthaltsstatus bestehen und ihre Sprachlernereffizienz möglicherweise durch traumatische Fluchterlebnisse und psychische Gesundheitsprobleme beeinträchtigt ist (vgl. Renaud und Godin 2002). Sprachkurse stellen eine Möglichkeit dar, den Spracherwerb der Geflüchteten zu unterstützen. Allerdings bestehen bisher nur wenige Erkenntnisse über die Effektivität solcher Kurse, wobei die berichteten Effektstärken häufig sehr klein oder statistisch nicht signifikant sind (vgl. Beenstock 1996;

van Tubergen 2010; Hayfron 2001). Die vorliegende Arbeit soll anhand der Auswirkungen des sog. Asylverfahrensbeschleunigungsgesetzes (AsylVfBG) von 2015 die Frage beantworten, wie stark sich der Effekt einer frühen bzw. späten Integrationskursteilnahme auf den L2-Erwerb und die Beschäftigungschancen von Geflüchteten kausal auswirkt. Aus der Reform ergeben sich systematisch unterschiedliche Zugangsvoraussetzungen für frühe Sprach- und Integrationskurse: Diejenigen Geflüchteten, die bereits früher in Deutschland ankamen, fanden schlechtere Zugangsbedingungen vor, während insbesondere die später angekommenen Geflüchteten gute Aussichten auf einen frühen Kursbeginn hatten.

#### Methode und Daten

Wir benutzen die exogene Variation in Bezug auf die gesetzliche Teilnahmeberechtigung, die sich für unterschiedliche Gruppen von Geflüchteten durch die Gesetzesreform ergeben hat, als Instrumente für eine frühe Kursteilnahme. Hieraus schätzen wir den kausalen Effekt der Kursteilnahme auf die Deutschkenntnisse und die Beschäftigungschancen mit Hilfe von instrumentellen Variablenschätzungen. Dabei benutzen wir die Daten von 2.300 Befragten der IAB-BAMF-SOEP-Geflüchteten-Stichprobe (Kroh et al. 2017), die sich hinsichtlich ihrer Ankunftsdaten in Deutschland systematisch unterscheiden und die somit von der Gesetzesreform unterschiedlich stark betroffen waren.

Unsere abhängigen Variablen sind zum einen die selbsteingeschätzten Deutschkenntnisse der Geflüchteten (Gesamtskala der Dimensionen Lesen, Schreiben und Sprechen). Zum anderen untersuchen wir die Chancen für den Arbeitsmarkteintritt in Form einer Erwerbstätigkeit, einer beruflichen Ausbildung oder eines Praktikums. Unsere zentrale unabhängige Variable ist der Abschluss eines Sprachkurses, wobei hier nach einer frühen (innerhalb der ersten sechs Monate nach Ankunft in Deutschland) und späten Kursaufnahme (später als sechs Monate) unterschieden wird. Die Vergleichsgruppe stellen Befragte dar, die noch keinen Kurs besucht hatten. Die unterschiedlichen Determinanten des Spracherwerbs werden in den Modellen berücksichtigt, indem wir für die Aufenthaltsdauer in Deutschland, Bleibeabsicht, die vorherigen Muttersprach-, Deutsch- sowie Englischkenntnisse, Qualifikationen, soziodemografische Charakteristika und den physische sowie psychische Gesundheitszustand kontrollieren.

#### Ergebnisse

Unsere Ergebnisse zeigen, dass die Teilnahme an Sprachkursen die Deutschkenntnisse fördert und sehr viel stärkere positive Effekte aufweist, als dies aufgrund bisheriger Ergebnisse zu erwarten war. Ein besonderer Vorteil früher Sprachkurse (innerhalb der ersten sechs Monate nach Ankunft) lässt sich jedoch nicht feststellen, da sowohl die frühe als auch späte Kursteilnahme in etwa den gleichen Effekt auf den Spracherwerb zeigen. Allerdings ergeben sich deutliche Unterschiede hinsichtlich der Auswirkungen auf die Erwerbschancen. Jene Befragte, die einen frühen Sprachkurs belegten, haben deutlich höhere Chancen, eine Erwerbstätigkeit aufzunehmen, als die Geflüchteten, die noch keinen Kurs oder einen späten Sprachkurs belegten. Insgesamt sprechen die Ergebnisse dafür, dass ein früher Zugang zu Sprachkursen die Integration von Geflüchteten in den Arbeitsmarkt tatsächlich zu beschleunigen scheint.

## B17

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H114

### Lehrkräfte und Unterricht als Wirkfaktoren auf die motivationale Entwicklung

Chair(s): **Rebecca Lazarides** (Universität Potsdam), **Hanna Gaspard** (Universität Tübingen)

DiskutantIn(nen): **Tina Hascher** (Universität Bern)

Motivationale Merkmale wie die fachbezogenen Wertüberzeugungen Lernender, ihre Interessen und ihr Fähigkeitsselbstkonzept sind von zentraler Bedeutung für erfolgreiche Lernergebnisse (Wentzel & Miele, 2016; Wigfield et al., 2015). Die gegenwärtige Forschung verdeutlicht dabei, dass professionelles Handeln und professionelle Kompetenzen die Entwicklung motivationaler Merkmale beeinflussen können (Mayer & Alexander, 2016; Woolfolk Hoy, Davis, & Pape, 2006). Das Symposium widmet sich daher einer differenzierten Analyse der Bedeutung von Merkmalen der Professionellen Kompetenz von Lehrpersonen (z.B. Motivation, Enthusiasmus, Leistungsbeurteilung) für die motivationale Entwicklung im Lehr-Lernkontext. Dabei werden verschiedene Altersgruppen untersucht (Primarstufe, Sekundarstufe & Hochschule). Das Symposium ist interdisziplinär angelegt (Psychologie, Erziehungswissenschaft & Schulpädagogik) und nutzt verschiedene methodische Zugänge (experimentelles Design, Interventionsstudie & Fragebogensurvey).

Der erste Beitrag befasst sich basierend auf Forschung zu Lehrerenthusiasmus sowie der Erwartungs-Formations-Hypothese im Rahmen einer experimentellen Studie mit der Untersuchung der Effekte des gezeigten Dozentenenthusiasmus und der Dozentenmotivation auf die Motivation und das emotionale Erleben von Lernenden. Im Rahmen von Vorlesungstutorien wurden Studierenden (N = 377) Videos mit manipulierten Ausprägungen des Enthusiasmus und der motivationalen Grundhaltung des Dozenten gezeigt. Die Ergebnisse zeigen Effekte der wahrgenommenen Dozentenmotivation auf das emotionale Erleben der Lernenden sowie Interaktionseffekte im Sinne eines Zusammenwirkens von Dozentenmotivation und Enthusiasmus.

Der zweite Beitrag stellt die Ergebnisse einer Interventionsstudie vor, die gezielt die wahrgenommene Relevanz des Mathematikunterrichts zu fördern versucht. Theoretische Grundlage der Intervention ist die Erwartungs-Wert-Theorie von Eccles und Kollegen (1983). Die Ergebnisse einer ersten Studie mit Lernenden der Sekundarstufe (N=1916) verdeutlichen langfristige positive Effekte der Intervention im Klassenkontext auf die Wertüberzeugungen und Leistungen in Mathematik. Der Enthusiasmus der Lehrkraft für das Fach Mathematik sowie normative Vorstellungen der Lehrkraft werden als Wirkfaktoren auf den Erfolg der Intervention untersucht.

Der dritte Beitrag befasst sich unter Bezugnahme auf die Erwartungs-Wert-Theorie von Eccles und Kollegen (1983) und auf Marsh's Forschung zum akademischen Selbstkonzept (Marsh, 1986) mit den Effekten der Noten Lernender und der Leistungsbeurteilungen von Lehrkräften auf das akademische Selbstkonzept sowie auf die subjektiven fachbezogenen Werte von Schülerinnen und Schülern in den Domänen Mathematik und Lesen. Besonders der big-fish-little-pond effect (BFLPE) wird untersucht. Datengrundlage bilden Schülerdaten (N=678 in Klassenstufen 3, 4, and 6) und Lehrerdaten (N=68) einer US-amerikanischen Längsschnittstudie. Für die Leistungsbeurteilungen der Lehrkräfte zeigten sich generell und spezifisch für den BFLP stärkere Effekte auf das akademische Selbstkonzept als für die Noten.

Der vierte Beitrag untersucht basierend auf dem Stage-Environment Fit Ansatz (Eccles et al., 1993) sowie auf Modellen der Professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2006) die Effekte des von Lehrkräften berichteten Enthusiasmus auf die Veränderung im Mathematikinteresse und in der wahrgenommenen Lehrerunterstützung nach dem Schulübergang in die Sekundarstufe. Ergebnisse der Auswertung von Daten Lernender der 5. und 6. Jahrgangsstufe (N = 1,000) und ihrer 42 Klassenlehrkräfte zeigen ein Absinken des Interesses sowie der wahrgenommenen Unterstützung auf Klassenebene. Der von der Klassenlehrkraft berichtete Enthusiasmus zu Unterrichten wirkte dem Absinken des Mathematikinteresses auf Klassenebene entgegen.

Die Ergebnisse aller Beiträge werden abschließend in Hinblick auf mögliche pädagogische Implikationen für Lern- und Instruktionsprozesse diskutiert.

#### Beiträge des Symposiums

### Macht Lernen mehr Spaß, wenn der Dozent enthusiastisch ist? Einblicke aus einem Feldexperiment

**Anne C. Frenzel<sup>1</sup>, Jamie L. Taxer<sup>2</sup>, Carolin Schwab<sup>1</sup>, Christof Kuhbandner<sup>3</sup>**

<sup>1</sup>Ludwig-Maximilians-Universität München, <sup>2</sup>Stanford University, <sup>3</sup>Universität Regensburg

Theoretischer Hintergrund und Fragestellung. In diesem Feldexperiment zielten wir darauf, Ideen aus zwei Forschungssträngen zu integrieren: Forschung zu Lehrerenthusiasmus sowie der Erwartungs-Formations-Hypothese (expectation formation hypothesis), die im Rahmen der Selbstbestimmungstheorie von Deci & Ryan (1987) formuliert wurde. Lehrerenthusiasmus ist unumstritten ein zentraler Bestandteil von hoher Lehrqualität (für einen Überblick siehe z.B. Keller, Woolfolk Hoy, Goetz, & Frenzel, 2016). Die Erwartungs-Formations-Hypothese besagt, dass Hinweise zur motivationalen Grundhaltung von Interaktionspartnern (intrinsisch versus extrinsisch) für die eigene Motivation und das eigene emotionale Erleben in der Interaktion bedeutsam sind (Deci & Ryan, 1987; Wild & Enzle, 2002). Diese Hypothese insbesondere für Lehr-Lernkontexte belegend, konnten Wild und Kollegen in einer Reihe von Studien zeigen, dass Lernende bei angeblich intrinsisch motivierten Lehrkräften selbst eine höhere intrinsische Motivation sowie größere Freude am Lernen erleben als beim Lernen mit angeblich extrinsisch motivierten Lehrkräften – trotz faktisch identischer Unterrichtssequenzen (Radel, Sarrazin, Legrain, & Wild, 2010; Wild, Enzle, & Hawkins, 1992; Wild, Enzle, Nix, & Deci, 1997). Zentrale Idee der vorliegenden Studie war es, gleichzeitig den tatsächlich gezeigten Dozentenenthusiasmus und die angebliche Dozentenmotivation zu manipulieren, um deren einzelnen und gemeinsamen Effekte auf die Motivation und das emotionale Erleben von Lernenden zu untersuchen.

Methode. N = 483 Studierenden (59% weiblich) wurde im Rahmen von Vorlesungstutorien eine von insgesamt vier Varianten eines kurzen instruktionalen Videos gezeigt. In diesen Videos präsentierte ein professioneller Schauspieler mit entweder hohem oder niedrigem Enthusiasmus aber ansonsten identischem Skript Informationen zu effektivem Lernen. Seine angebliche motivationale Grundhaltung wurde manipuliert, indem er entweder als bezahlt oder als ehrenamtlich vorgestellt wurde. Im Anschluss füllten die Teilnehmer einen kurzen Fragebogen aus, in dem die wahrgenommene Motivation und Unterrichtsfreude des Dozenten, sowie die eigene Lernmotivation, emotionales Erleben beim Lernen und der subjektive Lernerfolg im Selbstbericht erfasst wurde (alle Cronbach's Alpha  $\geq$ .71).

Ergebnisse. Wir verwendeten 2x2 ANOVAs zur Überprüfung der Haupt- und Interaktionseffekte von gezeigtem Lehrerenthusiasmus sowie angeblicher Dozentenmotivation. In der Hoch-Enthusiasmus-Bedingung sowie in der „Ehrenamts-Bedingung“ berichteten die Teilnehmer signifikant höhere Dozentenmotivation und –freude und auch positiveres eigenes

emotionales Erleben beim Lernen. Zudem ergab sich Interaktionseffekte für Dozentenmotivation und –freude, die dadurch charakterisiert waren, dass in der Hoch-Enthusiasmus Bedingung der Dozent gleichermaßen als motiviert und freudig erlebt wurde, aber in der Niedrig-Enthusiasmus-Bedingung der angeblich ehrenamtliche Dozent als motivierter und freudiger erlebt wurde. Für den subjektiven Lernerfolg ergab sich eine disordinale Interaktion: Dieser war in den jeweils „kongruenten“ Bedingungen (hoch enthusiastisch/ehrenamtlich sowie niedrig enthusiastisch/bezahlt) höher ausgeprägt als in den inkongruenten Bedingungen.

### **Auf die Motivation der Lehrkraft kommt es an?! Merkmale der Lehrkraft als Moderatoren der Effektivität von Nützlichkeitsinterventionen im Klassenkontext**

**Cora Parrisius, Hanna Gaspard, Heide Kneissler, Eike Wille, Benjamin Nagengast, Ulrich Trautwein**  
Universität Tübingen

Theoretischer Hintergrund und Fragestellung. Das Fach Mathematik wird von vielen Schülerinnen und Schülern als wenig relevant für ihr aktuelles und späteres Leben empfunden. Zusätzlich findet mit zunehmendem Alter ein Abfall der Motivation für Mathematik statt (Fredricks & Eccles, 2002; Frenzel, Goetz, Pekrun, & Watt, 2010; Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2017). Daran knüpfen erfolgreich Interventionen an, die gezielt verschiedene Aspekte der Motivation zu fördern versuchen (z.B. Hulleman & Harackiewicz, 2009; Morisano, Hirsh, Peterson, Pihl, & Shore, 2010). Die Nützlichkeit als eine der vier Komponenten der Wertüberzeugung im Sinne der Erwartungs-Wert-Theorie von Eccles und Kollegen (1983) ist eine von außen veränderbare Komponente der Motivation (Wigfield, Tonks, & Lutz Klauda, 2009). Hier setzte die Studie „Motivationsförderung im Mathematikunterricht“ (MoMa; Brisson et al., 2017; Gaspard et al., 2015) an zur Steigerung der wahrgenommenen Relevanz von Mathematik. Die Autoren konnten zeigen, dass innerhalb ihrer Stichprobe von 1916 Schülerinnen und Schülern in 82 Klassen aus 25 baden-württembergischen Gymnasien eine 90-minütige Intervention im Klassenkontext einen langfristigen positiven Effekt auf die Wertüberzeugungen und Leistungen in Mathematik hatte. Fraglich ist, ob der Effekt von weiteren Einflussfaktoren abhängt. Auf Ebene der Schülermerkmale finden sich viele Variablen wie das Geschlecht oder die anfängliche Motivation (Rosenzweig & Wigfield, 2016), die die Effekte solcher Motivationsinterventionen moderieren. Es ist jedoch weitestgehend unklar, ob auch Merkmale des Unterrichts oder der Lehrkraft einen Einfluss auf den Effekt der Intervention haben. Epistemologische Überzeugungen (z.B. Praxisrelevanz; Schreier et al., 2014) und motivationale Merkmale der Lehrkraft (z.B. Enthusiasmus; Frenzel, Goetz, Lütke, Pekrun, & Sutton, 2009; Patrick, Hissley, & Kempler, 2000) können ebenso subjektive Wertüberzeugungen und Emotionen der Schülerinnen und Schüler beeinflussen wie das Lehrer-Verhalten (Eccles & Wigfield, 2002). Möglicherweise bewirkt eine höhere Motivation auf Lehrerseite eine verstärkte Bereitschaft zur Auseinandersetzung mit den Interventionsinhalten (Assor, Kaplan, & Roth, 2002). Dies wirft die Frage auf, ob die Effekte einer Nützlichkeitsintervention mit der Motivation und dem Verhalten der Lehrkraft im Unterricht interagieren.

Methode. Diese Fragestellung soll mit Hilfe von Reanalysen der MoMa-Daten sowie durch eine MoMa-Folgestudie beantwortet werden, die von Oktober 2017 bis Februar 2018 durchgeführt wird. An dieser Interventionsstudie werden voraussichtlich ca. 2000 Neuntklässlerinnen und Neuntklässler aus 75 Klassen und 25 Gymnasien teilnehmen. Gegenüber der Vorläuferuntersuchung, in der die Intervention durch Projektmitarbeiterinnen durchgeführt wurde, wird in der MoMa-Folgestudie die Intervention durch die reguläre Lehrkraft (Interventionsbedingung 1) oder durch geschulte Masterstudierende (Interventionsbedingung 2) durchgeführt. Die Klassen werden randomisiert einer der drei Bedingungen (Wartekontrollgruppe als Bedingung 3) zugeteilt. Neben Wertüberzeugungen der Schülerinnen und Schüler, Leistungen und der Qualität des Unterrichts aus Schülersicht werden Wertüberzeugungen der Lehrkraft und die Unterrichtsqualität aus Lehrkraftperspektive zu drei Testzeitpunkten (Prätest-Posttest-Follow-Up-Design) erhoben. Zusätzlich wird die Intervention durch zwei Rater beobachtet, um die Qualität der Implementation durch die durchführende Person (Lehrkräfte versus Masterstudierende) zu beurteilen.

Ergebnisse. Die Reanalysen geben Hinweise darauf, dass lehrkraftspezifische Variablen den Effekt einer Nützlichkeitsintervention moderieren können. Motivationale Charakteristika wie der Enthusiasmus für das Fach Mathematik sowie normative Vorstellungen der Lehrkraft wie die berichtete Praxisrelevanz des Mathematikunterrichts scheinen Einflussvariablen für den Erfolg der Intervention zu sein. So verringert sich der von Gaspard und Kollegen (2015) berichtete Effekt der Intervention zum Posttest, wenn die Lehrkraft eine überdurchschnittliche Praxisrelevanz berichtet ( $\beta = .19, p < .001$ ), und fällt höher aus, wenn sie eine unterdurchschnittliche Praxisrelevanz berichtet ( $\beta = .44, p < .001$ ). Zum Follow-Up hingegen ist der Effekt der Intervention größer, wenn die Lehrkraft einen überdurchschnittlichen Fachenthusiasmus berichtet ( $\beta = .38, p < .001$ ), und umgekehrt geringer, wenn sie einen unterdurchschnittlichen Wert berichtet ( $\beta = .16, p < .05$ ). Auf diesen Ergebnissen aufbauend soll auch mit der MoMa-Folgestudie untersucht werden, ob die Effektivität der Intervention von Merkmalen der Lehrkraft abhängt. Implikationen der bisherigen Ergebnisse und der bis Februar verfügbaren Daten der MoMa-Folgestudie werden diskutiert.

### **How school grades and teachers' judgments of student aptitude shape students' motivations in math and reading**

**Fani Laueremann<sup>1</sup>, Jacquelynn S. Eccles<sup>2</sup>**  
<sup>1</sup>Universität Bonn, <sup>2</sup>UC-Irvine

Theory. School grades are an important lever of motivational change and perhaps one of the most powerful tools teachers can use to influence their students. Drawing on Eccles' expectancy-value theory (EVT, Eccles et al., 1983) and Marsh's internal/external frame of reference model (IE-FRM, Marsh, 1986), we examine the predictive effects of school grades and of teachers' judgments of their students' aptitudes on students' self-concepts of ability and subjective valuing of math and reading. We focus in particular on the so called big-fish-little-pond effect (BFLPE), according to which the predictive effect of students' achievement on their ability self-concept is positive, but the corresponding effect of the average classroom achievement is negative (Marsh et al., 2005). Being in a high-achieving classroom, for instance, sets a high standard against which individual achievement is compared; consequently, relatively high average classroom achievement can negatively impact students' own self-concept of ability and potentially their subjective valuing of math and reading. Thus, teacher-rated achievement indicators can impact students' motivations not only on the individual level, but also via aggregated classroom-level achievement.

Objectives. We test potential variance in the BFLPE as a function of whether students' achievement is assessed via school grades versus via teachers' (presumably privately-held) judgments of their students' aptitude in math and reading. Substantial evidence suggests that grades are more powerful predictors of students' ability-related beliefs than standardized test scores, likely because students tend to be more aware of their grades than of their standardized performance, and because grades might be comparatively more sensitive to motivational characteristics (e.g., rewarding effort in addition to achievement; Möller et al., 2009). However, by and large, the BFLPE-literature has not focused on teachers' judgments of student aptitude, even

though teachers' judgments of achievement can be stronger powerful predictors of students' motivations than teachers' judgments of student motivation (e.g., Rakoczy et al., 2008).

Methods: We reanalyze data from the Childhood and Beyond study (CAB; <http://www.rcgd.isr.umich.edu/cab>), focusing on wave 4, which included not only students (n=678 in grades 3, 4, and 6) but also their teachers (n=68). The analyses include classic expectancy-value assessments (Wigfield et al., 1997), teacher-rated student aptitude, school grades obtained from school records, parent-reported level of education, students' grade level, and gender. All motivational and achievement indicators are domain-specific for math and reading.

Results. We use a manifest-latent multi-level analysis (Marsh et al., 2009) to examine the BFLPE across math and reading, including parental education, students' grade level, and gender as covariates. First, school grades and teachers' judgments of student aptitude were only moderately correlated ( $r = .32$  to  $.44$ ). Second, separate analyses including either school grades or teacher judgments and focusing either on math or reading confirmed BFLPEs of both indicators on students' self-concept of ability. No significant BFLPEs were found on students' task values, which were comparatively less affected by ability and achievement ratings. Third, and somewhat surprisingly, teachers' judgments produced substantially higher BFLPEs than school grades across both math and reading. Fourth, when both teacher judgments and school grades were included as predictors, teacher judgments had stronger effects on students' self-concepts and were associated with stronger BFLPEs.

Our results underscore the importance of teachers' judgments of student aptitude in shaping students' motivations. Whereas much of the extant literature has focused on the effects of grades versus standardized performance (reviews in Marsh, 2007; Möller et al., 2009), teachers' judgments have received less attention. Yet, such judgments can be even more powerful predictors of motivations than grades. The antecedents of teachers' judgments and the mechanisms through which they affect students beyond grades warrant further investigation (see Kaiser et al., 2014).

### **Welche Rolle spielt der Enthusiasmus von Lehrpersonen für die Entwicklung von Interesse und wahrgenommener Unterstützung im Mathematikunterricht?**

**Rebecca Lazarides<sup>1</sup>, Hanna Gaspard<sup>2</sup>, Anna-Lena Dicke<sup>3</sup>**

<sup>1</sup>Universität Potsdam, <sup>2</sup>Universität Tübingen, <sup>3</sup>UC-Irvine

Theoretischer Hintergrund und Fragestellung. Der stage-environment fit Ansatz (Eccles et al., 1993) geht davon aus, dass eine mögliche Erklärung für das durch viele Längsschnittstudien aufgezeigte Absinken der Motivation Lernender nach dem Übergang in die Sekundarstufe (Frenzel, Goetz, Pekrun, & Watt, 2010; Watt, 2004) die mangelnde Passung des unterrichtlichen Lernkontextes zu den Bedürfnissen der Lernenden ist. Die Studie greift diesen Ansatz auf und untersucht inwiefern die Veränderung im Mathematikinteresse Lernender zwischen 5. und 6. Jahrgangsstufe mit der Veränderung in der wahrgenommenen Unterstützung durch die Lehrkraft in Zusammenhang steht. Ausgehend von Modellen der Professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2006) wird des Weiteren untersucht, welche Effekte die Motivation von Lehrpersonen, operationalisiert als Enthusiasmus, auf die Veränderungen in Interesse und wahrgenommener Unterstützung hat. Theoretische Ansätze zur emotionalen Transmission im Unterrichtskontext (Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009) beschreiben, dass der Enthusiasmus der Lehrperson sich auf das Interesse und die Freude Lernender im Unterricht überträgt. Die Studie fokussiert dabei insbesondere motivationale Veränderungen auf Klassenebene, da bislang nur wenige empirische Studien motivationale Veränderungen auf Klassenebene und ihre Zusammenhänge zu Veränderungen der auf Klassenebene wahrgenommenen Merkmale von Lehrpersonen und Unterricht untersuchen. Gleichzeitig ermöglichen solche Studien jedoch ein besseres Verständnis der Dynamiken im Klassenkontext und implizieren Handlungsmöglichkeiten für Lehrkräfte, die auf eine effiziente Motivationsförderung im Unterrichtskontext abzielen.

Methode. Die Studie basiert auf Daten einer Teilstichprobe der längsschnittlich angelegten TRAIN Studie (TRAIN = „Tradition und Innovation: Entwicklungsverläufe an Haupt- und Realschulen in Baden-Württemberg und Mittelschulen in Sachsen“). TRAIN ist eine Mehrkohorten-Längsschnittstudie, die die Entwicklung von Schulleistung, Motivation, Leistungsbereitschaft und Wohlbefinden während der Sekundarstufe I fokussiert und dabei insbesondere prüft, wie stark diese von unterschiedlichen Lernumgebungen beeinflusst wird. Die vorliegenden Analysen beziehen sich auf eine Teilstichprobe (n = 1,000, 53.6% Jungen; M Alter = 11.14 Jahre, SD = 0.57) der Originalstichprobe von Fünft- und Sechstklässlern (N = 3,160). Dabei wurden nur die Lernenden und ihre 42 Klassenlehrkräfte in die Analysen einbezogen, die keinen Wechsel der Klassenlehrkraft von der fünften zur sechsten Jahrgangsstufe erlebten und deren Klassenlehrkräfte angaben, das Fach Mathematik in der Klasse zu unterrichten. Die Lernenden wurden jeweils zu Beginn der fünften Jahrgangsstufe (T1) und zu Beginn der sechsten Jahrgangsstufe (T2) befragt. Die Stichprobe umfasste Lernende aus drei Schultypen in zwei Bundesländern (Hauptschule: 50.6%; Realschule: 10.4%; Mittelschule: 37.8%). Die mittlere Klassengröße war 21.28 Lernende/Klasse. Die Klassenlehrkräfte hatten durchschnittlich 19.97 Jahre Lehrererfahrung (SD = 12.55) und 70.7% der befragten Klassenlehrkräfte waren weiblich.

Ergebnisse. Die Ergebnisse zeigen wie erwartet ein Absinken des Interesses sowie der wahrgenommenen Unterstützung der Klassenlehrkraft auf Klassenebene. Die Veränderung im Interesse steht dabei jedoch in keinem signifikanten Zusammenhang zu der Veränderung in der wahrgenommenen Unterstützung der Lehrkraft. Der von der Klassenlehrkraft berichtete Enthusiasmus zu Unterrichten hat einen signifikant positiven Effekt auf die Veränderung des Mathematikinteresses und wirkt somit einem Absinken des Mathematikinteresses auf Klassenebene entgegen. Kein signifikanter Effekt zeigt sich vom durch die Klassenlehrkraft berichteten Enthusiasmus auf die Veränderung in der von Lernenden wahrgenommenen Unterstützung auf Klassenebene. Des Weiteren wirkte das Ausgangsniveau der wahrgenommenen Unterstützung der Lehrkraft dem Absinken des Interesses auf Klassenebene entgegen, während das Ausgangsniveau des Mathematikinteresses auf Klassenebene keinen Effekt auf die Veränderung der von den Lernenden wahrgenommene Unterstützung der Lehrkraft hatte. Im Hinblick auf pädagogische Implikationen verweisen die Ergebnisse auf die wichtige Rolle der professionellen Kompetenz von Lehrpersonen im Sinne positiver Effekte des Lehrerenthusiasmus, die einer Abnahme im Interesse Lernender effektiv entgegenwirken sowie auf die Bedeutung der Unterstützung Lernender im Klassenkontext.



## B18

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H102

### Professionelle Angebotsentwicklung in der Erwachsenen- und Weiterbildung zur Förderung gesellschaftlicher Partizipation

Chair(s): **Josef Schrader** (Deutsches Institut für Erwachsenenbildung - Leibniz-Zentrum für Lebenslanges lernen e.V., Deutschland & Universität Tübingen)

DiskutantIn(nen): **Harm Kuper** (Freie Universität Berlin)

Erwachsenen- und Weiterbildung ist in geringerem Maße institutionalisiert als Schule und Hochschule: Was in ihr gelehrt und gelernt wird, ist zumeist nicht durch administrative Vorgaben für Beteiligung, Ziele oder Inhalte bestimmt. Konstitutiv für das Gelingen der Erwachsenen- und Weiterbildung ist daher eine professionelle Angebots- und Programmplanung, die immer wieder neu bedürfnis- und bedarfsgerechte Gelegenheitsstrukturen für Bildungs- und Lernprozesse entwickelt, die die Wiederherstellung, die Aufrechterhaltung und die Weiterentwicklung der Handlungsfähigkeit im Umgang mit beruflichen, öffentlichen und individuellen Handlungsanforderungen zu unterstützen vermag (Gieseke, 2009). Während im politischen und akademischen Diskurs lange Zeit vor allem die Verbesserung der Erwerbsfähigkeit (von Teilgruppen) Erwachsener im Fokus stand, rückt derzeit vor dem Hintergrund zunehmender kultureller Heterogenität auch die Förderung der zivilgesellschaftlichen Partizipation ins Blickfeld (United Nations Assembly 2015; Schmidt-Hertha, 2017). Zivilgesellschaft bildet das Fundament einer demokratischen Gesellschaft, da sie der primäre Ort der Interessensvertretung und des Informationsaustausches ist (Klein et al., 2004). Die Förderung von Individuen, sich zivilgesellschaftlich zu engagieren, stellt somit eine entscheidende gesellschaftliche und politische Herausforderung dar. Oft wird argumentiert, dass Bildung, Wissen und Fähigkeiten fördert sowie zu einem ausreichenden Selbstbewusstsein beiträgt, welche die Teilhabe am gesellschaftlichen Leben ermöglichen und auslösen (Schuller et al., 2004). Während die Effekte von Schulbildung auf zivilgesellschaftliche Partizipation relativ gut identifiziert sind, gibt es bisher wenig Forschung zu den Effekten des Lernens im Erwachsenenalter. Darüber hinaus basieren die bisherigen, oft sehr positiven Befunde zum größten Teil auf nicht-repräsentativen Daten und deskriptiven Analysen von Querschnittsdaten. Um die Frage beantworten zu können, ob es der organisierten Erwachsenen- und Weiterbildung tatsächlich gelingt, Angebote und Programme zu entwickeln, die die zivilgesellschaftliche Partizipation der Teilnehmenden fördern, ist es unumgänglich, die Effekte organisierten Lernens zu schätzen sowie die dafür notwendigen Wirkmechanismen zu identifizieren. Hier setzen die Beiträge dieses Symposium an. Sie suchen Antworten auf die Frage, ob die Teilnahme an Erwachsenen- und Weiterbildung verschiedene Formen individueller zivilgesellschaftlicher Partizipation fördert. Vorgestellt werden Forschungsbefunde aus dem BMBF-geförderten interdisziplinären Projekt Nicht monetäre Erträge der Weiterbildung: zivilgesellschaftliche Partizipation (NEWZ). Ein weiteres verbindendes Element ist der Fokus auf die Re-Analyse von Daten aus der Bildungs- und Sozialberichterstattung, darunter das Nationale Bildungspanel (NEPS) des IIfBi, das Sozioökonomische Panel des DIWs und der Freiwilligensurvey des DZAs. Der erste Beitrag setzt den Fokus auf die kausale Modellierung der Effekte von berufsbezogener Erwachsenenbildungsteilnahme auf drei breit angelegte Formen der zivilgesellschaftlichen Partizipation, d.h. auf soziale, kulturelle und politische Teilhabe. Im zweiten Beitrag werden verschiedene Dimensionen der politischen Partizipation identifiziert und im Hinblick auf Bildungsbiografien im genaueren untersucht. Während die ersten beiden Vorträge die Effekte auf die Aufnahme zivilgesellschaftlicher Partizipation fokussieren, beleuchtet der vierte Beitrag Effekte der Erwachsenenbildung, die innerhalb einer ehrenamtlichen Tätigkeit stattfinden, auf die Wahrscheinlichkeit der Fortführung des Engagements.

*Beiträge des Symposiums*

### Soziale, kulturelle und politische Teilhabe als Erträge von Weiterbildung

**Jens Ruhose, Insa Weilage, Stephan Thomsen**  
Leibniz-Universität Hannover

Bildungsprozesse, deren Grundlagen früh gelegt werden, benötigen kontinuierliche Anpassungs- und Lernprozesse über die allgemeinbildenden und beruflichen Schulen hinaus, um den ständig verändernden Anforderungen im Lebensverlauf gerecht zu werden. Nach Daten des Adult Education Surveys (AES) aus dem Jahr 2011 haben in Europa rund 40,3% der Erwachsenen im Alter von 25 bis 64 Jahren in den letzten 12 Monaten vor der Befragung an einer Weiterbildungsmaßnahme teilgenommen (Europäische Kommission/EACEA/Eurydice, 2015). Die Europäische Kommission hat Weiterbildung im Erwachsenenalter zu einer der Prioritäten im Bildungsbereich gemacht. Demnach sollen Aktivitäten des Lebenslangen Lernens in der EU stetig ausgeweitet werden (Europäische Kommission, 2001). Neben Einkommen, Beschäftigung und ökonomischen Wachstum soll diese Strategie auch die zivilgesellschaftliche Partizipation und die soziale Kohäsion stärken (Europäischer Rat/Europäische Kommission, 2015).

Als zentrale Fragestellung in diesem Projekt untersuchen wir daher, ob die Teilnahme an Weiterbildung eine größere Bereitschaft für zivilgesellschaftliche Partizipation bewirkt. Zivilgesellschaftliche Partizipation umfasst in diesem Sinne die kulturelle, soziale und politische Teilhabe (Putnam, 1995, 2002). Weiterbildung definiert sich als die Teilnahme an Weiterbildungsmaßnahmen zur beruflichen Bildung. Die bisherige Literatur weist darauf hin, dass es einen positiven Zusammenhang zwischen Weiterbildung und zivilgesellschaftlicher Partizipation gibt (siehe Field (2011) für einen Überblick).

Inwiefern die Korrelationen aber auch eine kausale Interpretation zulassen, ist bisher offen. Die umfassende Literatur zu den monetären Effekten der Weiterbildung (siehe z.B. Ashenfelter (1978), Ashenfelter und Card (1985), LaLonde (1986) sowie für Deutschland z.B. Lechner (1999), Pischke (2001), Büchel und Pannenberg (2004)) zeigt zumindest, dass die größte methodische Herausforderung die Kontrolle der Selbstselektion in die Weiterbildungsmaßnahme darstellt (De Grip und Sauermann, 2013). Insbesondere gut gebildete und hoch motivierte Personen nehmen häufiger an Weiterbildungen teil. Diese positive Selbstselektion führt dann fast zwangsläufig zu einer Überschätzung der Weiterbildungserträge.

Unsere Analyse basiert auf Daten des deutschen Sozioökonomischen Panels (SOEP) für die Jahre 1994 bis 2014. Das SOEP bietet eine repräsentative jährliche Befragung von mehr als 20.000 Personen in mehr als 10.000 Haushalten. Neben detaillierten Informationen zu den Lebensumständen, Bildungs- und Arbeitsmarkthistorien werden u.a. Informationen zur Teilnahme an beruflicher Weiterbildung und Informationen zu kultureller (Besuch von kulturellen Veranstaltungen – Oper, Theater, Kino, Disco, usw. – sowie die aktive Beteiligung an kulturellen Aktivitäten), sozialer (ehrenamtliche Tätigkeit, Treffen mit und Hilfeleistungen für Freunde, Nachbarn und Verwandte) und politischer Teilhabe (Interesse an Politik sowie Partizipation in der Kommunalpolitik) erhoben.

Wir nutzen einen regressions-basierten Differenzen-in-Differenzen Ansatz mit einem vorgelagerten Propensity-Score Matching Verfahren (Heckman et al., 1997, 1998; Dehejia und Wahba, 2002; Todd, 2008), um die kontrafaktische Situation zu konstruieren, wie sich Weiterbildungsteilnehmer entwickelt hätten, wenn sie nicht an der Weiterbildung teilgenommen hätten. In einem ersten Schritt wird dabei ein Matching-Verfahren angewendet, das basierend auf einer Vielzahl von Variablen Personen identifiziert, die nicht an einer Weiterbildung teilgenommen haben, die aber ansonsten in den beobachtbaren Kriterien mit Weiterbildungsteilnehmern vergleichbar sind („statistische Zwillinge“). Diese Vergleichsgruppe nutzen wir dann in einem Differenzen-in-Differenzen Ansatz, der den Effekt der Weiterbildung zwischen Teilnehmern und Nichtteilnehmern vor und nach der Weiterbildungsteilnahme schätzt. Auf diese Weise eliminieren wir zusätzlich zeitpersistente (unbeobachtbare) Heterogenität zwischen den Personengruppen. Zuzüglich kontrollieren wir für einen Hauptselektionskanal, in dem wir Arbeitsmarktvariablen vor und nach der Teilnahme im Schätzansatz berücksichtigen.

Unsere Resultate zeigen zunächst, dass es eine starke Selbstselektion hinsichtlich monetärer und auch nicht-monetärer Kriterien in die Weiterbildung gibt. Typischerweise sind Teilnehmer wesentlich besser gebildet, haben höhere Löhne, sind jünger und haben eine höhere Wahrscheinlichkeit, in großen Firmen zu arbeiten. Außerdem weisen sie eine höhere Zufriedenheit mit ihrer Einkommenssituation und eine wesentlich höhere zivilgesellschaftliche Partizipation bereits vor der Programmteilnahme auf. Trotz dieser starken Selbstselektion finden wir, dass die Teilnahme an Weiterbildung zu einem Anstieg der kulturellen und sozialen Partizipation führt. Wir finden hingegen nur temporäre Effekte auf die politische Partizipation.

### **Dimensionen politischer Partizipation und ihr Zusammenhang mit unterschiedlichen Bildungsphasen**

**Robin Busse<sup>1</sup>, Julia Lischewski<sup>2</sup>, Susan Seeber<sup>1</sup>, Martin Baethge<sup>2</sup>**

<sup>1</sup>Georg-August-Universität Göttingen, <sup>2</sup>Soziologisches Forschungsinstitut Göttingen an der Georg-August-Universität

Der Zusammenhang zwischen Bildung und politischer Partizipation ist seit den 1960er Jahren viel diskutiertes Thema. Zahlreiche Untersuchungen beschäftigten sich über die Jahrzehnte mit der Identifikation möglicher Kausalbeziehungen. Für den US-amerikanischen Kontext konnte der theoretisch hergeleitete Zusammenhang von allgemeiner Schulbildung und Wahlbeteiligung (als eine Facette der politischen Partizipation) mehrfach bestätigt werden. Der Zusammenhang von Bildung und politischer Partizipation ist hierzulande erst in den letzten Jahren stärker in den Fokus der Forschung gerückt.

Allerdings wird eine kohärente Diskussion zum Themenbereich Bildung und politischer Partizipation durch einige Probleme erschwert: Die Operationalisierung der politischen Partizipation folgt oft keinem theoretischen Konstrukt, was dazu führt, dass in empirischen Untersuchungen einzelne Facetten betrachtet werden, bei denen nur unzureichend belegt ist, ob sie miteinander in einem Zusammenhang stehen und ob Ergebnisse zu einer bestimmten Partizipationsform auf andere transferierbar sind. Während bei Hadjar und Becker (2007) beispielsweise politisches Interesse als Bedingung für politische Partizipation im Vordergrund der Untersuchung steht, betrachtet Gille et al. (2011) die Mitgliedschaft in Vereinen. Recht häufig ist die Wahlbeteiligung als Form politischer Partizipation Gegenstand von Untersuchungen. Über die genannten Formen hinaus gibt es eine Vielzahl an Möglichkeiten, politisch zu partizipieren. Diese Komplexität politischer Partizipation ist bisher wenig systematisch untersucht, weder theoretisch-konzeptuell noch empirisch. Daher ist auch offen, wie die unterschiedlichen Formen politischer Partizipation zusammenhängen und welche Relevanz den verschiedenen Beteiligungsformen für aktive Mitgestaltung einer demokratischen Gesellschaft zuzuschreiben ist.

Ein weiteres Problem stellt die unterschiedlichen Operationalisierungen von Bildung dar. Diese reichen von hoch differenzierten Systematiken, wie sie beispielsweise der CASMIN-Klassifikation unterliegen, bis hin zu Ansätzen, die lediglich zwischen hoher und niedriger Bildung unterscheiden. In der Mehrzahl der Studien zu Zusammenhängen von Bildungsstand und Formen politischer Partizipation wird auf die Allgemeinbildung abgestellt und der Einfluss beruflicher Bildung oder gar Weiterbildung nicht beachtet. Die Folge der konzeptionellen und methodischen Schwächen ist, dass die Stabilität und Verallgemeinerbarkeit der gefundenen Effekte weiterhin schwer einzuschätzen bleibt.

Vor diesem Hintergrund untersucht der Beitrag welche Dimensionen von politischer Partizipation sich identifizieren lassen und durch den Verlauf unterschiedlicher Bildungsbiografien erklärt werden können. Im ersten Schritt werden ein- und mehrdimensionale Konstrukte der politischen Partizipation auf Basis des Jugendsurvey und des ALLBUS mittels konfirmatorischer Faktorenanalysen mit dem Ziel der Identifikation zentraler Dimensionen untersucht. Im Anschluss wird der Zusammenhang zwischen einzelnen Dimensionen und ausgewählten Merkmale unterschiedlicher Bildungsverläufe analysiert. Die Vertreter des „absolute education model“ gehen davon aus, dass Bildung einen direkten positiven Einfluss auf zivilgesellschaftliche Kompetenzen und politisches Wissen hat. Dies wird unter anderem mit der Zunahme verbaler und kognitiver Fähigkeiten mit den besuchten Schuljahren begründet (Nie, Junn, & Stehlik-Barry, 1996). Sprachliche Fähigkeiten, so die Annahme, bilden die Grundlage, um dem politischen Diskurs folgen zu können. Neben der allgemeinen Schulbildung sind in Deutschland die berufliche und/oder universitäre Bildung ein zentraler Bestandteil von Bildungsverläufen. In Anbetracht dieser Differenzierung des deutschen Bildungssystems ist es nicht unbedingt zielführend, wenn sich bisherige Analysen auf die allgemeine Schulbildung konzentrieren. Aber auch die Bedeutung von Weiterbildung als einer Form der Erwachsenenbildung sollte nicht unterschätzt werden. Das Verständnis von Weiterbildung als Bestandteil einer Bildungskette, die lebenslanges Lernen ermöglicht, wirft die Frage auf, ob sich mit dieser – neben beschäftigungsbezogenen Effekten - nicht auch Erträge für die gesellschaftliche Partizipation ergeben. Ziel des vorliegenden Beitrags soll es deshalb sein, mittels Regressionsanalysen den Einfluss der drei Bildungsphasen (Schulbildung, berufliche und/oder universitäre Bildung und Erwachsenenbildung) auf die politische Partizipation unter Berücksichtigung der unterschiedlichen Beteiligungsformen anhand empirischer Befunde aus verschiedenen Surveys zu diskutieren. Als Datengrundlage werden zusätzlich zu Querschnittserhebungen wie dem ALLBUS auch Längsschnittdaten des NEPS herangezogen. Im Anschluss an die Diskussion Ergebnisse zu den aufgeworfenen Fragen der Dimensionalität des Konstrukts und des Einflusses von Bildungsstand und verschiedener Bildungsmuster/-phasen wird zudem auf Limitationen aktueller Datensätze eingegangen.

### **Erwachsenenbildung und die Förderung des Freiwilligenengagements**

**Ina Rüber<sup>1</sup>, Josef Schrader<sup>2</sup>, Dogus Güler<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Erwachsenenbildung - Leibniz-Zentrum für Lebenslanges Lernen e.V., <sup>2</sup>Deutsches Institut für Erwachsenenbildung - Leibniz-Zentrum für Lebenslanges Lernen e.V., Deutschland & Universität Tübingen

Erwachsenenbildung wurde und wird im politischen und akademischen Diskurs nicht nur im Blick auf ihre Leistungen für das Wirtschaftssystem, sondern auch für eine gelingende Zivilgesellschaft erörtert. Die tragende Säule der Zivilgesellschaft ist das individuelle ehrenamtliche Engagement, welches sich nach dem gesellschaftlichem Wohl orientiert (Priemer et al., 2017). Dieser Beitrag geht der Frage nach, ob die Teilnahme an Erwachsenenbildung individuelles ehrenamtliches Engagement fördert. Der positive Zusammenhang von Erwachsenenbildung in Form von Teilnahme an Kursen oder Fortbildungen und

Ehrenamt in Form von aktiven Vereinsmitgliedschaften wurde bereits in einigen Studien belegt (e.g. Schuller et al. 2004, Vera-Toscano 2017). Die Studien geben allerdings bisher wenig differenzierten Aufschluss darüber, (1) ob ein Wirkungszusammenhang von Erwachsenenbildung auf Engagement besteht und (2) welche Kontexte und Themen innerhalb der Erwachsenenbildung entscheidend für einen potentiellen Effekt sind. Dabei bemerkt zum Beispiel Field (2005; 2011), dass Erwachsenenbildung oft innerhalb eines Engagements stattfindet und nicht Engagement fördert, sondern selbst durch Engagement begünstigt wird. Bisher besteht ebenfalls der Fokus darauf zu testen, ob die Teilnahme an Erwachsenenbildung die Aufnahme eines Engagements begünstigt. Nicht weniger interessant für die Zivilgesellschaft ist jedoch, ob Erwachsenenbildung auch einen Effekt auf die Beständigkeit eines Engagements hat.

Aus der Ehrenamtsforschung geht hervor, dass das freiwillige Engagement stark mit ökologischen Faktoren verknüpft ist (Bekkers, 2005; Einolf & Chambré, 2011; Hustinx et al., 2010; Wilson, 2012). Darunter fallen die Strukturen und Charakteristika sozialer Netzwerke und naher Kontexte. Sind alle Personen im persönlichen Umfeld freiwillig engagiert, entsteht ein starker sozialer Druck, sich selbst ebenfalls zu engagieren. Auch eine offen zutage tretende gesellschaftliche Notwendigkeit freiwilligen Engagements erhöht die Bereitschaft zum Ehrenamt. Sozial erwünscht und notwendig zugleich ist das freiwillige Engagement, wenn spezifische Ressourcen vorausgesetzt werden, die der (potentiell) Engagierte besitzt. Ein klassisches Beispiel ist das Engagement im Sportverein. Um eine Jugendgruppe zu leiten, werden nicht nur zeitliche Ressourcen notwendig, sondern auch pädagogische Kompetenz und Selbstdisziplin. Jemand, der diese Kompetenzen besitzt, steht unter besonderem sozialen und intrinsischem Druck, sich zu engagieren. Gleichzeitig können entsprechende Kompetenzen und Persönlichkeitsmerkmale durch Fortbildungen innerhalb des Engagements gefördert werden. Die Vermittlung und Aneignung der ehrenamtsspezifischen Kompetenzen stellt eine Investition sowohl der freiwillig Engagierten als auch der betreffenden Organisationen wie z.B. der Vereine dar. Klassischen ökonomischen Theorien zur Folge steigern diese derartige Investitionen den Wert des Engagements. Eine mögliche Folge ist, dass die Wahrscheinlichkeit, das Engagement zu beenden, sinkt. Wir nehmen entsprechend an, dass die Teilnahme an Erwachsenenbildung innerhalb eines Ehrenamts die Dauer der ehrenamtlichen Tätigkeit erhöht.

Zur Überprüfung der Annahme nutzen wir den scientific use file des Freiwilligensurvey (FWS) 2014 (Müller & Hameister, 2016), Version 1.2. Der Freiwilligensurvey wird von dem Deutschen Zentrum für Altersfragen (DZA) durchgeführt und ist eine repräsentative Befragung zum Freiwilligenengagement in Deutschland. Wir nutzen diese Daten, da diese sowohl ehrenamtliche Tätigkeiten in verschiedenen Organisationen erfassen als auch die Dauer des Engagements. Zudem wird erfasst, ob die Befragten innerhalb des Engagements an Erwachsenenbildung teilgenommen haben. Wir modellieren die Dauer des Engagements in Abhängigkeit von der Teilnahme an Erwachsenenbildung mit Hilfe von linearen Regressionen. Hier kontrollieren wir für Faktoren, für die bereits gezeigt wurde, dass sie sowohl Erwachsenenbildungsteilnahme als auch ehrenamtliches Engagement beeinflussen wie z.B. Bildungsbeteiligung, Einkommen und Erwerbsstatus. In einem weiteren Schritt untersuchen wir, ob der organisationale Rahmen, in dem das Engagement stattfindet, mit dem Effekt von Erwachsenenbildungsteilnahme auf die Dauer des Engagements interagiert.

Vorläufige Ergebnisse zeigen, dass es einen starken Zusammenhang zwischen der Teilnahme an Erwachsenenbildung innerhalb eines Engagements und der Dauer des Engagements gibt. Die Stärke des Zusammenhangs variiert, je nach subjektiv eingeschätzter Notwendigkeit der Teilnahme an Erwachsenenbildung und je nach Ehrenamtskontext. Geplante vertiefende Analyse versuchen die unterschiedlichen Zusammenhangsstärken zu erklären und die Richtung des Zusammenhangs theoretisch und soweit möglich statistisch zu erörtern.

## B19

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H001

### Nicht-monetäre Erträge von der Adoleszenz bis ins Erwachsenenalter

Chair(s): **Dave Möwisch** (DIPF Berlin, Deutschland), **Luise von Keyserlingk** (DIPF Berlin, Deutschland)

DiskutantIn(nen): **Beatrice Rammstedt** (GESIS, Mannheim)

Welche Einflüsse hat die erworbene Bildung auf das Individuum und die Gesellschaft? Diese Frage stellen Bildungsforscher aus verschiedenen Disziplinen, wie der Ökonomie, der Soziologie, der Psychologie und der Erziehungswissenschaft. Während monetäre Erträge von Bildung bereits in vielen Studien beforscht wurden, gibt es bisher weniger Arbeiten zu dem breiten Feld der nicht-monetären Erträge von Bildung (OECD, 2007). Nicht-monetäre Erträge lassen sich dabei aus unterschiedlicher Perspektive betrachten: Sie können u.a. auf individueller (z.B. Bildungsaspirationen, Gesundheitsverhalten) oder gesellschaftlicher Ebene (z.B. politische Partizipation) sowie auf berufsbezogener (z.B. Bildungsentscheidungen, größere berufliche Wahlmöglichkeiten) oder berufsferner Ebene (z.B. Wohlbefinden) untersucht werden (Vila, 2000). In diesem Symposium werden Arbeiten zu verschiedenen nicht-monetären Erträgen vorgestellt, die Einflüsse von Bildung und Bildungskontexten auf soziale, motivationale und gesundheitliche Outcomes untersuchen. Dabei werden verschiedene Zeitpunkte in der Bildungsbiografie im Verlauf von der Adoleszenz bis ins Erwachsenenalter betrachtet. Alle vier Beiträge betrachten nicht-monetäre Erträge auf individueller Ebene und fokussieren dabei sowohl unmittelbare (Aspirationen am Ende der Schulzeit), als auch verzögerte Outcomes (Wohlbefinden im Erwachsenenalter). Zudem werden in den Beiträgen neben direkten Zusammenhängen von Bildung und Bildungsergebnissen auch vermittelnde Prozesse (z.B. über entstehende soziale Kosten oder Selbsteinschätzungen) untersucht und diskutiert.

Der erste Beitrag beschäftigt sich mit der Frage, inwieweit aufwärts gerichtete Bildungsmobilität in bestimmten Peergruppen auch mit Kosten assoziiert sein kann. Anhand der Daten der Startkohorte Klasse 9 des Nationalen Bildungspanels wird gezeigt, dass Bildungsmobilität, die durch vom Netzwerk abweichende Bildungsaspirationen operationalisiert wurde, auch mit sozialen Kosten zusammenhängt.

Der zweite Beitrag widmet sich den Effekten des akademischen Selbstkonzepts auf Bildungsentscheidungen beim Übergang von der Schule in die tertiäre Bildung. Zudem wird anhand längsschnittlicher Daten untersucht, wie der Schulkontext (Big-Fish-Little-Pond Effekt) auf Bildungsentscheidungen junger Erwachsener wirken kann.

Im dritten Beitrag werden typische und untypische Übergangsverläufe junger Erwachsener nach dem Schulaustritt untersucht. Der Fokus liegt dabei ebenso auf emotionalem Wohlbefinden als Outcome Variable von Bildungsverläufen, wie auch auf der Frage, wie emotionales Wohlbefinden durch die Herausforderungen des Überganges nach dem Schulaustritt beeinflusst werden kann.

Der vierte Beitrag geht der Frage nach, wie Bildung als Ressource für affektives Wohlbefinden und Lebenszufriedenheit wirken kann. Dazu wurden Daten aus dem Sozio-oekonomischen Panel genutzt, das eine repräsentative Stichprobe beinhaltet und Analysen über die gesamte Spanne des Erwachsenenalters ermöglicht.

Alle vier Beiträge werden abschließend hinsichtlich praktischer und bildungspolitischer Implikationen diskutiert.

#### Beiträge des Symposiums

### Negative soziale Effekte von Bildungsmobilität? Eine empirische Überprüfung der Aufwärtsmobilitätshypothese

**Marina Lagemann<sup>1</sup>, Peter Winker<sup>2</sup>**

<sup>1</sup>Universität Hamburg, <sup>2</sup>Justus-Liebig-Universität, Gießen

Zur Erklärung ethnischer Bildungsdisparitäten wird häufig auf Boudons (1974) Ansatz zur Erklärung sozialer Bildungsungleichheiten zurückgegriffen. Dieser modelliert soziale Bildungsungleichheiten als Resultat ungleich verteilter Erfolgsvoraussetzungen (primäre Stratifikationseffekte) sowie systematisch unterschiedlicher Bildungsentscheidungen von Familien in unterschiedlichen sozialen Positionen (sekundäre Stratifikationseffekte).

Nach Boudon lassen sich sekundäre Effekte vor allem auf eine sozial stratifizierte Struktur der mit zusätzlicher Bildung verbundenen sozialen Kosten und Erträge zurückführen: Zum einen sind ambitionierte Bildungsentscheidungen für sozial besser gestellte Familien aufgrund des sonst drohenden Risikos eines Statusverlustes mit vergleichsweise hohen erwarteten Erträgen verbunden. Zum anderen sind Bildungsentscheidungen, die weniger ambitioniert sind als die des Freundeskreises, für Schüler(innen) in höheren sozialen Positionen mit sozialen Kosten verbunden. Für Schüler(innen) mit niedrigem sozialem Status ist hingegen die Wahl von Bildungsalternativen, die die Bildungsentscheidungen ihres sozialen Netzwerks übersteigen, mit sozialen Kosten verbunden. Ähnlich stärken ambitionierte Bildungsentscheidungen die Familiensolidarität in höheren sozialen Positionen, während sie sie in sozial niedrigeren Positionen schwächen.

Zwar lässt sich ein Großteil der ethnischen Bildungsdisparitäten durch Unterschiede im sozialen Status erklären. Die Beobachtung signifikant höherer Bildungsaspirationen in Migrantenfamilien bereits vor Kontrolle der sozialen Position deutet jedoch darauf hin, dass die Bildungsorientierungen von Migranten in niedrigen sozialen Positionen nicht denen von Familien ohne Migrationshintergrund entsprechen. Erklärt wird dieses Phänomen durch die Existenz sekundärer ethnischer Effekte, welche die Bildungsentscheidungen von Migranten über soziale Effekte hinaus beeinflussen (Heath & Brinbaum 2007).

Die derzeit populärste Erklärung für die Existenz positiver sekundärer ethnischer Effekte bezieht sich auf das Streben nach Aufwärtsmobilität als zentrales Migrationsmotiv: Da Aufwärtsmobilität von der ersten Generation meist nicht erreicht wird, wird eine hohe Bildung der Kinder als Weg zur Erreichung dieses Ziels angesehen und führt daher auch in Familien in niedrigen sozialen Positionen zu vergleichsweise hohen Bildungsaspirationen (Heath & Brinbaum 2007).

Vorliegende empirische Befunde sind zwar mit der Aufwärtsmobilitätshypothese, jedoch auch mit alternativen Erklärungsansätzen für die vergleichsweise hohen Aspirationen in Migrantenfamilien kompatibel (Trebbels 2015). Ziel des Beitrags ist es, die Aufwärtsmobilitätshypothese möglichst direkt zu überprüfen und damit zum Verständnis der Formationsmechanismen von Bildungsentscheidungen beizutragen.

Grundlage der Analysen sind die Daten der Startkohorte Klasse 9 des Nationalen Bildungspanels. In einem ersten Schritt wird die Verteilung der Bildungsaspirationen und -einstellungen von Schüler(inne)n und Eltern in Familien mit und ohne Migrationshintergrund nach elterlichem Bildungsstand betrachtet. In einem zweiten Schritt werden soziale Erträge u.a. als

Funktion von (angestrebter) Aufwärtsmobilität modelliert, um die Hypothese zu testen, dass Aufwärtsmobilität bei Nichtmigranten, nicht aber bei Migranten in niedrigen sozialen Positionen mit sozialen Kosten verbunden ist.

Bildungsmobilität wird modelliert als vom sozialen Netzwerk (nach oben) abweichende Bildungsaspirationen/-einstellungen bzw. -übergänge. Dazu werden neben den Informationen über die Schüler(innen) selbst Informationen über Bildungsaspirationen und -verhalten von Eltern sowie im Freundeskreis einbezogen. Zur Abbildung sozialer Kosten bzw. Erträge werden unterschiedliche zur Verfügung stehende Konstrukte verwendet – wie die Zufriedenheit mit der Familie, dem Freundeskreis und Problemen im Umgang mit Gleichaltrigen.

Zur Schätzung sekundärer ethnischer Effekte wird u.a. für ungleiche Erfolgswahrscheinlichkeiten von Migranten und Nichtmigranten (durch Maße wie Testleistungen und Noten sowie durch wahrgenommene Erfolgswahrscheinlichkeiten) und für den sozialen Status kontrolliert. Systematische Unterschiede der Effekte von (angestrebter) Bildungsmobilität bei Migranten und Nichtmigranten in niedrigen sozialen Positionen modellieren wir durch die Interaktion von sozialem Status bzw. elterlichem Bildungsstand und Bildungsmobilität sowie durch die Schätzung der Effekte von Bildungsmobilität an verschiedenen Stellen der Verteilung des sozialen Status (Quantilsregressionen, marginale Effekte).

In Einklang mit der Aufwärtsmobilitätshypothese zeigen sich systematische Unterschiede in den Bildungsaspirationen und -einstellungen von Migranten und Nichtmigranten in Familien mit niedrigerem, aber nicht mit hohem Bildungsniveau. Die multivariaten Analysen zeigen, dass angestrebte Bildungsmobilität (nach oben) bei alternativen Operationalisierungen von Bildungsmobilität mit sozialen Kosten verbunden ist. Diese fallen entgegen der in der Literatur diskutierten Hypothese jedoch nicht nur bei Nichtmigranten, sondern auch bei Migranten an.

### **Effekte des Schulkontextes auf Studienaspirationen und Studienfachwahlen**

**Luise von Keyserlingk<sup>1</sup>, Michael Becker<sup>2</sup>, Malte Jansen<sup>3</sup>, Kai Maaz<sup>1</sup>**

<sup>1</sup>DIPF, Berlin, <sup>2</sup>DIPF, Berlin, IPN; Kiel, <sup>3</sup>IQB, Berlin

Das akademische Selbstkonzept (ASK) als nicht-monetärer Ertrag von Bildung ist ein relevanter Prädiktor für Aspirationen und Bildungsentscheidungen von Schülerinnen und Schülern. Dabei besitzt das ASK für die Wahl eines Studienfaches sogar eine höhere Vorhersagekraft, als die individuelle Leistung (Parker et al., 2012). Die bisherige Selbstkonzept-Forschung bestätigt zudem die hohe Relevanz des Lernumfeldes für das ASK, da Schülerinnen und Schüler u.a. soziale Vergleiche mit ihren Peers als Informationsquelle für die Einschätzung der eigenen Fähigkeiten nutzen. Offen bleibt jedoch, welche Rolle das Lernumfeld, vermittelt über das akademische Selbstkonzept, für die Aspirationen und Studienfachwahlen junger Erwachsener spielt. Der Big-Fish-Little-Pond Effekt (BFLPE) beschreibt den wohl bekanntesten Kontexteffekt: Den negativen Effekt der gemittelten Schulleistung der Peers auf das ASK unter Kontrolle der individuellen Leistung (Marsh & Hau, 2003). Dieser Effekt gilt innerhalb der Schulzeit als gut repliziert und erste Studien konnten zudem zeigen, dass auch die Aspirationen von Schülerinnen und Schülern von diesem negativen Kontexteffekt betroffen sind (Nagengast & Marsh, 2012). Ob jedoch die konkreten Studienfachwahlen am Ende der Schullaufbahn ebenfalls den Einflüssen des BFLPE unterliegen, ist ungeklärt. Erste Studien lieferten ein inkonsistentes Bild, mit überwiegend kleinen oder negativen Effekten (Marsh, 1991; Marsh & O'Mara, 2010). Hierbei wurde jedoch nicht die konkrete Studienfachwahl im Anschluss an die Schulzeit erfasst, sondern der höchste erreichte Bildungsabschluss als Outcome betrachtet.

In diesem Beitrag adressieren wir 1) die Fragestellung, welchen Einfluss der Lernkontext am Ende der Schulzeit (Sekundarstufe II) auf das ASK und die Studienaspirationen als nicht-monetäre Erträge von Bildung hat. Darüber hinaus untersuchen wir 2) die Auswirkungen des BFLPE auf die konkrete Studienfachwahl nach dem Schulabschluss.

Anhand der Daten der längsschnittlichen BIJU-Studie (Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter) untersuchten wir, welche Einflüsse die individuelle Leistung in Mathematik, die aggregierte Leistung des Lernumfeldes in Mathematik und das mathematische Selbstkonzept 1) auf die Aspirationen ein MINT-Fach aufzunehmen und 2) auf die tatsächliche Wahl eines MINT-Studienfaches nach der Schule haben. Dabei betrachteten wir sowohl die direkten Effekte der Leistungsvariablen auf die Bildungsaspirationen und Bildungsentscheidungen, als auch die indirekten Effekte, die über das akademische Selbstkonzept vermittelt werden. Wir verwendeten logistische Regressionen unter Verwendung der KHB-Methode. In allen Analysen wurden das Geschlecht und der sozioökonomische Status der Personen kontrolliert. Die Stichprobe umfasste N = 3008 Personen (62,3% weiblich). Wir verwendeten Daten von mehreren Messzeitpunkten: In der 12. Klasse in der Oberstufe, zwei Jahre nach Schulabschluss und 10 Jahre nach Schulabschluss.

Es zeigte sich wie erwartet ein positiver Einfluss des akademischen Selbstkonzeptes in Mathematik auf die Aspiration ein MINT-Fach zu wählen. Die positiven Effekte zeigten sich ebenso für die tatsächliche Aufnahme eines Studienfaches im MINT-Bereich nach dem Schulabschluss. Unter Kontrolle individueller Leistungen, konnte kein direkter Effekt der aggregierten Leistung in Mathematik auf die Bildungsaspiration oder tatsächliche Studienfachwahl gefunden werden. Jedoch zeigten sich für beide Outcomes negative indirekte Effekte der aggregierten Leistung, die jeweils vollständig über das akademische Selbstkonzept vermittelt wurden.

Diese Ergebnisse verdeutlichen, dass das Lernumfeld nicht nur für das akademische Selbstkonzept und die Aspirationen während der Schulzeit entscheidend ist, sondern – vermittelt über das ASK – auch einen Einfluss auf die Wahl des Studienfaches hat. In anderen Worten: Eine Schülerin oder ein Schüler in einem leistungsstarken Umfeld im Bereich Mathematik wählte mit geringerer Wahrscheinlichkeit ein Studienfach im MINT-Bereich, als eine Schülerin oder ein Schüler der gleichen individuellen Leistungsfähigkeit in einem leistungsschwächeren Umfeld.

### **Typische und untypische Übergangsverläufe: Das emotionale Wohlbefinden junger Erwachsener nach dem Schulaustritt unter Berücksichtigung multipler Herausforderungen**

**Sabine Walper, Gabriela Gniewosz**

DJI, München

Die Planung des eigenen beruflichen Werdegangs, darin eingeschlossen die eigene Ausbildung, berufliche Karrierewege und Eintritt in den Arbeitsmarkt, sind Grundbausteine für die eigene (finanzielle) Unabhängigkeit und zunehmende Selbstständigkeit junger Erwachsener (Schulenberg & Schoon, 2012). Neben der beruflichen / akademischen Entwicklung stehen jedoch noch weitere zu meistern Aspekte der individuellen Lebensführung an diesem Übergang in das Erwachsenenalter im Vordergrund, wie es der Auszug aus dem Elternhaus und damit einhergehende Mobilitätsanforderungen darstellen (Blossfeld & Maurice, 2011). Einerseits kann die Bewältigung der genannten Herausforderungen durchaus als belastend erlebt werden und ist herausfordernd für die eigene psychische Verfasstheit. Andererseits ist davon auszugehen, dass individuelle

(gesundheitsbezogene) Ressourcen beeinflussen, wie Jugendliche diese multiplen Anforderungen meistern (Hale et al., 2015; Haas & Fosse, 2008).

Mit Fokus auf das emotionale Wohlbefinden untersucht diese Studie, 1) wie Jugendliche ihre Bildungs- und Berufswege gestalten und welche Rolle dabei der Aus- und Umzug aus dem Elternhaus spielt. Hierbei soll geprüft werden, 2a) welchen Einfluss das Wohlbefinden auf die Bewältigung der genannten wohnortsbezogenen und beruflichen Herausforderungen hat und 2b) wie das Wohlbefinden wiederum von den unterschiedlichen Anforderungen beeinflusst wird. Die Datenbasis bilden Informationen von 1.532 Jugendlichen (geb. 1991-1993), die an den jährlichen Befragungen des Deutschen Familienpanels "pairfam" teilnahmen. Die Daten wurden derart umstrukturiert, dass Informationen zum Wohlbefinden ein Jahr vor Schulaustritt (T0) sowie vier weitere Wellen nach Schulabschluss zur Verfügung standen.

Erste Analysen verweisen auf Basis der jährlich erfassten Erwerbs- und Bildungstätigkeiten auf fünf, über latente Klassenanalysen ermittelte, Verlaufstypen: Typischer akademischer (37,4%) und ausbildungsbezogener Verlauf (37,93%), später Ausbildungseinstieg mit vorheriger (weiterer) Bildungsepisode (7,69%) sowie einer Gruppe von Studienabbrechern (9,48%) und einem Verlaufsmuster von sich abwechselnden Berufs- und Erwerbslosigkeitsphasen (7,49%). Unter Berücksichtigung von Geschlecht, Migrationshintergrund und Schulabschluss, fand sich zunächst kein Einfluss des Wohlbefindens vor Schulaustritt auf die Bildungs- und Berufskarrieren. Hingegen zeigten sich deutliche negative Zusammenhänge zwischen den identifizierten Typen der Studienabbrecher und der unsteten Bildungs- und Erwerbsverläufen mit dem emotionalen Wohlbefinden. Demnach berichten Befragte dieser „Risikogruppen“ häufiger von einem verringerten Wohlbefinden bis zu drei Jahre nach dem Schulaustritt als Befragte der anderen Verlaufsmuster. Diese Erkenntnisse sollen in weiteren Analysen um die Erfahrungen des Wohnortwechsels, darin eingebunden das Alter bei Auszug (z. B. frühe oder später Auszug), Wohnortsituation (z. B. WG, Wohnheim) und Entfernung vom familiären Wohnort ergänzt werden. Dies soll, über die bestehenden Analysen hinaus helfen, ein vertieftes Verständnis über das Zusammenwirken multipler Herausforderungen für das Wohlbefinden an diesem Übergang von Jugend in das frühe Erwachsenenalter zu erhalten.

### **Bildung als Ressource für affektives Wohlbefinden im Alltag im Erwachsenenalter**

**Dave Möwisch<sup>1</sup>, Annette Brose<sup>2</sup>, Florian Schmiedeck<sup>3</sup>**

<sup>1</sup>DIPF, Berlin, <sup>2</sup>Humboldt-Universität zu Berlin, Berlin, <sup>3</sup>DIPF, Frankfurt a.M.

Subjektives Wohlbefinden (SWB) ist eines der einflussreichsten psychologischen Konstrukte zur Bewertung von Lebensqualität (Diener, 2000) und steht im Zusammenhang zu vielen wichtigen positiv assoziierten Variablen, wie Berufserfolg, sowie physischer und psychischer Gesundheit (Diener & Chan, 2011). Zum Zusammenhang von Wohlbefinden und Bildung gibt es diverse Modelle und empirische Studien (Desjardins, 2008; Jenkins & Wiggins, 2015), die versuchen, mögliche Wirkmechanismen zwischen Bildung und SWB zu beschreiben.

Zwei dafür relevante Bereiche stellen Arbeitszufriedenheit und Freizeitgestaltung dar, die elementare Bestandteile des Alltages sind. Empirische Studien konnten zeigen, dass Personen mit einem höheren Bildungsgrad eine höhere Arbeitszufriedenheit berichten (Hartog & Oosterbeck, 1998), die wiederum ein wichtiger Prädiktor für globalere Einschätzungen von SWB ist (Diener, Suh, Lucas & Smith, 1999). Hinsichtlich der Freizeitgestaltung sowie dem Wohlbefinden bei Freizeitaktivitäten und dem Zusammenhang zu Bildung gibt es nur vereinzelte Belege. Dye (1980) konnte zeigen, dass Personen mit einem höheren Bildungsgrad häufiger ehrenamtliche Tätigkeiten für die Gemeinschaft ausführen. Solche Aktivitäten, die von einem hohen Level an Bedeutsamkeit und Relevanz charakterisiert sind, fungieren als möglicher Mechanismus zwischen Bildung und Wohlbefinden.

Neben der kognitiven Komponente von SWB (Lebenszufriedenheit), kann ebenfalls dessen affektive Komponente (positiver und negativer Affekt) als Outcome untersucht werden, hat in der bisherigen Literatur im Kontext von Bildung allerdings nur vergleichsweise wenig Aufmerksamkeit erfahren.

Ziele der Analysen sind (1) Beziehungen zwischen Bildung und subjektivem Wohlbefinden (positiver und negativer Affekt, Lebenszufriedenheit) zu untersuchen und (2) mögliche Zusammenhänge von Bildung, der täglichen Freizeitgestaltung und des affektiven Erlebens täglicher Aktivitäten zu identifizieren.

In einer national repräsentativen Stichprobe des Sozio-oekonomischen Panels (SOEP, Richter & Schupp, 2012) wurden Daten von insgesamt 2401 Personen zu zwei Messzeitpunkten (2012: N = 2234; 2013: N = 1886) mithilfe der Day Reconstruction Method (DRM; Kahneman, Krueger, Schkade, Schwarz, & Stone, 2006) erhoben. Die Teilnehmer wurden im Rahmen der Datenerhebung instruiert, ihren vorangegangenen Tag in Episoden einzuteilen und diese mithilfe einer Liste von 24 möglichen Aktivitäten (z.B. „Arbeiten“ oder „Hausarbeit“) näher zu beschreiben. Außerdem wurden drei Episoden eines Tages zufällig ausgewählt und sollten von den Probanden hinsichtlich 10 Affektitems wie „Glück“, „Trauer“ oder „Ärger“ auf einer Skala von 1 = „gar nicht“ bis 7 = „sehr stark“ eingeschätzt werden. Zudem wurden personenspezifische Variablen wie Lebenszufriedenheit, Schul- und Hochschulbildung, sowie demografische Variablen erfasst. Bildung wurde durch die International Standard Classification of Education (ISCED) operationalisiert. Zur Modellierung der Affektstruktur wurden konfirmatorische Mehrebenenfaktoranalysen durchgeführt. Zur Untersuchung der Zusammenhänge zwischen positiven und negativen Affekt, Lebenszufriedenheit und Bildungsvariablen wurden Strukturgleichungsmodelle und Regressionsanalysen via Mplus verwendet.

Die Ergebnisse der Regressionsanalysen zeigen, dass es einen direkten negativen Zusammenhang von Bildung und negativem Affekt im Alltag gibt. Demnach berichten beispielsweise Personen mit einem höheren Bildungsgrad im Durchschnitt weniger Trauer und Einsamkeit bei ihren Alltagsaktivitäten als Personen mit einem niedrigeren Bildungsgrad, selbst wenn für den Erwerbsstatus kontrolliert wurde. Ebenfalls zeigte sich, dass die Effekte von Arbeitslosigkeit auf Affekt bei höherer Bildung signifikant schwächer waren, was auf eine puffernde Wirkung höherer Bildung hinweist.

## B20

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H120

### Bedrohung durch Geschlechterstereotype im Bildungsverlauf

Chair(s): **Andreas Sander** (TU Dortmund/Institut für Schulentwicklungsforschung), **Johannes Hasselhorn** (TU Dortmund/Institut für Schulentwicklungsforschung)

DiskutantIn(nen): **Ursula Kessels** (Freie Universität Berlin)

In dem Symposium werden exemplarische Beispiele für die Auswirkung von Geschlechterstereotypen in verschiedenen Etappen des Bildungsverlaufs thematisiert (Stereotype Threat). Vor dem Hintergrund einer stetigen gesellschaftlichen, aber auch wissenschaftlichen Diskussion über wiederkehrende empirische Befunde, dass Mädchen und Frauen geringere durchschnittliche Mathematikleistungen zeigen (für die Grundschule: z.B. Wendt et al., 2016) sowie geringere Anteile an Frauen in MINT Studiengängen (Quaiser-Pohl, 2012) vorzufinden sind, wird in diesem Symposium die Rolle von Stereotype Threat als ein Erklärungsansatz diskutiert (Spencer et al., 1998). Die Stereotype Threat Theorie geht davon aus, dass das Wissen über ein negatives Stereotyp über eine Gruppe, der man selbst angehört (z.B. Frauen in der mathematischen Domäne) dazu führt, dass das Stereotyp beispielsweise in Leistungssituationen bestätigt wird (Steele & Aronson, 1995). Im Gegensatz dazu können nicht stereotypisierte Personen bei der Aktivierung von Stereotypen der anderen Gruppe gegenüber paradoxerweise höhere Leistungen zeigen (Stereotype Lift; Walton & Cohen, 2003). In diesem Symposium werden in drei Beiträgen aktuelle Studien vorgestellt, die die Auswirkungen von Geschlechterstereotypen auf Leistungen im mathematischen Bereich in der Grundschule, der Sekundarstufe und im Studium beleuchten.

Im ersten Beitrag wird die Grundschule in den Blick genommen. In einem experimentellen Design wurden 116 Viertklässlerinnen und -klässler hinsichtlich des Einflusses von geschlechterstereotypisierten Testmaterialien und von Lösungsstrategien auf die Fähigkeiten zur mentalen Rotation untersucht. Es wurden unter anderem schnellere Reaktionszeiten bei Jungen im Vergleich zu Mädchen sowie Vorteile von Kindern mit holistischen Lösungsstrategien beobachtet.

Der zweite Beitrag fokussiert Schülerinnen und Schüler in der Sekundarstufe I. Stereotype Threat/Lift Effekte wurden hier unter Berücksichtigung des wichtigen Moderators der Domänenidentifikation untersucht. Über eine implizite Stereotypenaktivierung wurden 97 Zehntklässlerinnen und -klässler bei der Lösung eines Mathematiktests beeinflusst. Es zeigten sich Stereotype Threat Effekte bei Mädchen, jedoch keine Stereotype Lift Effekte bei Jungen. Mädchen mit hoher Mathematikidentifikation identifizierten sich nach Stereotypenaktivierung weniger mit ihrem Geschlecht und vice versa.

Der dritte Beitrag nimmt die Hochschule in den Fokus. In einer experimentellen Onlinestudie mit 129 Studierenden wurde untersucht, ob sich differentielle Stereotype Threat und Lift Effekte zwischen Männern und Frauen mit mathematiknäheren und -ferneren Studiengängen identifizieren lassen. Auch in diesem Beitrag wurde eine implizitere Stereotypenaktivierung implementiert. Frauen zeigten in mathematiknahen Studiengängen schwächere Testleistungen in der Experimentalbedingung.

Abschließend wird das Symposium aus inhaltlicher Perspektive und in Bezug auf die verwendeten Designs sowie Methoden diskutiert. Mit den Themenschwerpunkten Stereotype, Geschlecht und Mathematik in verschiedenen Bildungsetappen werden potenziell kumulierende Auswirkungen von Geschlechterstereotypen in den Blick genommen, die eine besondere Herausforderung für professionelles Handeln von Akteuren im Bildungssystem, aber auch für die Bildungsforschung beispielsweise hinsichtlich des methodischen Umgangs in Fragebogenstudien bietet.

#### Beiträge des Symposiums

### Der Einfluss geschlechterstereotypisierter Stimuli auf Lösungsstrategien von Viertklässlern bei Aufgaben zur mentalen Rotation

**Vera Ruthsatz, Martina Rahe, Linda Schürmann, Claudia Quaiser-Pohl**  
Universität Koblenz-Landau, Campus Koblenz, Institut für Psychologie

Zur Messung der mentalen Rotationsfähigkeit wird üblicherweise auf die Würfelfiguren der Pionierstudie von Shepard und Metzler (1971) zurückgegriffen. Damit konnte in zahlreichen Untersuchungen ein reliabler Geschlechterunterschied zugunsten von Männern festgestellt werden (Voyer, Voyer & Bryden, 1995). Neben biologischen Ursachen für diesen Geschlechterunterschied finden sich auch Einflüsse von sozio-kognitiven Faktoren und verschiedenen Leistungsfaktoren (z.B. Zeitlimit, Rotationsachse) auf die Fähigkeit zur mentalen Rotation (Estes & Felker, 2012; Goldstein, Haldane & Mitchell, 1990; Hausmann, Slabbekoorn, van Goozen & Güntürkün, 2000; A. Moè & Pazzaglia, 2010; Angelica Moè, 2016). Zum Beispiel kann sich ein stereotype threat Effekt nachteilig auf die Leistung auswirken (Moè & Pazzaglia, 2006). Dieser Effekt kann außer durch Instruktionen (Neuburger, Jansen, Heil & Quaiser-Pohl, 2012), auch durch den Test selbst hervorgerufen werden (Ramirez, Gunderson, Levine & Beilock, 2012). Studien legen nahe, dass die Vertrautheit mit dem Stimulusmaterial auch die Wahl der Lösungsstrategie und deren Effizienz beeinflussen kann (Jansen, Quaiser-Pohl, Neuburger & Ruthsatz, 2014; Ruthsatz, Neuburger, Rahe, Jansen & Quaiser-Pohl, 2017).

Hauptfragestellung der vorliegenden Studie ist, ob geschlechterstereotypisierte Stimuli Lösungsstrategien und Reaktionszeiten von Jungen und Mädchen beeinflussen. Dazu wurde eine chronometrische Version aus den Stimuli des männlich stereotypisierten M-MRT-3D (z.B. Säge, Zug) und des weiblich stereotypisierten F-MRT-3D (z.B. Nagellack, Puppe) von Ruthsatz et al. (2017) erstellt. Für den mentalen Rotationstest wurde eine schnellere Reaktionszeit für geschlechterkongruente Stimuli erwartet (H1). Bezugnehmend auf frühere Studien mit geschlechterstereotypen Testmaterialien (Ruthsatz et al., 2014, 2017), wurde nur für die Bedingung mit männlich stereotypisierten Items ein Geschlechterunterschied in der Fehlerrate und der Reaktionszeit angenommen (H2). Da sich holistische Strategien als effizienter gegenüber analytischen erwiesen haben (z.B. Geiser, Lehmann & Eid, 2006), wurde eine schnellere Reaktionszeit bei holistischen Lösungsstrategien erwartet (H3).

Als Versuchspersonen wurden Viertklässler ausgewählt, weil in diesem Alter das Bewusstsein für Stereotype immer stärker wird und vermehrt die Handlungen der Kinder beeinflusst (Clark McKown & Weinstein, 2003). Dies zeigt sich unter anderem darin, dass räumliche Aktivitäten (z.B. Spielen mit LEGO®) in diesem Alter geschlechterübergreifend als typische Jungenaktivitäten angesehen werden (Neuburger, Jansen, Heil & Quaiser-Pohl, 2013). Daraus resultiert ein unterschiedlicher Umgang mit räumlichen Aufgaben, der die Erfahrungshorizonte und zudem Leistungen und Erfolge im MINT-Bereich beeinflusst (Fennema & Sherman, 1977; Quaiser-Pohl & Lehmann, 2002).

Erfasst wurden Reaktionszeit, Fehlerrate und Rotationsgeschwindigkeit von 65 Mädchen und 51 Jungen (Alter: M=9,58; SD=0,6) in einem 2 (Geschlecht) x 2 (Stimulusmaterial)-Zwischensubjekt-Design. Anschließend beantworteten alle Kinder einen Fragebogen zu ihren Lösungsstrategien und den Zahlenverbindungstest (ZVT, Oswald & Roth, 1987).

Wie erwartet wurden in geschlechterkongruenten Bedingungen signifikant schnellere Reaktionszeiten gemessen. Geschlechterunterschiede in der Reaktionszeit zeigten sich nur bei den männlich stereotypisierten Stimuli, bei welchen Mädchen signifikant langsamere Reaktionszeiten aufwiesen als Jungen. Für die Fehlerquote wurde nur ein Haupteffekt des Stimulusmaterials gefunden: Jungen schnitten in der Bedingung mit weiblichen stereotypisierten Stimuli besser ab als Jungen in der Bedingung mit männlich stereotypisierten Stimuli. Bei Mädchen zeigten sich keine signifikanten Unterschiede. Mit Bezug auf Lösungsstrategie und Geschlechterkongruenz der Stimuli wurden alle Hypothesen bestätigt. Kinder, die holistische Strategien berichteten, zeigten schnellere Reaktionszeiten als Kinder mit analytischen Strategien. Eine signifikante Interaktion von Strategie und Geschlechterkongruenz der Stimuli zeigte die Überlegenheit einer holistischen Strategie nur für Bedingungen, in denen Geschlecht und Stereotypisiertheit des Stimulusmaterials nicht übereinstimmten. In geschlechterkongruenten Bedingungen zeigten Kinder mit analytischen Strategien keine signifikant schlechteren Reaktionszeiten als Kinder mit holistischen Strategien.

Die Ergebnisse zeigen einen Einfluss der Übereinstimmung von Geschlecht und Stereotypisiertheit des Stimulusmaterials auf die Effizienz von Lösungsstrategien. Da üblicherweise Shepard und Metzlers Würfelfiguren verwendet werden, um die mentale Rotationsleistung zu erfassen, scheint es bedeutsam zu hinterfragen, ob Würfelfiguren für Jungen auch einen Vorteil bieten, da sie in der Regel vertrauter mit ähnlich aussehenden Konstruktionsspielen wie beispielsweise LEGO® sind (Kersh, Casey & Mercer Young, 2008).

## **„Bin ich Frau genug?“ – Stereotype Threat im Mathematikunterricht als Quelle und Bedrohung der Identifikation mit dem eigenen Geschlecht**

**Johanna Hermann, Regina Vollmeyer**

Goethe-Universität Frankfurt, Institut für Psychologie, Arbeitsbereich Pädagogische Psychologie

Ziel dieser Studie war es zu untersuchen, wie Schüler/-innen auf die Aktivierung des Stereotyps „nur Jungen können Mathe“ im Unterricht reagieren. Obwohl Geschlechtsunterschiede in Mathematik geringer geworden sind (Klieme et al., 2010), entscheiden sich immer noch wenige Mädchen für entsprechende Studiengänge oder Berufe (Quaiser-Pohl, 2012). Eine Erklärungsmöglichkeit bietet der Stereotype-Threat-Ansatz (Steele & Aronson, 1995). Das Phänomen beschreibt einen situativen Leistungseinbruch in Folge negativer Stereotypisierung, der langfristig auch zur Desidentifikation mit einer Domäne führen kann (Deemer, Lin & Soto, 2016). Sind Personen hingegen von Stereotypen indirekt positiv betroffen, wie in Mathematik die Jungen, kann ein Leistungsanstieg erfolgen (Walton & Cohen, 2003). Stereotype-Threat-Effekte zeigen sich zudem gerade bei Mädchen, die sich stark mit Mathematik identifizieren (Hermann & Vollmeyer, 2017), also geeignet für ein mathematisches Studium wären. Haben Mädchen Mathematik „bereits abgeschrieben“, scheinen sie „immun“ gegenüber stereotyper Bedrohung. Mädchen mit hoher Mathematikidentifikation können sich hingegen nur von der eigenen Gruppe distanzieren, um einem Stereotype Threat zu entgehen. So zeigt sich, dass sich Frauen mit hoher Mathematikidentifikation nach stereotyper Bedrohung von femininen Eigenschaften distanzieren (Pronin, Steele & Ross, 2004). Daher soll auch geprüft werden, wie sich die stereotype Bedrohung der Schülerinnen auf deren Identifikation mit dem eigenen Geschlecht auswirkt, das gerade in der Pubertät hohe Selbstwertrelevanz besitzt (Witt & Wood, 2010).

Hypothesen: Wir gehen davon aus, dass eine Stereotyp-Aktivierung zu einer schlechteren Mathematikleistung im Vergleich mit einer weiblichen Kontrollgruppe führt (Hypothese 1a). Für Jungen wird in der Experimentalgruppe hingegen ein Leistungsanstieg im Vergleich mit der Kontrollgruppe erwartet (Hypothese 1b). Zudem erwarten wir, dass negativ stereotypisierte Mädchen mit hoher Mathematikidentifikation eine geringere Identifikation mit ihrem Geschlecht aufweisen (Hypothese 2a). Bei Mädchen mit geringer Mathematikidentifikation nehmen wir hingegen an, dass die stereotype Bedrohung mit einer stärkeren Identifikation mit dem eigenen Geschlecht einhergeht (Hypothese 2b).

Methode: An der Studie nahmen 97 Schüler/-innen der 10. Jahrgangsstufe teil (55m, 42w; Alter  $M = 15.37$ ,  $SD = .63$ ), die einen Fragebogen zur Mathematikidentifikation ausfüllten. Die Stereotyp-Aktivierung erfolgte anschließend über einen Zeitungsartikel, der entweder die Mathematiküberlegenheit der Jungen oder Mathematik in der Alltagssprache thematisierte. Danach wurde ein Mathematiktest absolviert, von dem nur die schwierigen Aufgaben in die Analyse eingingen (Lösungswahrscheinlichkeit  $< 50\%$ ), bei denen Stereotype Threat verstärkt auftritt (Keller, 2007). Abschließend wurde die Identifikation mit dem Geschlecht mittels Fragebogen erfasst.

Ergebnisse: Erwartungskonform zeigte sich hinsichtlich der Mathematikleistung unter Kontrolle der Mathematikidentifikation eine signifikante Interaktion zwischen Geschlecht und Experimentalgruppe,  $F(1,92) = 4.56$ ,  $p = .04$ . Mädchen der Stereotype-Threat-Gruppe schnitten signifikant schlechter ab ( $M = .21$ ,  $SD = .21$ ) als Mädchen der Kontrollgruppe ( $M = .37$ ,  $SD = .23$ ),  $t(93) = 2.10$ ,  $p = .04$ , sodass Hypothese 1a bestätigt werden konnte. Hypothese 1b musste hingegen verworfen werden, da kein Lift-Effekt bei den Jungen vorlag. Hinsichtlich der Effekte stereotyper Bedrohung auf die Geschlechtsidentifikation der Mädchen, ergab sich eine signifikante Interaktion zwischen Mathematikidentifikation und Experimentalgruppe,  $F(1,38) = 7.56$ ,  $p = .01$ . Konform der Hypothese 2a identifizierten sich Mädchen mit hoher Mathematikidentifikation nach stereotyper Bedrohung signifikant weniger mit ihrem Geschlecht ( $M = 4.97$ ,  $SD = .80$ ) als unter neutralen Bedingungen ( $M = 5.75$ ,  $SD = .56$ ),  $t(38) = 2.19$ ,  $p = .03$ . Mädchen, die sich nur wenig mit Mathematik identifizierten, verfügten nach stereotyper Bedrohung hingegen über eine höhere Geschlechtsidentifikation ( $M = 5.18$ ,  $SD = .74$ ) als in der Kontrollgruppe ( $M = 4.52$ ,  $SD = .99$ ), wobei es sich jedoch um keinen signifikanten Unterschied handelte,  $t(38) = 1.72$ ,  $p = .09$ , sodass Hypothese 2b verworfen werden muss.

Diskussion: Implikationen für einen gendergerechten Mathematikunterricht werden diskutiert. Dabei soll im Fokus stehen, an wen sich Fördermaßnahmen richten und wie diese gestaltet werden sollten.

## **Geschlechteridentität und -orientierung im Studium – Stereotypenbedrohung beim Lösen von mathematischen Aufgaben**

**Andreas Sander, Johannes Hasselhorn**

TU Dortmund/Institut für Schulentwicklungsforschung

Theoretischer Hintergrund

Frauen bzw. Mädchen werden in mathematisch-naturwissenschaftlichen Bereichen durchschnittlich geringere Leistungen als Männern bzw. Jungen nachgesagt. Dies wurde z.T. auch bereits empirisch nachgewiesen (z.B. Wendt et al. 2016). Eine Erklärung für geringere Outcomes im MINT Bereich und auch weiterführend für Interessen-/Studien-/Berufswahl können in gesellschaftlich verankerten Stereotypen liegen. Diese Stereotype können sich bei Aktivierung in Test- und Lernsituationen negativ auf die individuellen Outcomes auswirken (z.B. Stereotype Threat; Steele & Aronson, 1995). Hingegen können nicht Betroffene bei der Aktivierung eines gegen eine andere Gruppe gerichteten Stereotypes situative Leistungssteigerungen zeigen (Stereotype Lift; Walton & Cohen, 2003). Eine aktuelle Experimentalstudie von Tempel und Neumann (2016) konnte darlegen,



dass Frauen mit einer femininen Geschlechtsrollenorientierung in einem Mathematiktest von negativen Stereotypen bedroht wurden und daher schlechtere mittlere Mathematikleistungen erzielten. Auch wird in der Stereotype Threat Forschung davon ausgegangen, dass domänenspezifische Identifikation eine große Rolle hinsichtlich der Stereotypenbedrohungsvulnerabilität spielt und hierbei stärker identifizierte eher anfällig für stereotypeninduzierte Leistungsreduktion sind als weniger stark identifizierte Personen (z.B. Keller, 2007), obwohl dies in empirischen Studien nicht immer Berücksichtigung findet.

#### Fragestellung

In dieser Studie soll die Rolle der Domänenidentifikation für Stereotype Threat bzw. Lift zentral in den Blick genommen werden. Am Beispiel Mathematik und Geschlechterstereotype wird daher untersucht, ob eine Identifikation mit der Domäne Mathematik theoriekonform Stereotype Threat bzw. Lift begünstigt.

#### Methode

In einem Online-Experiment wurden Anfang 2017 N = 129 freiwillig teilnehmende Studierende (59.7% weiblich, MAlter = 25.89 [SD = 5.59]; 74.4% mathematiknahe Studienfächer) zufällig auf zwei Experimentalbedingungen (EB) aufgeteilt und absolvierten einen zweiteiligen, 40 Items umfassenden Mathematiktest ( $\alpha_{\text{Gesamt}} = .89$ ; angelehnt an Rechenaufgaben des I-S-T 2000 R; Liepmann et al., 2007). In Anlehnung an Tempel und Neumann (2016) beantworteten Teilnehmende in EB1 zur Geschlechtsidentitätsaktivierung vor den Testaufgaben Fragen zu Geschlechtsrollen ( $\alpha_{\text{Gesamt}} = .88$ ; Deutsche Version des Bem Sex-Role-Inventory [BSRI]; Schneider-Düker & Kohler, 1988) und eine direkte Frage zum Geschlecht. Teilnehmende in EB2 beantworteten diese Fragen nach beiden Mathematiktestteilen. Darüber hinaus wurde nach dem aktuellen Studiengang gefragt. Diese Angabe wurde nach Abschluss der Datenerhebung in eine dichotome Variable mit den Ausprägungen mathematiknaher bzw. mathematikferner Studiengang überführt.

#### Ergebnisse

Frauen zeigten in dieser Stichprobe durchschnittlich statistisch signifikant weniger korrekt gelöste Rechenaufgaben ( $M = 24.95$  [SD = 6.81]) als Männer ( $M = 28.92$  [SD = 7.17];  $t(127) = 3.18$ ,  $p < .01$ ,  $d = 0.57$ ) sowie Studierende mathematiknaher Studiengänge höhere Resultate ( $M = 27.28$  [SD = 6.83]) als Studierende anderer Studiengänge ( $M = 24.42$  [SD = 7.92];  $t(127) = 1.99$ ,  $p = .05$ ,  $d = 0.40$ ).

Zur Prüfung der zugrundeliegenden Fragestellung wurden aufeinander aufbauende T-Tests berechnet. Da für Frauen ein Stereotype Threat und für Männer ein Stereotype Lift vermutet wurde, wurden die Analysen getrennt nach den Geschlechtern und mit einseitigem Signifikanztest berechnet. Ein Vergleich der Testleistungen von Frauen zwischen den Experimentalbedingungen ergab keinen statistisch signifikanten Unterschied ( $t(75) = 1.17$ ,  $p = .12$ ). Eine Eingrenzung der Datenbasis auf solche Frauen, die in mathematiknahen Studiengängen studierten ergab wiederum einen bedeutsamen Unterschied bezogen auf die Experimentalbedingung ( $t(53) = 1.92$ ,  $p = .03$ ,  $d = 0.52$ ). Für die Substichprobe der Männer konnten keinerlei Unterschiede in der Testleistung zwischen den Experimentalbedingungen gefunden werden.

Insgesamt konnten in der vorliegenden Studierendenstichprobe Hinweise auf einen Stereotype Threat Effekt bezogen auf Mathematikleistungen von Frauen gefunden werden, der sich allerdings ausschließlich bei Frauen in mathematiknahen Studiengängen zeigte. Dies kann dahingehend interpretiert werden, dass Stereotype Threat vor allem bei hoher Domänenidentifikation wirksam wird. Für Stereotype Lift konnten keinerlei Belege gefunden werden.

## B21

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 30, H148

### Individuelle Förderung an Ganztagschulen als zentrale Herausforderung für professionelles Handeln

*Chair(s):* **Désirée Theis** (Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland), **Heinz Günter Holtappels** (Institut für Schulentwicklungsforschung, TU Dortmund), **Markus Sauerwein** (Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland)

*DiskutantIn(nen):* **Ferdinand Stebner** (Ruhr-Universität Bochum)

Lehrkräfte sind durch Transformationsprozesse von Schule mit einer stetigen Ausweitung ihres Tätigkeitsspektrums konfrontiert – insbesondere an Ganztagschulen (Sauerwein, 2017). Die Lehrkräfte sind hier nicht nur als unterrichtende Fachkraft gefragt, sondern sollen ebenfalls Schüler\*innen in Ganztagsangeboten individuell fördern, Integrationsprozesse anstoßen und zugleich mit weiteren pädagogischen Fachkräften kooperieren (Böttcher, Maykus, Altermann, & Liesegang, 2014). Dieses zusätzliche Aufgabenspektrum wird auch in politischen Zielvorgaben festgehalten, wie etwa in Erklärungen der ‚Ständige Konferenz der Kultusminister der Länder‘ (KMK), die sich mehr individuelle Förderung durch Ganztagschule erhofft (z.B., Fischer, Radisch, Theis, & Züchner, 2012). Demnach soll durch ein Mehr an Zeit und lernwirksame außerunterrichtliche Lerngelegenheiten mehr individuelle Förderung ermöglicht werden. Multiprofessionelle Kooperation ist dabei eine unverzichtbare Voraussetzung (Böttcher et al., 2014). Die Einführung der Ganztagschule mit Fokus auf individuelle Förderung der Schüler\*innen ist somit als zentrale Herausforderungen für professionelles Handeln zu sehen.

Aktuelle Ergebnisse der internationalen Vergleichsstudien PISA 2015 (Reiss et al., 2016) und TIMSS 2015 (Wendt et al., 2016) weisen erneut darauf hin, dass individuelle Förderung im deutschen Bildungssystem zu verbessern ist. Allerdings zeigen sich unterschiedliche Konzepte und Verständnisse individueller Förderung: Leistungsdefizite sollen kompensiert werden, besondere Begabungen und unterschiedliche Interessen sowie Potenziale der Schüler\*innen gefördert werden (Klieme & Warwas, 2011).

An fast allen Ganztagschulen werden außerunterrichtliche Fördermaßnahmen bzw. individuelle Lerngelegenheiten angeboten (StEG-Konsortium, 2015). Gleichzeitig nimmt jedoch nur eine Minderheit der ganztägig Lernenden an diesen Angeboten teil (StEG-Konsortium, 2016) und eine Verbesserung fachlicher Kompetenzen konnte bisher nicht nachgewiesen werden (Lossen et al., 2016; Fischer et al., 2016). Hierzu passend geben ca. 50% der Eltern an, mit der Förderung in Ganztagschulen unzufrieden zu sein. Dementsprechend stehen „mehr Förderangebote“ an der Spitze der Verbesserungsforderungen der Eltern (Hollenbach-Biele & Zorn, 2016; Holtappels, 2013). Folglich besteht in Bezug auf individuelle Förderung ein erhöhter Entwicklungs- (Holtappels, 2014) und Untersuchungsbedarf (Fischer et al., 2014).

Basierend auf Daten verschiedener Teilstudien der Studie zur Entwicklung von Ganztagschulen (StEG) zielt das vorliegende Symposium darauf ab, einen Überblick zur Rolle der individuellen Förderung an Ganztagschulen als zentrale Herausforderung für professionelles Handeln zu geben und Gelingensbedingungen individueller Förderung zu untersuchen. Beitrag 1 liefert Befunde zur Sichtweise der Eltern auf individuelle Förderung und Ganztagschule. Es wird der Frage nachgegangen, welche Vorstellungen Eltern von Ganztagschulen haben und inwieweit diese Schulform in ihren Augen ein Ort für individuelle Förderung ist. Der zweite Beitrag untersucht den Zusammenhang zwischen professionellen Handlungskomponenten von Lehrkräften mit Fokus auf individuelle Förderung und Kooperation und der Entwicklung von Fachkompetenzen der Schülerinnen und Schüler an Ganztagsgrundschulen. Der dritte Beitrag ermittelt, inwiefern die individuell wahrgenommene Unterstützung durch das Personal in Leseangeboten mit der Lesekompetenz und dem Leseselbstkonzept von Fünftklässlern zusammenhängt. Im Rahmen des vierten Beitrags wird ein Entwicklungsmodell zur Abstimmung professionellen Handelns von Lehrerinnen und Lehrern sowie Personen des weiteren pädagogisch tätigen Personals an Ganztagschulen aus der Praxis herausgearbeitet. Die Beiträge werden abschließend diskutiert.

*Beiträge des Symposiums*

### Elterliche Vorstellungen von individueller Förderung und Ganztagschule

**Bettina Arnoldt, Peter Furthmüller, Susanne Gerleigner**  
Deutsches Jugendinstitut

Zwei Ziele des Ganztagschulenausbaus in Deutschland waren, die Leistungen der Schüler/-innen zu verbessern und die soziale Ungleichheit im Bildungssystem zu verringern (BMBF, 2003). Erwartet wurde, dass durch das Mehr an Zeit an Ganztagschulen eine bessere individuelle Förderung möglich ist und dadurch die beiden Ziele erreicht werden können. Bisherige Studien belegen, dass dies unter bestimmten Voraussetzungen auch gelingt. Allen voran gelten qualitativ hochwertige Angebote von entsprechend qualifiziertem Personal sowie eine intensive Teilnahme an den Angeboten als wichtiger Grundstein für positive Wirkungen von Ganztagschulen (Fischer et al., 2011). Eine Voraussetzung für gelingende Ganztagschulen sind somit motivierte Kinder, die dauerhaft an den Angeboten teilnehmen. Hierfür ist einerseits eine positive Einstellung gegenüber Ganztagschule auf Seiten der Kinder wichtig. Bedenkt man, dass insbesondere im Grundschulalter Kinder kaum eigenmächtig Entscheidungen in schulischen Belangen treffen (Quellenberg, 2010), ist andererseits auch die Einstellung der Eltern gegenüber Ganztagschule ein wichtiges Kriterium dafür, wie Kinder an Angeboten des Ganztags teilnehmen und ob Ganztagschule dadurch individuelle Förderung optimieren kann.

Deshalb geht vorliegender Beitrag zunächst der Frage nach, wie Kinder nach Ansicht ihrer Eltern von Schule gefördert werden sollten. Anschließend wird untersucht, welche Vorstellungen Eltern von Ganztagschulen haben und inwieweit diese Schulform in ihren Augen ein Ort für individuelle Förderung ist. Darauf aufbauend wird abschließend überprüft, ob und wenn ja wie sich diese Vorstellungen in Familien systematisch unterscheiden.

So wird beispielsweise davon ausgegangen, dass Eltern mit höherem Bildungsstatus in Anlehnung an Bourdieus Habituskonzept (Bourdieu, 1987) vom Ganztagsschulbesuch weniger eine gezielte zusätzliche Leistungsverbesserung ihres Kindes erwarten, als vielmehr auf eine ganzheitliche und allumfassende Persönlichkeitsentwicklung des Kindes Wert legen. Dahinter steckt die Annahme, dass höher gebildete Eltern mit entsprechender Kapitalausstattung für die reine Leistungsverbesserung zum einen auch andere Ressourcen haben, die sie bei Bedarf einsetzen können (z.B. die Möglichkeit, das Kind bei den Hausaufgaben selbst zu unterstützen oder die Finanzierung von Nachhilfe). Zum anderen werden sie der formellen Bildung in der Schule nicht dasselbe Gewicht zukommen lassen, wie die bei Bourdieu sogenannten Bildungsaufsteiger, deren Hoffnung es ist, durch (Schul-)Bildung in der Gesellschaft aufzusteigen. Eltern mit geringerer Bildung würden somit die Möglichkeiten der Ganztagschule eher in Anspruch nehmen; unter anderem, weil sie selbst den Kindern weniger Unterstützung zukommen lassen können.

Zur Beantwortung der oben skizzierten Forschungsfragen wurde eine Stichprobe verwendet, die auf der Einwohnermeldeamtsstichprobe des DJI-Surveys Aufwachsen in Deutschland: Alltagswelten (AID:A) basiert und sich aus Teilnehmer/-innen zusammensetzt, die Angaben, mit einem Kind aus den Geburtsjahrgängen 2003 bis 2009 im Haushalt zu leben. Ein Teil dieser Personen wurde im Rahmen des Teilprojekts Bildungsorte der Studie zur Entwicklung von Ganztagsgrundschulen (StEG) erneut telefonisch befragt, so dass insgesamt Angaben von 2.679 Eltern und 2.145 Schülern zum Themenkomplex Ganztagsgrundschule zur Verfügung stehen. Für vorliegenden Beitrag wurden die Angaben der im Frühjahr 2017 befragten Erziehungsberechtigten von Grundschüler/-innen in Halb- und Ganztagsgrundschulen verwendet (n=1387).

Mittels Gruppenvergleichen sowie multipler Regressionsanalysen soll untersucht werden, inwieweit sich Elterngruppen (z.B. nach Bildungsstand, nach besuchter Klassenstufe der Kinder oder nach bisherigen Erfahrungen mit Ganztagsgrundschulen) hinsichtlich ihrer Einstellungen zu individueller Förderung an Schulen allgemein, sowie der Akzeptanz von Ganztagsgrundschulen im Besonderen unterscheiden.

Erste deskriptive Ergebnisse weisen darauf hin, dass Eltern die Persönlichkeitsentwicklung der Kinder als oberstes Ziel schulischer Förderung betrachten und nicht primär die Verbesserung von Schulleistungen. Unterscheidet man jedoch nach Berufsabschluss, zeigt sich, dass Eltern ohne akademische Ausbildung der Verbesserung von Schulleistungen signifikant häufiger als wichtiges Ziel schulischer Förderung zustimmen als Eltern mit akademischer Ausbildung. Auch bezüglich der Einstellung gegenüber Ganztagsgrundschule verweisen erste deskriptive Analysen auf Unterschiede nach Berufsabschluss. Während nur gut ein Drittel derer mit akademischer Ausbildung der Aussage ganz oder teilweise zustimmen, dass Ganztagsgrundschule für die Kinder zu anstrengend ist, sind dies bei Nichtakademikern rund die Hälfte der Befragten.

### **Individuelle Förderung und Kompetenzentwicklung an Ganztagsgrundschulen: In welchem Zusammenhang stehen förderungsrelevante Lehrermerkmale professionellen Handelns mit der Entwicklung der Schülerkompetenzen?**

**Katja Tillmann, Karin Lossen, Heinz Günter Holtappels, Frederik Osadnik**  
Institut für Schulentwicklungsforschung, TU Dortmund

Seit PISA 2000 ist die Forderung nach individueller Förderung zentraler Bestandteil bildungspolitischer Diskurse (Klieme & Warwas, 2011), nicht zuletzt mit Blick auf ganztägig organisierte Schulen (Aktionsrat Bildung, 2013). Diese Forderung ist verknüpft mit der Annahme, dass individualisierte Förderung im unterrichtlichen und außerunterrichtlichen Bereich von ganztägig organisierten Schulen zu verbesserter Lern- und Kompetenzentwicklung der Schüler\*innen beitragen kann (KMK, 2002). Um Kinder mit unterschiedlichen Lernvoraussetzungen optimal zu unterstützen, ist eine inhaltlich-konzeptionelle und organisatorische Verzahnung der unterrichtlichen und außerunterrichtlichen Lerngelegenheiten notwendig. Dies stellt hohe Anforderungen an professionelles Lehrerhandeln und setzt die Kooperation der Lehrkräfte (LK) untereinander und mit dem weiteren pädagogisch tätigen Personal voraus (Böttcher et al., 2014). In der schulischen Praxis ist jedoch sowohl die systematische Verankerung individueller Förderung (Maykus et al., 2011) als auch die Kooperationskultur (Tillmann & Rollett, 2010) durch hohe Heterogenität gekennzeichnet.

Im Rahmen dieses Beitrags wird untersucht, inwieweit sich Ganztagsgrundschulen in professionellen Handlungskomponenten von LK mit Fokus auf individuelle Förderung und Kooperation unterscheiden und in welchem Zusammenhang dies mit der Entwicklung von Fachkompetenzen der Schüler\*innen (SuS) steht.

Forschungsfragen:

- (1) Lassen sich die Lehrkräfte an Ganztagsgrundschulen anhand ihrer eigenen Angaben zu verschiedenen Aspekten professionellen Handelns mit Fokus auf individuelle Förderung in unterschiedliche Typen gruppieren? Wie lassen sich diese Typen beschreiben?
- (2) Wie verteilen sich die ‚Lehrkrafttypen‘ auf die Schulen? Sind die Typen an verschiedenen Schulen unterschiedlich stark vertreten?
- (3) Zeigen sich auf Schulebene systematische Unterschiede hinsichtlich der Kompetenzentwicklung der SuS in den Bereichen Lesen und Naturwissenschaft in Abhängigkeit von der Ausrichtung des professionellen Handelns der Lehrerschaft auf individuelle Förderung?

Im Rahmen der Studie StEG-P wurden von 2013 bis 2014 an Ganztagsgrundschulen LK zu verschiedenen Aspekten professionellen Handelns befragt. Die Analysen basieren auf Angaben von 636 LK an 60 Schulen zu folgenden professionellen Handlungskomponenten mit Fokus auf individuelle Förderung: Bezugsnorm bei Leistungsbewertung, professionelle Teamarbeit mit Fokus auf Schülerlernen sowie auf Analyse, Diagnose und Evaluation, schülerbezogene und unterrichtsbezogene multiprofessionelle Kooperation, Schulentwicklung bzgl. individueller Förderung und Einbindung der Lehrkräfte in den Ganztagsbetrieb. Alle Skalen weisen gute Reliabilitäten auf (alpha .74-.92).

Zur Erfassung der Kompetenzentwicklung werden die über standardisierte Testverfahren zu jeweils zwei Messzeitpunkten erhobenen naturwissenschaftlichen Kompetenzen sowie die Lesekompetenzen der SuS des dritten und vierten Jahrgangs verwendet (Lossen et al., 2016; Tillmann et al., 2018).

Nach der Analyse korrelativer Beziehungen zwischen den Konstrukten des professionellen Lehrerhandelns mit Fokus auf individuelle Förderung und den Schülerkompetenzen werden die Lehrkräfte mithilfe einer Latent-Class-Analyse (LCA) (Hagenaars & McCutcheon, 2002; Muthén & Muthén, 2008) zu Typen gruppiert. Mit diesem Verfahren wird datengestützt eine latente Variable gebildet, die die Merkmalsträger auf Grundlage der verwendeten Konstrukte in homogene Gruppen einteilt (Willemss & Spillebeen, 2014). Mit regressionsanalytischen Verfahren wird untersucht, inwiefern die Ausrichtung des professionellen Handelns der Lehrerkollegien auf individuelle Förderung einen Effekt auf die Kompetenzentwicklung der SuS zeigt.

Erwartete Ergebnisse:

Es wird erwartet, dass sich verschiedene charakteristische ‚Lehrkrafttypen‘ finden lassen, die bezüglich ihrer förderungsrelevanten Lehrermerkmale deutlich differieren. Zudem wird angenommen, dass sich die Zusammensetzung der Lehrerschaft hinsichtlich der unterschiedlichen Typen zwischen verschiedenen Schulen wesentlich unterscheidet. Im Vortrag werden vertiefende Analysen zur Frage des Zusammenhangs zwischen Ausrichtung der Lehrerschaft auf individuelle Förderung und den Lernergebnissen der Kinder vorgestellt. Annahmen sind hier in zweierlei Richtung plausibel: Möglicherweise erzielen SuS an Schulen mit einem hohen Anteil an Lehrkräften, die in ihrem professionellen Handeln stark auf individuelle Förderung orientiert sind, im Durchschnitt einen höheren Leistungszuwachs. Denkbar ist ebenso, dass gerade

Lehrerkollegien an Schulen mit durchschnittlich eher leistungsschwächerer Schülerschaft besonders stark auf individuelle Förderung setzen. Implikationen für Forschung und Praxis werden diskutiert.

## **Leseförderung in Ganztagsangeboten und im Unterricht an Ganztagschulen der Sekundarstufe I: Wie wirkt die erlebte individuelle Unterstützung durch das Personal?**

**Markus Sauerwein<sup>1</sup>, Désirée Theis<sup>1</sup>, Natalie Fischer<sup>2</sup>, Heyl Katrin<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, <sup>2</sup>Universität Kassel

Ganztagsangebote sollen helfen, Schüler\*innen individuell zu fördern und so Leistungsdefizite auszugleichen (Arbeitsstab Forum Bildung, 2001). Bisher gibt es jedoch wenig Hinweise auf Leistungssteigerungen durch den Besuch von Ganztagsangeboten (Decristan & Klieme, 2016). Die Teilnahme der Schüler\*innen unterliegt Selektionsmechanismen (Tillmann et al., 2017) und Fischer et al. (2016) zeigen, dass sich die Lesekompetenz der Schüler\*innen nur dann verbessert, wenn sie freiwillig an Leseangeboten teilnehmen. Dies tun jedoch eher leseaffine Schüler\*innen. Bei Schüler\*innen, die hingegen von ihren Eltern oder Lehrkräften zum Besuch eines Leseangebots aufgefordert werden, verbessert sich die Lesekompetenz nicht, obwohl dies häufig Schüler\*innen sind, bei denen eine Förderung und Steigerung ihrer Lesekompetenzen notwendig wäre. Eine mögliche Erklärung für diese Effekte liefern Rabenstein und Reh (2013): Schüler\*innen erleben durch die Teilnahme an Ganztagsangeboten bestimmte Formen von Zuschreibungen, die sie wiederum auch als leistungsschwache Schüler\*innen konstituiert. Es kann also angenommen werden, dass die nicht freiwillige Teilnahme an Ganztagsangeboten das Selbstkonzept im Lesen der Schüler\*innen negativ beeinflusst. Zugleich ist bisher unklar, ob Förderangebote überhaupt im Sinne von kompensatorischen oder ergänzenden Maßnahmen geeignet sind, Schüler\*innen individuell zu fördern. Das wäre nur dann möglich, wenn die Angebote eine besonders gute Passung zu individuellen Lernvoraussetzungen bieten (Fischer et al., 2014) und die wahrgenommene Unterstützung entsprechend hoch ist.

Diese Ausgangslage führt zu folgenden Fragestellungen:

1. Beeinflusst die Teilnahme an Leseangeboten das Leseselbstkonzept?
2. Wie bewerten Schüler\*innen die individuelle Unterstützung in Leseangeboten im Vergleich zum Unterricht?
3. Fördern Leseangebote bzw. die individuell wahrgenommene Unterstützung in den Leseangeboten die Lesekompetenz und das Leseselbstkonzept?

Basierend auf Daten von 2105 Schüler\*innen der 5. Jahrgangsstufe werden die Fragestellungen beantwortet. Hierfür werden Fragebogendaten von zwei Messzeitpunkten (zu Beginn und am Ende der 5. Klasse) herangezogen. Neben dem Leseselbstkonzept wurde die erlebte Unterstützung durch das Personal in Leseangeboten und im Deutschunterricht erhoben. Die Lesekompetenz wurde über standardisierte Tests erfasst. Den Analysen liegen Pfadmodelle zugrunde. Die Mehrebenenstruktur der Daten wird berücksichtigt.

Erste Ergebnisse zeigen, dass weder das Leseselbstkonzept, noch die Lesegeschwindigkeit und das Leseverstehen durch die Teilnahme an Leseangeboten beeinflusst werden. Der freiwillige Besuch von Ganztagsangeboten hat lediglich einen Effekt auf das Leseverstehen (Fischer et al., 2016), aber nicht, wie erwartet, auf das Leseselbstkonzept. Es zeigt sich außerdem, dass Schüler\*innen in Leseangeboten nicht mehr Unterstützung erfahren als im Unterricht und sich Schüler\*innen die Leseangebote besuchen hinsichtlich der im Unterricht erlebten Unterstützung im Unterricht nicht von den anderen Schüler\*innen unterscheiden.

Die wahrgenommene Unterstützung in Leseangeboten kann mit einer positiveren Entwicklung im Leseselbstkonzept in Verbindung gebracht werden. Dieser Effekt zeigt sich auch für den Unterricht jedoch nur bei den Schüler\*innen, die nicht zusätzlich an Leseangeboten teilnehmen, wie Interaktionsanalysen zeigen. Nur bei diesen steht die wahrgenommene individuelle Unterstützung im Unterricht auch mit einer Verbesserung des Leseverstehens der Schüler\*innen in Zusammenhang.

Die Ergebnisse zeigen, dass Schüler\*innen, die Leseangebote besuchen, hinsichtlich ihrer Leistung weder im Unterricht noch in den Leseangeboten von der wahrgenommenen individuellen Unterstützung profitieren, obwohl sie die gleiche individuelle Unterstützung erfahren wie ihre Mitschüler\*innen. Die individuelle Unterstützung in Leseangeboten wirkt sich bei diesen Schüler\*innen zwar positiv auf ihr Leseselbstkonzept aus, hat aber keinen leistungssteigernden Effekt. Für dieselbe Stichprobe wurde bereits festgestellt, dass Leseangebote in der Regel von leistungsschwächeren Schüler\*innen besucht werden (Fischer et al., 2016). Solche Selektionseffekte können mit einem höheren Unterstützungsbedarf zusammenhängen. Die Ergebnisse zeigen aber, dass es den betreffenden Schüler\*innen nicht unbedingt gelingt, diese Unterstützung auch im Sinne einer Leistungsverbesserung zu nutzen. Dies kann auch erklären, weshalb bisher keine leistungssteigernden Effekte durch den Besuch von Förderangeboten an Ganztagschulen nachgewiesen werden konnten.

## **Wege zur Verbesserung multiprofessioneller Kooperation an Ganztagschulen**

**Stephan Kielblock, Johanna Gaiser**

Universität Gießen

Deutschland ist noch weit davon entfernt, inklusive Bildung, bei der Kinder mit unterschiedlichen Bedürfnissen und Voraussetzungen gemeinsam lernen, vollständig umzusetzen. Dringender Handlungsbedarf im Bereich inklusiver Bildung wurde Deutschland sowohl ‚von extern‘ (UN, 2015) als auch ‚von intern‘ (Klemm, 2015) diagnostiziert. Bundesweit repräsentative Daten von Ganztagschulleitungen belegen das grundsätzliche Potenzial des Ganztagsbetriebs im Hinblick auf die Umsetzung inklusiver Bildung. Dieses ist in der Fachliteratur insbesondere in Bezug auf die multiprofessionelle Kooperation hervorgehoben worden (Vollstädt 2009; Kielblock, Gaiser & Stecher, 2017), die – neben anderen Themen – den Ganztagschuldiskurs mit dominiert. Forschungsergebnisse in diesem Bereich weisen darauf hin, dass strukturelle Hürden sowie das Konkurrenz- und Hierarchiedenken der Beteiligten, das Gelingen der Zusammenarbeit – und damit letztlich die Weichenstellungen zu inklusiveren Praktiken – beeinträchtigen (Breuer & Reh, 2010; Buchna, Coelen, Dollinger, & Rother, 2016; Dizinger, Fussangel, & Böhm-Kasper, 2011; Reh & Breuer, 2012; Speck, Olk, & Stimpel, 2011). Vielfältige empirische Forschungen zu multiprofessioneller Kooperation liegen mittlerweile vor und in jüngeren Studien wird das Potenzial zur Umsetzung von inklusiver Bildung eindringlich hervorgekehrt. Weitgehend unbeantwortet bleibt allerdings die Frage, wie die multiprofessionelle Kooperation an Ganztagschulen in der Praxis zu entwickeln wäre, um professionelles Handeln in multiprofessionellen Teams tatsächlich derart aufeinander abzustimmen, dass inklusive Bildung für alle Kinder ermöglicht wird.

Aus dieser Situation ergibt sich das Forschungsdesiderat, ein Entwicklungsmodell zur Abstimmung professionellen Handelns von Lehrerinnen und Lehrern sowie Personen des weiteren pädagogisch tätigen Personals an Ganztagschulen aus der Praxis

herauszuarbeiten. Ein solches Modell könnte entsprechend in Fortbildungsmaßnahmen und Interventionen in diesem Bereich Anwendung finden, und entsprechend kontinuierlich weiter verfeinert werden.

Um dieses Desiderat zu bearbeiten, wurden Expertinnen und Experten aus unterschiedlichen Bereichen – u. a. Bildungspolitik, Serviceagenturen, Beratungs- und Förderzentren, Schulleitungen – wiederholt zu gemeinsamen Diskussionsrunden eingeladen. Die Treffen wurden ausgiebig protokolliert, und zusätzliche Materialien (z. B. Folien, Flipchart-Notizen, etc.) wurden dokumentiert. Nach jedem Treffen wurden die Protokolle und Materialien analysiert, im Forschungsteam diskutiert und theoretisch angereichert, in verdichteter Form beim nächsten Treffen zurückgespiegelt und weiterbearbeitet. Alle im Verlauf dieses Prozesses entstandenen Materialien wurden mit der Qualitativen Inhaltsanalyse ausgewertet (Gläser & Laudel, 2010). Im Vordergrund der Analysen stand dabei, eine ‚realistische Synthese‘ (Pawson, 2006) zu generieren, die die Komplexität der Praxis abzubilden versucht. Dabei greifen wir auf Methoden der ‚erklärenden Analyse‘ (Yin, 2009) zurück.

Erste analytische Iterationen lassen bereits ein Prozessmodell erkennen, das eine Bandbreite von relevanten Facetten umfasst: Als unumgänglicher Ausgangspunkt für die Entwicklung der multiprofessionellen Kooperation wurde die Bereitschaft zu ebendieser identifiziert. Darauf aufbauend können Kontext und Ressourcen vor Ort gemeinsam geklärt werden, um weitere Entwicklungsmöglichkeiten auszuloten. Die Kommunikation zwischen den Beteiligten kann mit einem Fokus auf Effektivität und einen wertschätzenden Umgang miteinander gestärkt werden. Ebenso gilt es für das Gelingen der Kooperation die Zusammenarbeit in den multiprofessionellen Teams zu stärken. Schließlich muss die Umsetzung in der Praxis stets aufs Neue reflektiert werden, um innerhalb des Prozesses dort ansetzen zu können, wo Bedarf für ein Nachsteuern besteht. Aus der Analyse der Treffen von Praxisexpertinnen und -experten ergeben sich einige interessante Überschneidungen zu bereits bekannten Modellen des Change Managements, sowie zu Modellen der (Ganztags)Schulentwicklung. Diese Anschlussmöglichkeiten, allerdings auch die Unterschiede zu etablierten Modellen, werden in der vorliegenden Studie diskutiert und zur Ausarbeitung des Modells genutzt. Das Ergebnis der Studie eröffnet fundierte Sichtweisen auf die Ganztagschulentwicklung und die Entwicklung inklusiver Bildung. Sich ergebende Möglichkeiten, dieses Modell zu entsprechenden Fortbildungsmaßnahmen und/oder Interventionsstudien weiterzuentwickeln, wird einen Teil der Diskussion darstellen.

## B22 Invited Symposium

Zeit: Donnerstag, 15.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus, Aula 33

### Heterogenität als Herausforderung für pädagogisches Handeln: Ergebnisse des IQB-Bildungstrends 2016

*Chair(s):* **Camilla Rjosk** (Humboldt-Universität zu Berlin, Deutschland), **Nicole Haag** (Humboldt-Universität zu Berlin, Deutschland), **Sebastian Weirich** (Humboldt-Universität zu Berlin, Deutschland), **Stefan Schipolowski** (Humboldt-Universität zu Berlin, Deutschland), **Petra Stanat** (Humboldt-Universität zu Berlin, Deutschland)

*DiskutantIn(nen):* **Eckhard Klieme** (Deutsches Institut für Internationale Pädagogische Forschung)

Das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) führt im Auftrag der Kultusministerkonferenz (KMK) regelmäßig Studien durch, um zu überprüfen, inwieweit in Deutschland zentrale Kompetenzziele erreicht werden, die durch die Bildungsstandards der KMK definiert sind. Im IQB-Bildungstrend 2016 (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017) wurden die Kompetenzen von Viertklässlerinnen und Viertklässlern in den Fächern Deutsch und Mathematik nach dem Jahr 2011 zum zweiten Mal erfasst, so dass es anhand der Daten möglich ist, Entwicklungstrends zu beschreiben. In den Analysen wird neben dem erreichten Kompetenzniveau auch die Heterogenität der Schülerschaft in den Blick genommen, die eine besondere Herausforderung für das professionelle Handeln aller Akteure im Bildungssystem darstellt.

Ziel des Symposiums ist es, zentrale Ergebnisse des IQB-Bildungstrends 2016 und vertiefender Analysen der Daten zu verschiedenen Aspekten von Heterogenität zu berichten, um die bestehenden Herausforderungen genauer zu lokalisieren und mögliche Implikationen für professionelles Handeln abzuleiten. Als Heterogenitätsdimensionen werden die von Schülerinnen und Schülern erreichten Kompetenzen, individuelle Lernvoraussetzungen sowie Rahmenbedingungen schulischen Lernens untersucht. Dabei wird auf Unterschiede eingegangen, die zwischen Ländern, zwischen Schulen und innerhalb von Schulen bestehen. Der Grad und die Verteilung von Heterogenität können zudem einem zeitlichen Wandel unterliegen, der sich anhand von Daten des IQB-Bildungstrends für die Jahre 2011 und 2016 beschreiben lässt.

Im ersten Beitrag wird der IQB-Bildungstrend 2016 als Teil des nationalen Bildungsmonitorings vorgestellt. Es werden die wichtigsten Ergebnisse der Studie zum Erreichen der Bildungsstandards und zu den untersuchten Disparitäten berichtet. Ferner werden Veränderungen in der Heterogenität der Schülerschaft beschrieben, die zwischen den Jahren 2011 und 2016 stattgefunden haben.

Der zweite Beitrag berichtet Analysen zuwanderungsbezogener Disparitäten, wobei der Fokus auf der Schulebene liegt. Insbesondere wird der Frage nachgegangen, inwieweit Unterschiede in der zuwanderungsbezogenen Heterogenität der Schülerschaft an Schulen mit den erreichten Kompetenzen in Zusammenhang stehen und ob Veränderungen in der Zusammensetzung der Schülerschaft, die sich zwischen den Jahren 2011 und 2016 ergeben haben, mit Veränderungen in den erreichten Kompetenzen assoziiert sind.

Im dritten und im vierten Beitrag geht es um spezifische Schülergruppen, die in besonderem Maße zur Heterogenität der Schülerschaft an Schulen beitragen: leistungsstarke Schülerinnen und Schüler und Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF). Im dritten Beitrag werden Analysen zur operationalen Definition von Leistungsstärke vorgestellt. Der vierte Beitrag untersucht, über welche Kompetenzen und motivationalen Merkmale Kinder mit SPF verfügen, die an Förderschulen oder an allgemeinen Schulen unterrichtet werden. Weiterführend wird dabei auch der Fragen nachgegangen, inwieweit die Klassenzusammensetzung mit den Kompetenzen und der Motivation assoziiert ist.

*Beiträge des Symposiums*

### **IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich**

**Petra Stanat, Stefan Schipolowski, Camilla Rjosk, Sebastian Weirich, Nicole Haag**

Humboldt-Universität zu Berlin, Deutschland

Die Bildungsstandards der Kultusministerkonferenz (KMK) für den Primarbereich beschreiben, welche Kompetenzen Schülerinnen und Schüler bis zum Ende der 4. Jahrgangsstufe in den Fächern Deutsch und Mathematik in der Regel entwickelt haben sollen (KMK, 2005a, 2005b). In den Studien des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB), die in regelmäßigen Abständen überprüfen, inwieweit die mit den Bildungsstandards festgelegten Kompetenzziele erreicht werden, wird neben den erreichten Kompetenzen auch untersucht, unter welchen Bedingungen die Schülerinnen und Schüler lernen. Diese Informationen sollen dazu dienen, Stärken und Schwächen zu identifizieren und Prozesse der Weiterentwicklung von Bildungsqualität anzustoßen. Für den Primarbereich fand die erste Studie zur Überprüfung des Erreichens der Bildungsstandards im Jahr 2011 statt (IQB-Ländervergleich 2011; Stanat, Pant, Böhme & Richter, 2012); der IQB-Bildungstrend 2016 hat die Kompetenzen von Viertklässlerinnen und Viertklässlern zum zweiten Mal untersucht.

Im Rahmen des IQB-Bildungstrends 2016 (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017) wurden im Fach Deutsch Kompetenzen in den Bereichen Lesen, Zuhören und Orthografie getestet. Im Fach Mathematik erfasste die Studie Kompetenzen in fünf inhaltlichen Bereichen, die zu einer Globalkala zusammengefasst wurden. Insgesamt nahmen 29.259 Schülerinnen und Schüler der 4. Jahrgangsstufe aus 1.508 Schulen an der Erhebung teil. Neben den Kompetenztests wurden Befragungen der teilnehmenden Schülerinnen und Schüler, Eltern, Lehrkräfte und Schulleitungen durchgeführt.

In diesem Beitrag wird zunächst dargestellt, in welchem Maße sich die von Schülerinnen und Schülern erreichten Kompetenzen zwischen den Jahren 2011 und 2016 verändert haben. Anschließend wird der Frage nachgegangen, inwieweit die beobachteten Trends in den erreichten Kompetenzen mit Veränderungen in der Heterogenität der Schülerschaft einhergehen. Als Heterogenitätsmerkmale werden dabei der sozioökonomische Hintergrund (HISEI bzw. EPG-Klassen), der Zuwanderungshintergrund (Geburtsland der Eltern und des Kindes, Sprachhintergrund) und der sonderpädagogische Förderbedarf der Schülerinnen und Schüler betrachtet. In den Analysen der Verteilung und Veränderung von Heterogenität wird vor allem die Ebenen der Bundesländer in den Blick genommen.

Die Ergebnisse des IQB-Bildungstrends 2016 zeigen, dass im Fach Deutsch im Jahr 2016 bundesweit im Bereich Lesen knapp 66 Prozent, im Bereich Zuhören gut 68 Prozent und im Bereich Orthografie rund 54 Prozent der Viertklässlerinnen und Viertklässler den Regelstandard erreichen oder übertreffen. Im Fach Mathematik erreichen oder übertreffen gut 62 Prozent der Schülerinnen und Schüler in Deutschland den Regelstandard. Die entsprechenden Anteile variieren jedoch erheblich zwischen den Ländern.

Im Fach Deutsch sind die Ergebnisse für Deutschland insgesamt im Bereich Lesen zwischen den Jahren 2011 und 2016 weitgehend stabil geblieben. Hingegen ist der Anteil der Viertklässlerinnen und Viertklässler, die mindestens den

Regelstandard erreichen, im Zuhören um etwa 5 Prozentpunkte, in der Orthografie um etwa 10 Prozentpunkte und im Fach Mathematik um knapp 6 Prozentpunkte gesunken.

Für die untersuchten Heterogenitätsmerkmale ist festzustellen, dass sich der sozioökonomische Hintergrund der Schülerschaft im Mittel kaum verändert hat. Der Anteil der Schülerinnen und Schüler aus zugewanderten Familien ist hingegen um 9 Prozentpunkte auf fast 34 Prozent angestiegen, wobei dies hauptsächlich Kinder der zweiten Generation bzw. Kinder mit einem im Ausland geborenen Elternteil betrifft. Der Anteil der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) hat sich insgesamt wenig verändert, allerdings ist der Anteil inklusiv beschulter Kinder bundesweit um 17 Prozentpunkte auf nunmehr 44 Prozent angestiegen.

Innerhalb der Länder sind die sozialen und zugewanderungsbezogenen Disparitäten weitgehend stabil geblieben. In Übereinstimmung mit früheren Untersuchungen lassen sich zugewanderungsbezogene Disparitäten teilweise auf den geringeren sozialen Status zugewandelter Familien zurückführen. Schülerinnen und Schüler mit SPF, die an allgemeinen Schulen unterrichtet werden, weisen sowohl im Jahr 2011 als auch im Jahr 2016 Kompetenzvorteile gegenüber an Förderschulen unterrichteten Kindern mit SPF auf. Die Vorteile fallen im Jahr 2016 jedoch geringer aus als im Jahr 2011. Zusammengefasst weisen die Befunde auf eine steigende Heterogenität der Schülerschaft in Grundschulen hin, mit der die Lehrkräfte in ihrem Unterricht umgehen müssen.

## **Zuwanderungsbezogene Disparitäten und Effekte der Klassenzusammensetzung im IQB-Bildungstrend 2016**

**Camilla Rjosk, Nicole Haag, Birgit Heppt, Petra Stanat**  
Humboldt-Universität zu Berlin, Deutschland

Nationale und internationale Schulleistungsstudien finden übereinstimmend Kompetenznachteile für Schülerinnen und Schüler mit Zuwanderungshintergrund (z.B. Haag, Böhme, Rjosk & Stanat, 2016; Rauch, Mang, Härtig & Haag, 2016). Die Forschung konnte zudem zeigen, dass sowohl individuelle Lernvoraussetzungen als auch die Zusammensetzung der Klasse bzw. der Schule die von Schülerinnen und Schülern erreichten Kompetenzen beeinflussen. Demnach besteht zwischen dem Anteil benachteiligter Schülerinnen und Schüler in der Klasse und der individuellen Leistungsentwicklung ein negativer Zusammenhang (Dumont, Neumann, Maaz & Trautwein, 2013; Van Ewijk & Sleegers, 2010a, 2010b). Dabei konnten für Deutschland bislang kaum eigenständige Effekte der zugewanderungsbezogenen Zusammensetzung nachgewiesen werden, da sich diese in der Regel auf eine ungünstigere soziale und leistungsbezogene Klassenzusammensetzung zurückführen ließen (z.B. Stanat, Schwippert & Gröhlich, 2010).

Der IQB-Bildungstrend 2016 bietet für die Primarstufe die Möglichkeit, sowohl die zugewanderungsbezogene Zusammensetzung der Schülerschaft als auch zugewanderungsbezogene Disparitäten für die Jahre 2011 und 2016 vergleichend zu analysieren. Das allgemeine Befundmuster der Studie weist darauf hin, dass die Schülerschaft im Jahr 2016 im Durchschnitt in mehreren Bereichen ein geringeres Kompetenzniveau erreicht als im Jahr 2011. In der Öffentlichkeit wurde dieses Befundmuster teilweise auf den gestiegenen Anteil von Kindern mit Zuwanderungshintergrund in den Schulen zurückgeführt. Daran knüpft der vorliegende Beitrag an und untersucht Effekte der zugewanderungsbezogenen Zusammensetzung auf die erreichten Kompetenzen anhand der aktuellen Daten.

Der Beitrag stellt zunächst aktuelle Befunde des IQB-Bildungstrends 2016 zu Veränderungen der zugewanderungsbezogenen Zusammensetzung der Schülerschaft und zugewanderungsbezogenen Disparitäten zwischen den Jahren 2011 und 2016 zusammenfassend dar. Im Anschluss gehen wir in vertiefenden Analysen folgenden Fragestellungen nach: (1) Welche Zusammenhänge bestehen zwischen der zugewanderungsbezogenen Klassenzusammensetzung und den von Viertklässlerinnen und Viertklässlern erreichten Kompetenzen im Fach Deutsch (Lesen, Zuhören, Orthografie) und im Fach Mathematik im Jahr 2016? (2) Unterscheiden sich diese Zusammenhänge zwischen Kindern mit Zuwanderungshintergrund und Kindern ohne Zuwanderungshintergrund? Zudem sind vergleichende Analysen für Schulen vorgesehen, die sowohl im Jahr 2011 als auch 2016 an der Studie teilgenommen haben, um zu untersuchen, inwieweit auf Schulebene Veränderungen in der Zusammensetzung der Schülerschaft mit Veränderungen in den erreichten Kompetenzen einhergehen.

Die vertiefenden Analysen beziehen sich auf Daten des IQB-Bildungstrends 2016 (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017) mit rund 24.000 Viertklässlerinnen und Viertklässlern aus ca. 1.300 Klassen. Der Zuwanderungshintergrund der Kinder wird anhand des Geburtslands ihrer Eltern operationalisiert (mindestens ein Elternteil nicht in Deutschland geboren). Zur Vorhersage der Kompetenzen im Lesen, im Zuhören und in der Orthografie sowie der Globalskala mathematischer Kompetenz werden Mehrebenenanalysen (Random-Intercept- und Random-Slope-Modelle) durchgeführt. Zentrale unabhängige Variable ist der Anteil von Kindern mit Zuwanderungshintergrund an der Schule. Als Kontrollvariablen auf Individualebene werden der Zuwanderungshintergrund, der sozioökonomische Status, die kognitiven Grundfähigkeiten und das Geschlecht der Kinder in die Analysen einbezogen. Auf Klassenebene werden die Mittelwerte und Streuungen des sozioökonomischen Status und der kognitiven Grundfähigkeiten berücksichtigt.

Die Ergebnisse des IQB-Bildungstrends 2016 zeigen, dass der Anteil von Kindern mit Zuwanderungshintergrund gegenüber dem Jahr 2011 um rund neun Prozentpunkte zugenommen hat und dass zwischen den Jahren 2011 und 2016 die Anzahl der Schulen mit einem vergleichsweise hohen Anteil von Kindern mit Zuwanderungshintergrund angestiegen ist. Zuwanderungsbezogene Kompetenzunterschiede finden sich in allen vier untersuchten Bereichen und diese sind im Untersuchungszeitraum der fünf Jahre weitgehend stabil geblieben. In Übereinstimmung mit früheren Befunden weisen erste vertiefende Analysen darauf hin, dass der Anteil von Kindern mit Zuwanderungshintergrund negativ mit dem erreichten Kompetenzniveau assoziiert ist, dass aber nach Kontrolle weiterer Merkmale der Klassenzusammensetzung diese Zusammenhänge – mit Ausnahme der Effekte im Fach Mathematik – nicht mehr statistisch signifikant sind. Differenzielle Effekte für Kinder aus zugewanderten Familien im Vergleich zu Kindern ohne Zuwanderungshintergrund zeigen sich nicht. Mögliche Implikationen der Ergebnisse für den Umgang mit heterogen zusammengesetzten Klassen werden diskutiert.

## **Konzeptionelle Herausforderungen bei der Untersuchung leistungsstarker Schülerinnen und Schüler**

**Claudia Neuendorf<sup>1</sup>, Poldi Kuhl<sup>2</sup>, Malte Jansen<sup>1</sup>**

<sup>1</sup>Humboldt-Universität zu Berlin, Deutschland, <sup>2</sup>Leuphana Universität Lüneburg

In jüngster Zeit rückt die Gruppe leistungsstarker Schülerinnen und Schüler zunehmend in die Aufmerksamkeit der Bildungspolitik und Öffentlichkeit. So verabschiedete die Kultusministerkonferenz (KMK) im Jahr 2015 eine Strategie zur Förderung leistungsstarker Schülerinnen und Schüler (KMK, 2015), in deren Folge eine gemeinsame Initiative von Bund und Ländern mit der gleichen Zielsetzung beschlossen wurde (KMK, 2016). Das zunehmende Interesse spiegelt sich auch in den Ergebnisberichten großer Schulleistungsstudien wider (Köller & Baumert, 2017; Neuendorf, Kuhl & Jansen, 2017; Reiss &

Sälzer, 2016). Die Forschung zu leistungsstarken Schülerinnen und Schülern steht aber derzeit noch vor konzeptionellen Herausforderungen, von denen im vorliegenden Beitrag zwei näher betrachtet werden sollen.

Zum einen fehlt bislang eine einheitliche Definition schulischer Leistungsstärke (Köller & Baumert, 2017). In früheren Forschungsarbeiten wurden zur Operationalisierung sowohl Leistungs- und Kompetenztests als auch Noten herangezogen. Zum Teil wurden fachspezifische, zum Teil aber auch fächerübergreifende Kriterien genutzt. Des Weiteren lässt sich Leistungsstärke auf der Grundlage verschiedener Bezugsnormen (i.d.R. kriterial oder sozial) klassifizieren. Es ist davon auszugehen, dass unterschiedliche Herangehensweisen auch zu systematischen Unterschieden bei der Identifikation leistungsstarker Schülerinnen und Schüler führen.

Einhergehend mit dem Desiderat einer eindeutigen Definition stellt sich zudem die Herausforderung, Leistungsstärke klar von anderen Merkmalen schulischer und kognitiver Exzellenz abzugrenzen. In der Forschung zu Leistungsstärke wird häufig Bezug auf die (Hoch-)Begabungsforschung genommen, die innerhalb der Psychologie bereits eine lange Tradition hat (z.B. Seashore, 1922; Terman, 1925). In einigen Studien wurden leistungsstarke und hochbegabte Schülerinnen und Schüler vergleichend gegenübergestellt. Dabei zeigte sich, dass Leistungsstarke, Kinder mit hohen kognitiven Grundfähigkeiten und Schülerinnen und Schüler mit (Hoch-)Begabungsdiagnose unterschiedliche, aber sich auch teilweise überschneidende Teilpopulationen darstellen (z.B. Rost, 2009; Vock, Köller & Nagy, 2013; Wild, 1991).

Auf Grundlage der oben umrissenen Herausforderungen behandelt der Beitrag folgende Fragestellungen:

1. Wie unterscheiden sich verschiedene Operationalisierungen von Leistungsstärke im Hinblick auf den identifizierten Anteil Leistungsstarker und ihrer Zusammensetzung?
2. Welche Zusammenhänge bestehen zwischen Leistungsstärke, kognitiven Grundfähigkeiten und diagnostizierter Hochbegabung?

Grundlage der Analysen bilden Daten des IQB-Bildungstrends 2016 (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017). In der Studie wurde eine kriteriale Operationalisierung von Leistungsstärke auf Basis der auf den Bildungsstandards der KMK basierenden Kompetenztests gewählt (Erreichen des Optimalstandards) und der Überschneidungsbereich von Leistungsstärke in Deutsch und Mathematik untersucht (Neuendorf et al., 2017). Dieser Beitrag wendet im Rahmen weiterführender Analysen zusätzlich verschiedene fächerübergreifende Operationalisierungen an. Exemplarisch wird zum einen das Leistungskriterium (Kompetenzwerte vs. Noten) und zum anderen die Bezugsnorm (kriterial, d.h. Kompetenzstufe bzw. Note, vs. sozial, also Rangreihe innerhalb der Schulklasse) variiert. Die vier resultierenden Gruppen werden hinsichtlich ihrer Größe, ihres Überschneidungsbereichs und ihrer Zusammensetzung verglichen. Mit Blick auf die zweite Fragestellung werden Zusammenhänge zwischen den Merkmalen Leistungsstärke, kognitive Grundfähigkeiten sowie diagnostizierte Hochbegabung untersucht.

Bei einer kriterialen Operationalisierung auf Basis der Bildungsstandards zeigt sich zunächst, dass bei fast 30 Prozent der leistungsstarken Kinder sowohl in Deutsch als auch in Mathematik das Kriterium der Leistungsstärke zutrifft. Vergleicht man verschiedene Operationalisierungen von Leistungsstärke wird deutlich, dass die Operationalisierung von Leistungsstärke die Anteile leistungsstarker Schülerinnen und Schüler beeinflusst. Am größten ist die Gruppe der Leistungsstarken, wenn in jeder Klasse die Notenbesten als leistungsstark klassifiziert werden, am geringsten ist die Gruppengröße, wenn ausschließlich Kinder mit der Bestnote in beiden Fächern als leistungsstark gewertet werden. Bei Nutzung der Kompetenzwerte liegen die Anteile Leistungsstarker dazwischen. Auch Unterschiede in der Ausprägung zentraler Hintergrundmerkmale der Gruppen leistungsstarker Schülerinnen und Schüler werden im Beitrag dargestellt.

Zur zweiten Fragestellung wird unter anderem deutlich, dass mit 0,4 Prozent lediglich ein kleiner Anteil der Kinder eine diagnostizierte Hochbegabung aufweist, die der Schule bekannt ist. Die Ergebnisse werden im Hinblick auf ihre Relevanz für zukünftige Forschung zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler diskutiert.

## **Schulische Kompetenzen und schulische Motivation von Kindern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen**

**Aleksander Kocaj<sup>1</sup>, Poldi Kuhl<sup>2</sup>, Nicole Haag<sup>1</sup>, Pauline Kohrt<sup>1</sup>**

<sup>1</sup>Humboldt-Universität zu Berlin, Deutschland, <sup>2</sup>Leuphana Universität Lüneburg

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) werden zunehmend in allgemeinen Schulen anstatt in Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016). Dies wirft die Frage auf, inwieweit die Schullart mit den schulischen Kompetenzen und der schulischen Motivation von Kindern und Jugendlichen mit SPF zusammenhängt. Für eine Beschulung in Förderschulen spricht, dass die Unterrichtsgestaltung in Förderschulen stärker an die Lernvoraussetzungen von Kindern und Jugendlichen mit SPF angepasst sei (Ahrbeck, 2014). Für eine gemeinsame Beschulung spricht, dass Kinder und Jugendliche mit SPF in allgemeinen Schulen eine anregendere Lernumgebung mit leistungsstärkeren Mitschülerinnen und Mitschülern als in Förderschulen vorfinden (Peetsma et al., 2001; Wocken, 2005). Der Forschungsstand verweist auf gegenläufige Zusammenhänge zwischen der Schullart und den schulischen Kompetenzen bzw. der schulischen Motivation. Während Schülerinnen und Schüler mit SPF in allgemeinen Schulen höhere schulische Kompetenzen erzielen als in allgemeinen Schulen (Kocaj et al., 2014; Ruijs & Peetsma, 2009; siehe aber Lehmann & Hoffmann, 2009; Stranghöner et al., 2017), verweist die Mehrzahl der Studien auf eine höhere schulische Motivation zugunsten von Schülerinnen und Schülern in Förderschulen beziehungsweise in speziellen Förderklassen (Bear et al., 2002; siehe aber Elbaum, 2002; Peetsma et al., 2001; Wild et al., 2015).

In diesem Beitrag vergleichen wir die schulischen Kompetenzen und Merkmale der schulischen Motivation von Kindern mit SPF in Förderschulen und allgemeinen Schulen. Dabei untersuchen wir, ob sich die Zusammenhänge zwischen der Schullart und den schulischen Outcomes in Abhängigkeit vom Förderschwerpunkt unterscheiden. Der Beitrag knüpft an Analysen zum IQB-Ländervergleich 2011 an und erweitert diese um Ergebnisse zum Kompetenzbereich Orthografie und um differenzierte Ergebnisse für Schülerinnen und Schüler mit dem Förderschwerpunkt emotionale und soziale Entwicklung.

Die Analysen basieren auf einer Teilstichprobe des IQB-Bildungstrends 2016 von 2 456 Kindern mit einem SPF in den Bereichen Lernen, Sprache oder emotionale und soziale Entwicklung, von denen 1 196 Kinder eine Förderschule und 1 260 Kinder eine allgemeine Schule besuchten. Kinder mit mehreren Förderschwerpunkten wurden nicht in den Analysen berücksichtigt, um nach Förderschwerpunkt differenzierte Aussagen zu Schullartunterschieden treffen zu können. Zur Erfassung der im IQB-Bildungstrend 2016 untersuchten Kompetenzen in den Fächern Deutsch (Kompetenzbereiche Lesen, Zuhören und Orthografie) und Mathematik kamen bei Schülerinnen und Schülern mit SPF in Förderschulen und in allgemeinen Schulen angepasste Testhefte zum Einsatz, die leichtere Aufgaben beinhalteten. Als Merkmale der schulischen Motivation



wurden das akademische Selbstkonzept sowie das Interesse in den Fächern Deutsch und Mathematik erfasst. Bei der Auswertung kamen Propensity-Score-Matching-Verfahren (Full Matching) zum Einsatz, um potenzielle Unterschiede in Lernvoraussetzungen und weiteren Hintergrundmerkmalen zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen statistisch zu kontrollieren.

Kinder mit SPF in allgemeinen Schulen wiesen höhere schulische Kompetenzen auf als vergleichbare Kinder mit SPF in Förderschulen. Diese Schularnterschiede waren für Kinder mit dem Förderschwerpunkt Lernen stärker ausgeprägt als für Kinder mit dem Förderschwerpunkt Sprache. Bei Kindern mit dem Förderschwerpunkt emotionale und soziale Entwicklung zeigten sich nach dem Propensity-Score-Matching hingegen für keinen der getesteten Kompetenzbereiche statistisch signifikante Unterschiede zwischen Förderschulen und allgemeinen Schulen. Im Gegensatz zu den schulischen Kompetenzen berichteten Kinder mit SPF in Förderschulen insgesamt eine höhere schulische Motivation als vergleichbare Kinder mit SPF in allgemeinen Schulen. Diese motivationalen Unterschiede zugunsten von Förderschulen zeigten sich für alle untersuchten Förderschwerpunkte, waren jedoch für Kinder mit dem Förderschwerpunkt Lernen und emotionale und soziale Entwicklung tendenziell stärker ausgeprägt als für Kinder mit dem Förderschwerpunkt Sprache. In weiterführenden Analysen wird anhand von Mehrebenenanalysen geprüft, inwieweit die fähigkeitsbezogene und soziale Klassenkomposition in Förderschulen und allgemeinen Schulen mit den schulischen Kompetenzen und der schulischen Motivation der Kinder mit SPF zusammenhängt. Außerdem werden die Ergebnisse im Vergleich zu vorangegangenen Analysen anhand der Daten des IQB-Ländervergleichs 2011 diskutiert.

# Postersession

Zeit: Donnerstag, 15.02.2018: 17:15 - 18:45 · Ort: Kollegienhaus 1. Stock

## Evidenzbasierte Praxis in der Lehrerbildung: Die Perspektive der Lehrbildner/innen

**Annika Diery, Freydis Vogel, Maria Bannert, Eva-Maria Lankes, Tina Seidel**

Technische Universität München, Deutschland; [annika.diery@tum.de](mailto:annika.diery@tum.de)

### Theoretischer Hintergrund

Analog zur Ausbildung und praktischen Arbeit von Mediziner\*innen (vgl. Coe, Fitz-Gibbon, & Tymms, 2000) wird evidenzbasierte Praxis für das professionelle Handeln von Lehrkräften gefordert (Bromme, Prenzel & Jäger, 2014; KMK, 2004, 2014). Dies bedeutet generell, einschlägige wissenschaftliche Forschungsbefunde und Theorien beim professionellen Handeln im Schulkontext situations-angemessen zu berücksichtigen (Bauer, Prenzel & Renkl, 2015). So können aktuelle Forschungsbefunde beispielsweise als Grundlage dienen, um Handlungsentscheidungen zu treffen, die eigene Praxis rational zu begründen oder zu reflektieren (Seidel et al., 2017).

Jedoch scheinen Lehrkräfte nicht ausreichend für eine evidenzbasierte Praxis ausgebildet zu sein. Nur wenige nutzen fachwissenschaftliche Ressourcen für die Planung und Gestaltung ihres Unterrichts und sie stehen der Brauchbarkeit und Anwendung bildungswissenschaftlicher Forschung eher skeptisch gegenüber (Hetmanek et al., 2015).

Lehrbildner/innen stellen in ihrer Arbeit die Weichen, um Lehrkräften einen Zugang und Bezug zu evidenzbasierter Praxis zu ermöglichen. Daher fokussiert der vorliegende Beitrag Lehrbildner/innen und deren Perspektive auf evidenzbasierte Praxis.

### Fragestellungen

Wofür nutzen Lehrbildner/innen aktuelle Forschungsbefunde im Rahmen der Aus- und Weiterbildung von Lehrkräften?

Welche Herausforderungen und Schwierigkeiten sehen Lehrbildner/innen, wenn es um die Einbeziehung aktueller Forschungsbefunde in die Aus- und Weiterbildung von Lehrkräften geht?

### Methode

Zur Untersuchung der Fragestellungen wurden 200 Lehrbildner/innen zur Teilnahme an einer online-Fragebogenstudie eingeladen. Von den angeschriebenen Lehrbildner/innen füllten  $n = 74$  den kompletten Fragebogen aus (Alter:  $M = 45.84$ ,  $SD = 11.32$ ;  $n = 41$  Frauen;  $n = 33$  Männer;  $n = 52$  an einer Universität und  $n = 21$  an außeruniversitären Institutionen). Zur Erfassung der Nutzung aktueller Forschungsbefunde, wurde gefragt, ob und wofür diese eingesetzt werden. Die wahrgenommene Schwierigkeit wurde mit einer Skala aus fünf Items gemessen (z.B. „Ich fühle mich damit überfordert, aktuelle Forschungsbefunde in meine Lehrveranstaltungen miteinzubeziehen.“) ( $\alpha = 0,79$ ). Potentielle Schwierigkeiten bezüglich der Einbeziehung aktueller Forschungsbefunde wurden in einem geschlossenen Antwortformat abgefragt. Alle Items waren auf einer sechsstufigen Likert-Skala (1: „ich stimme überhaupt nicht zu“ bis 6: „ich stimme voll und ganz zu“) zu beantworten.

### Ergebnisse

Die Befragten gaben an, aktuelle Forschungsbefunde vor allem zu nutzen, um ihr eigenes Wissen zu erweitern ( $M = 5,59$ ,  $SD = 0,66$ ), um wissenschaftliche Belege für das Thema Unterricht vermitteln zu können ( $M = 5,07$ ,  $SD = 0,99$ ), und um Lehrkräfte auf den aktuellen Forschungsstand zu bringen ( $M = 4,89$ ,  $SD = 1,02$ ).

Die Frage, ob die Einbeziehung aktueller Forschungsbefunde in Lehrveranstaltungen eine Herausforderung darstellt, beantworteten  $n = 33$  Lehrbildner/innen mit „ich stimme eher zu“ bis „ich stimme voll und ganz zu“. Die Befragten fanden es schwierig, die nötige Zeit aufzubringen ( $M = 4,03$ ,  $SD = 1,46$ ) und die Fülle der Publikationen zu überblicken ( $M = 4,06$ ,  $SD = 1,43$ ). Explorative Analysen ergaben, dass Lehrbildner/innen an außeruniversitären Einrichtungen im Gegensatz zu Lehrbildner/innen an Universitäten es schwieriger einschätzen, die englische Sprache ( $M = 3,95$ ,  $SD = 1,54$ ,  $M = 1,82$ ,  $SD = 1,16$ ,  $t(68) = 6,20$ ,  $p = 0,031$ ) und die fachwissenschaftliche Sprache ( $M = 3,37$ ,  $SD = 1,34$ ,  $M = 1,56$ ,  $SD = 0,82$ ,  $t(69) = 6,84$ ,  $p = 0,005$ ) zu verstehen sowie immer auf dem aktuellen Stand der Forschung zu sein ( $M = 4,32$ ,  $SD = 0,94$ ,  $M = 3,75$ ,  $SD = 1,46$ ,  $t(68) = 1,57$ ,  $p = 0,032$ ).

Obwohl Lehrkräfte ihr professionelles Handeln auf empirischer Evidenz aufbauen sollen (Bromme, Prenzel & Jäger, 2014), haben Lehrbildner/innen offensichtlich Schwierigkeiten, evidenzbasierte Praxis an (zukünftige) Lehrkräfte zu vermitteln. Es besteht also ein Bedarf, hierfür Unterstützungsmaßnahmen zu entwickeln. Bei diesen sollte auf die spezifischen Schwierigkeiten eingegangen werden, vor allem hinsichtlich der unterschiedlichen Zielgruppen von Lehrbildner/innen. Im Hinblick auf die berichteten Schwierigkeiten (Zeitaufwand, Publikationsfülle, Verstehen englischer und fachwissenschaftlicher Sprache) stellt die strukturierte und zielgruppengerecht aufbereitete Darstellung von Forschungsbefunden in deutscher Sprache (z.B. „Clearing House Unterricht“, Seidel et al., 2017) einen vielversprechenden Ansatz dar. Die vorliegende Studie wird aktuell mit einer erweiterten Stichprobe durchgeführt, die auch zu den Angeboten des „Clearing House Unterricht“ befragt wird. Ergebnisse der Analyse dieser Daten sollen auf der Konferenz vorgestellt und diskutiert werden.

## Epistemic (conceptual) change bei Lehramtsstudierenden – Eine exemplarische Anwendung longitudinaler Messinvarianzanalysen zur Beschreibung von Veränderungsprozessen epistemologischer Überzeugungen

**Eric Klopp, Robin Stark**

Universität des Saarlandes, Deutschland; [e.klopp@mx.uni-saarland.de](mailto:e.klopp@mx.uni-saarland.de)

Epistemologische Überzeugungen spielen im Lehramtsstudium eine zentrale Rolle (Lunn-Brownlee, 2017), da diese eine wichtige Voraussetzung für die geforderte Kompetenz zur evidenzbasierten Reflexion von Lehr- und Lernprozessen sind (KMK, 2004). Daher ist es nicht nur von Interesse, Möglichkeiten zur Förderung epistemologischer Überzeugungen zu entwickeln, sondern auch deren Veränderung im Lehramtsstudium zu kennen. Ein Problem hierbei ist die duale Natur epistemologischer Überzeugungen, da diese sowohl domänenübergreifend als auch domänenspezifisch sind (Muis et al., 2006). Forschungsdesiderata sind daher eine methodisch angemessene Untersuchung des Entwicklungsprozesses sowie die Betrachtung der dualen Natur epistemologischer Überzeugungen.

Da epistemologische Überzeugungen i.d.R. mit Fragebogen erfasst werden, stellt sich die Frage, wie Veränderungsprozesse durch geeignete Modelle mit latenten Variablen abgebildet werden können. Veränderungsprozesse latenter Variablen lassen sich durch sogenannten  $\alpha$ -,  $\beta$ - und  $\gamma$ -Change beschreiben (Chan, 1998).  $\alpha$ -Change bezieht sich auf die Veränderungen der wahren Werte eines Konstrukts, sofern das Konstrukt selbst als auch die Messeigenschaften der Indikatoren über die Zeit

konstant bleiben.  $\beta$ -Change liegt vor, wenn das Konstrukt über die Zeit konstant bleibt, sich aber die Messeigenschaften der Indikatoren verändern.  $\gamma$ -Change hingegen liegt dann vor, wenn sich das Konstrukt über die Zeit verändert. Nach dieser Definition lassen sich  $\alpha$ -,  $\beta$ - und  $\gamma$ -Change durch das Vorliegen bestimmter Ebenen von Messinvarianz charakterisieren.  $\alpha$ -Change setzt longitudinale, skalare Messinvarianz voraus, wohingegen  $\beta$ -Change lediglich die Geltung konfigurationaler bzw. metrischer Messinvarianz voraussetzt.  $\gamma$ -Change dagegen beschreibt das Fehlen konfigurationaler Invarianz; bei  $\gamma$ -Change kann im Zusammenhang mit epistemologischen Überzeugungen auch von epistemic conceptual change (Sinatra & Chinn, 2011) gesprochen werden.

Die Studie stellt eine exemplarische Anwendung der beschriebenen Methoden zur Beschreibung der Entwicklung domänenübergreifender und domänenspezifischer epistemologischer Überzeugungen von Lehramtsstudierenden im ersten Semester dar. Insgesamt nahmen  $n=204$  Lehramtsstudierende (143 weiblich, MAlter=21.95 [sd=5.63]) teil, die Erhebung erfolgte zu drei Messzeitpunkten äquidistant verteilt über das erste Semester. Untersucht wurde die longitudinale Messinvarianz epistemologischer Überzeugungen, beim Vorliegen skalarer Invarianz wurde zur Beschreibung der Veränderung ein latentes Changescorere-Modell (LCS; Newsom, 2015) berechnet; Reliabilitäten wurden mittels des  $\omega$ -Koeffizienten (Raykov, 2001) bestimmt. Domänenübergreifende epistemologische Überzeugungen wurden mittels der adaptierten Skalen Sicherheit des Wissens (SW,  $\omega_{T1/T2/T2}=.66/.69/.73$ ), Umgang mit Autoritäten (UA,  $\omega_{T1/T2/T2}=.51/.53/.51$ ) sowie Reflexive Natur des Wissens (RN,  $\omega_{T1/T2/T2}=.54/.73/.68$ ) erfasst (Moschner & Gruber, 2017). Auf die Bildungswissenschaft bezogene domänenspezifischen epistemologische Überzeugungen wurden mit den adaptierten Skalen Textur des Wissens (TEX,  $\omega_{T1/T2/T2}=.71/.77/.77$ ) und Variabilität des Wissens (VAR,  $\omega_{T1/T2/T2}=.59/.67/.75$ ) erfasst (Stahl & Bromme, 2008).

Ein  $\alpha$ -Change lag bei Umgang mit Autoritäten und Textur des Wissens vor. Das skalare Invarianzmodell von UA wies einen ausgezeichneten Modellfit auf ( $\chi^2(23)=30.94$ ,  $p=.12$ , CFI=.986, RMSEA=0.041). Das Niveau von UA blieb über die Zeit konstant, die latenten Differenzwerte des LCS sind nicht signifikant ( $D12=-0.02$ ,  $p=.652$ ,  $D23=0.03$ ,  $p=.350$ ). Das skalare Invarianzmodell von TEX wies einen ausgezeichneten Modellfit auf ( $\chi^2(51)=54.50$ ,  $p=.343$ , CFI=.996, RMSEA=0.018). Das Niveau von TEX blieb von ersten zum zweiten Messzeitpunkt konstant ( $D12=0.06$ ,  $p=.278$ ), nahm aber vom zweiten zum dritten Messzeitpunkt signifikant zu ( $D23=0.147$ ,  $p=.006$ ).

$\beta$ -Change lag bei Sicherheit des Wissens, Reflexive Natur des Wissens und Variabilität des Wissens vor. Hier zeigte sich lediglich die Passung eines metrischen Invarianzmodells (SW:  $\chi^2(45)=54.42$ ,  $p=.108$ , CFI=.985, RMSEA=0.041; VAR:  $\chi^2(45)=47.50$ ,  $p=.371$ , CFI=.996, RMSEA=0.044) bzw. eines konfiguralen Invarianzmodells (RN:  $\chi^2(36)=50.63$ ,  $p=.052$ , CFI=.974, RMSEA=0.047).

Die Ergebnisse zeigen, dass bei der Erfassung von Veränderungsprozessen epistemologischer Überzeugungen die Beachtung von Messinvarianz erforderlich ist, um diese adäquat zu interpretieren; insbesondere wenn weitergehende Analysen mit latenten Changescorere-Modellen erfolgen sollen. In dieser Studie kann nur bei Umgang mit Autoritäten und Textur des Wissens im Sinne Niveauänderung ausgegangen werden, bei den anderen Dimensionen blieb das Konstrukt zwar konstant, es fand allerdings eine „Rekalibrierung“ der Indikatoren statt. Veränderungen im Sinne eines epistemic conceptual change konnten nicht beobachtet werden.

### **Fourth-graders' competence beliefs in mathematics**

**Denise Villányi, Rachel Wollschläger, Romain Martin, Antoine Fischbach**

University of Luxembourg, Luxembourg; [denise.villanyi@uni.lu](mailto:denise.villanyi@uni.lu)

#### Introduction

Competence beliefs and achievement are positively and reciprocally related (see e.g., Marsh & Martin, 2011; Huang, 2011) from elementary school on (Guay, Marsh, & Boivin, 2003). Although teachers should foster the development of positive competence beliefs, they often lack diagnostic competence and concrete tools to do so (Artelt & Gräsel, 2009; Schrader, 2009; Baumert & Kunter, 2006). Knowing this, the importance to investigate the structure of elementary school students' competence beliefs becomes evident, particularly as there are very few studies in that domain with young learners (Pinxten Wouters, Preckel, Niepel, De Fraine, & Verschuere, 2015). Developing and testing tools for the classroom that support teachers' diagnostic activities can take research and educational practice another step towards the above-mentioned objectives.

We investigated the structure of fourth-graders' competence beliefs in mathematics by closely drawing on Scherer's (2013) research related to secondary school students' competence beliefs in chemistry. We concretely studied:

- The internal structure of competence beliefs by hypothesizing a multidimensional structure including general academic self-concept (ACSC), mathematics self-concept (MSC), and mathematics self-efficacy (MSE).
- Domain specificity of competence beliefs in fourth-graders.
- The relationships of ACSC, MSC and MSE with following covariates: mathematics, reading, and listening performance; mathematics and general academic interest; mathematics and general academic anxiety.

#### Methodology

We assessed students' self-efficacy applying our newly developed mathematics self-efficacy scale (Villányi, Martin, Sonnleitner, Siry, & Fischbach, under review). The scale consists of 67 items to validly represent the primary education curriculum ( $\alpha = .96$ ; corrected item-to-total-correlations  $rit = .34 - .72$  with 6 curriculum relevant items showing lower values  $rit = .20 - .30$ ). The covariates of competence beliefs mentioned under point (c) stem from the national school monitoring program (Martin, Ugen, & Fischbach, 2015). Our sample consists of 392 fourth-graders out of 28 school classes randomly chosen all over Luxembourg's public elementary schools. We applied methods of structural equation modeling (SEM) to study the aspects mentioned above.

#### Results

- SEM confirmed, with acceptable to good model fit (CFI = .93; RMSEA = .07, SRMR = .05), the tridimensional structure of fourth-graders competence beliefs. One- and two-dimensional models showed poor model fit.
- When controlling for the general factor of competence beliefs (SEM, nested model), we found a significant correlation of moderate effect size between ACSC and MSC.
- Latent regression modelling showed that MSE and MSC were predicted by domain related covariates (performance, interest) and gender; ACSC was predicted by general interest. All models presented substantial variance explanation ( $R^2 = .35 - .60$ ).

#### Discussion and conclusion

a. Based on our results and in line with previous research confirming a common core for self-concept and self-efficacy (Bong & Skaalvik, 2003; Ferla, Valcke, & Cai, 2009; Scherer, 2013), we favoured the model with ACSC, MSC, and MSE being components of a second-order general factor of competence beliefs over a three-factor model assuming solely correlation between these constructs.

b. The moderate relationship between ACSC and MSC indicates less developed domain specificity in fourth-graders competence beliefs compared to secondary school students (see Scherer, 2013) which is in line with findings related to age differences in self-concept (Möller, Pohlmann, Köller, & Marsh, 2009; Pinxten et al., 2015; Marsh, 1989; Marsh, 1990).

c. Results revealed different predictors across constructs but similar ones within a domain.

In our study, we tested Scherer's (2013) hypotheses with fourth-graders in the domain of mathematics and reproduced his findings except for domain specificity of competence beliefs. We (re-)confirmed the acceptable to good test characteristics of our newly developed curriculum based mathematics self-efficacy scale and gave evidence on its construct validity.

## **Lehr- und Lernvorstellungen von Lehramtsstudierenden vor Beginn ihres Studiums – Eine Unterscheidung der Vorstellungen zum Lernen der Schüler/innen und zum eigenen Lernen**

**Fabian Schumacher, Melanie Basten, Matthias Wilde**

Universität Bielefeld, Deutschland; [fschumacher@uni-bielefeld.de](mailto:fschumacher@uni-bielefeld.de)

Der Analyse der Lehr- und Lernvorstellungen (LLV) von angehenden Lehrenden vor und im Verlauf des Studiums wird zur Optimierung der Lehramtsausbildung eine große Bedeutung zugeschrieben (Brown et al., 2013).

LLV werden nach Kleickmann et al. (2010) als die subjektiven Sichtweisen über das Lernen im Fach verstanden und umfassen epistemologisch validiertes Wissen sowie subjektive Überzeugungen. LLV sind bereits zu Beginn des Studiums recht persistent (Brauer et al., 2014; Pajares, 1992; Woolfolk Hoy et al., 2006) und oftmals traditionell bzw. transmissiv geprägt (Da-Silva et al., 2007; Koballa et al., 2000).

Der Fokus der meisten Untersuchungen liegt auf überfachlichen transmissiven und konstruktivistischen LLV (Buelens et al., 2002; Dubberke et al., 2008; Kleickmann et al., 2010). Diese LLV sind nicht gegensätzlich, sondern können parallel vorliegen (Brauer et al. 2014). In dieser Studie werden ergänzend zwei fachspezifische konstruktivistische LLV zu Vernetzung von Inhalten und zu Präkonzepten erfasst (vgl. Brauer et al. 2015). Die überfachlichen LLV werden größtenteils durch die schulischen Erfahrungen beeinflusst (Shulman, 1987); die fachspezifischen LLV mehrheitlich durch die universitäre Ausbildung (Cain, 2012).

Es konnte gezeigt werden, dass sich die LLV von Lehramtsstudierenden danach unterscheiden lassen, ob sie sich auf das Lernen ihrer zukünftigen Schüler/innen oder auf ihr eigenes Lernen beziehen (Brauer et al., 2015). Damit können zwei Perspektiven auf LLV unterschieden werden.

In Deutschland existieren nur wenige empirische Daten über die Ausprägung der LLV angehender Biologielehramtsstudierender. Markic et al. (2006) berichten von eher konstruktivistisch geprägten LLV. Internationale Erhebungen von Da-Silva et al. (2007) und Koballa et al. (2000) weisen eher auf transmissiv geprägte LLV hin.

Daraus ergibt sich Fragestellung 1: Welche der generalisierten und fachdidaktischen LLV dominieren bei den angehenden Biologielehramtsstudierenden?

Brauer et al. (2017) fanden Unterschiede in den LLV für die beiden Perspektiven zum eigenen Lernen und zum Lernen von Schüler/innen.

Daher lautet Fragestellung 2: Unterscheiden sich die LLV der Biologielehramtsstudierenden in der Einführungsveranstaltung zum eigenen Lernen und zum Lernen der Schüler/innen?

Für die vorliegende Studie wurden 70 Studierende der Biologie (74% weiblich) in der Einführungswoche (WiSe 2016/17), d.h. vor dem Beginn ihrer ersten Vorlesungen, befragt. Ihr Durchschnittsalter betrug 21,03 (SD=4,0) Jahre. Der Fragebogen zu Lehr-/Lernvorstellungen von Brauer et al. (2015) besteht aus identischen Skalen in zwei Perspektiven, den Vorstellungen zum Lernen der Schüler/innen und denen zum eigenen Lernen. Die Skalen umfassen jeweils zwei allgemeine Skalen zu konstruktivistischen (Kon) und transmissiven (Trans) Vorstellungen sowie zwei fachspezifische Skalen zur Vernetzung (Vern) und Präkonzepten (Präk). Das Antwortformat war eine siebenstufige Likert-Skala von „trifft gar nicht zu“=0 bis „trifft voll zu“=6.

Die Studierenden in der Einführungsveranstaltung verfügten über stark ausgeprägte konstruktivistische und eher gering ausgeprägte transmissive LLV (Lernen der Schüler/innen: Kon: M=5,00 (SD=0,78), Trans: M=1,79 (SD=0,97), Vern: M=5,25 (SD=0,68), Präk: M=4,26 (SD=0,88); Eigenes Lernen: Kon: M=4,90 (SD=0,95), Trans: M=2,25 (SD=1,14), Vern: M=5,24 (SD=0,79), Präk: M=4,25 (SD=1,10)). Die Vorstellungen zum eigenen Lernen und zum Lernen der Schüler/innen unterscheiden sich hinsichtlich der generalisierten transmissiven Vorstellungen (Trans:  $F(1;69)=29,25$ ,  $p<.001$ ,  $\eta^2=.30$ ). Die konstruktivistischen Vorstellungen stimmen sowohl für die generalisierte als auch für die beiden spezifischen Skalen für die beiden Perspektiven überein (Kon:  $F(1;69)=1,19$ ,  $p=ns$ ; Vern:  $F(1;69)=0,10$ ,  $p=ns$ ; Präk:  $F(1;69)=0,40$ ,  $p=ns$ ).

Die Vorstellungen der Studierenden sind im Gegensatz zu internationalen Befunden (Da-Silva et al., 2007; Koballa et al., 2000) stark konstruktivistisch ausgeprägt. Die Ergebnisse stimmen aber mit dem Befund einer nationalen Studie überein (Markic et al., 2006). Im Hinblick auf die Perspektivität der LLV zeigten Brauer und Wilde (2017) ähnliche Unterschiede, jedoch für Studierende in höheren Fachsemestern.

Die Befragung weiterer Kohorten Lehramtsstudierender der Biologie und des Sachunterrichts in der Einführungsveranstaltung erfolgt zu Beginn des WiSe 2017/18. Ebenso wird ein längsschnittliches Design verfolgt, um mögliche Entwicklungen der LLV empirisch zu beschreiben.

## **Lernförderliche Lehr-Lern-Gespräche: Soll-Ist-Vergleich und Implikationen für die Lehrpersonenbildung**

**Andrea Reichmuth**

ZHAW Zürcher Hochschule für Angewandte Wissenschaften, Schweiz; [reid@zhaw.ch](mailto:reid@zhaw.ch)

Lehr-Lern-Gespräche werden im Unterricht häufig eingesetzt (Dubs, 2009). Das Führen lernförderlicher Lehr-Lern-Gespräche ist jedoch nicht trivial (Bittner, 2006). Folglich wird viel Kritik an der Methode geübt (vgl. z.B. ebd.). Diese Kritik an der praktischen Umsetzung darf allerdings nicht mit Kritik am Konzept als solchem gleichgesetzt werden, denn dieses wird von der aktuellen Lehr-Lern-Forschung grundsätzlich nicht infrage gestellt und weitestgehend als sinnvolle Lehr-Lern-Form angesehen.

Vor diesem Hintergrund sollte sich Unterrichtsentwicklung nicht nur damit beschäftigen, welche Unterrichtsmethoden sich als Alternative zum mangelhaft geführten Lehr-Lern-Gespräch anbieten, sondern ebenso thematisieren, wie die Qualität von Lehr-Lern-Gesprächen verbessert werden könnte (Pauli, 2010).

Dieser Forderung nimmt sich der Beitrag an. Dazu wird ein Analyseraster für Lehr-Lern-Gespräche vorgestellt, das basierend auf theoretischen Überlegungen und empirischen Befunden zur Informationsverarbeitung, dem sozialkonstruktivistischen Lehr-Lern-Verständnis und der kognitiven Aktivierung entwickelt wurde. Dieses Analyseraster berücksichtigt neben der Gesprächsstruktur – bestehend aus der Abfolge von Initiierung, Antwort und Feedback (IRF) – zugleich den Gesprächsprozess. Als Indikatoren für dialogische und somit lernförderliche Lehr-Lern-Gespräche stehen in Anlehnung an beispielsweise Alexander (2008) sowie Michaels, O'Connor, Hall und Resnick (2012) eine erweiterte IRF-Struktur sowie als Prozessmerkmale eine hohe inhaltliche Progression, substantielle Beiträge der Lernenden und inhaltliche Vollständigkeit und Präzision.

Neben dem Analyseraster, das grundsätzlich allgemeindidaktisch nutzbar ist, wird zusätzlich die Gesprächsführungspraxis im kaufmännischen Fach «Wirtschaft & Gesellschaft» (W&G) thematisiert: Welche Einstellung haben Lehrpersonen gegenüber dem Lehr-Lern-Gespräch? Wie wird das eigene Gesprächsverhalten eingeschätzt? Welche Gesprächsführungspraxis lässt sich im Unterricht beobachten?

Datengrundlage bilden Befragungsdaten von 157 Deutschschweizer W&G-Lehrpersonen zu Einstellungen und Überzeugungen zum Lehr-Lern-Gespräch sowie eine Videostudie im Fach W&G in neun Berufsschulklassen, die in einem übergeordneten Projekt erhoben wurden. Analyseeinheit für die Untersuchung der Gesprächsführungspraxis ist in Anlehnung an Bauer-Klebl, Euler und Hahn (2001) die Gesprächssequenz, weil davon auszugehen war, dass das lernförderliche Potenzial zwischen unterschiedlichen Gesprächssequenzen variieren würde. Zudem kommt einer Sequenz nach Wells (1996) die grösste Bedeutung zu, weil sie die Verknüpfung der direkt aufeinander bezogenen Gesprächsaktivitäten aufzeigt, die durch eine erste Initiierung ausgelöst werden.

Die Kodierung der Gesprächssequenzen wurde basierend auf dem entwickelten Analyseraster vorgenommen und zeigte sich äusserst reliabel (Cohens Kappa  $\geq .90$ ). Da die untersuchten Gesprächssequenzen aus einer Klumpenstichprobe von neun Lehrpersonen stammen, wurde bei der Auswertung die hierarchische Struktur – bestehend aus der Verschachtelung von Sequenzen in Episoden in Lektionen von Lehrpersonen (vgl. z.B. Mehan, 1979; Wells, 1996) – mehrebenenanalytisch mitberücksichtigt. Dieses Vorgehen war sinnvoll, da insbesondere für die Ebene der Episode durchgängig ein Varianzanteil (ICC-Wert) von mindestens 5 % identifiziert werden konnte; das gilt in der Praxis als Richtwert für die Notwendigkeit von Mehrebenenanalysen (Dyer, Hanges, & Hall, 2005).

Die Ergebnisse zeigen, dass die befragten W&G-Lehrpersonen dem Lehr-Lern-Gespräch ein grosses Potenzial zur Initiierung verständnisbasierten Lernens zusprechen. Die eigene Gesprächsführungspraxis wurde zudem wenig kritisch und ausgeprägt sozialkonstruktivistisch orientiert eingeschätzt (bei vierstufiger Skala:  $M = 3.04\text{--}3.43$ ,  $SD = .52\text{--}.63$ ). Da die neun Lehrpersonen der Videostudie ebenfalls an der Befragung teilnahmen, war eine fallbezogene Datentriangulation bezüglich Selbst- und Fremdeinschätzung möglich. Dabei zeigte sich erwartungskonform eine zu positive Selbsteinschätzung der eigenen Gesprächsführungspraxis: Von 461 analysierten Lehr-Lern-Gesprächssequenzen konnte lediglich 8 % ein dialogischer Charakter zugesprochen werden. Die mehrebenenanalytisch angelegten Auswertungen ergaben zudem eine geringe Abhängigkeit von der Ebene der Lektion ( $ICC < 1\%$ ) und der Ebene der Lehrperson ( $ICC < 10\%$ ), was darauf hindeutet, dass die Lehr-Lern-Gespräche von den neun Lehrpersonen wenig systematisch geführt worden waren.

Die Befunde zum Lehr-Lern-Gespräch im Fach W&G, die weitgehend im Einklang mit der stufen- und fächerübergreifend geäusserten Kritik an der gängigen Gesprächsführungspraxis stehen, verweisen auf Handlungsbedarf im Bereich der Gesprächsführung im Unterricht. Eine Idee für ein Schulungskonzept für Lehrpersonen, das darauf abzielt, die Dialogizität im Klassenzimmer zu fördern und damit einen Beitrag zur Professionalisierung der Lehrpersonenbildung zu leisten, wird im Beitrag ebenfalls dargelegt.

### **Sagen kognitive und motivationale Faktoren eine erfolgreiche Teilnahme an einem mathematischen Schülerwettbewerb vorher?**

**Franziska Rebholz<sup>1,3</sup>, Korbinian Moeller<sup>2,3</sup>, Ulrich Trautwein<sup>1</sup>, Jessika Golle<sup>1,3</sup>**

<sup>1</sup>Universität Tübingen, Deutschland; <sup>2</sup>Leibniz-Institut für Wissensmedien, Tübingen, Deutschland; <sup>3</sup>LEAD Graduate School & Research Network, Eberhard Karls Universität Tübingen, Deutschland; [franziska.rebholz@uni-tuebingen.de](mailto:franziska.rebholz@uni-tuebingen.de)

Theoretischer Hintergrund:

In der Begabtenförderung gelten Schülerwettbewerbe als eine gute Möglichkeit herausfordernde Lernumgebungen anzubieten um damit die Kompetenzen der Schülerinnen und Schüler zu fördern (Fauser, Messner, Beutel, & Tetzlaff, 2007; Oswald, Hanisch, & Hager, 2005; Ozturk & Debelak, 2008a, 2008b; Petersen & Wulff, 2017; Pyrt, 2000). Dennoch werden, wie in allen akademischen Leistungen, auch in der Wettbewerbsleistung interindividuelle Unterschiede beobachtet. Um das Zusammenspiel verschiedener Variablen für das Erbringen einer bestimmten Leistung zu erklären, wurde von Eccles und Kollegen (1983) das Erwartungs-Wert-Modell vorgeschlagen, nach dem vor allem motivationale Faktoren (i.e. Erfolgserwartung und Wertüberzeugungen) einen direkten Einfluss auf die gezeigten Leistungen haben. Vorherige Leistungen und Leistungsvoraussetzungen scheinen dagegen eher indirekt durch ihren Einfluss auf die motivationalen Faktoren spätere Leistungen zu beeinflussen (vgl. Wigfield & Eccles, 2000).

Da es sich bei Schülerwettbewerben jedoch um freiwillige extracurriculare Enrichmentangebote handelt, stellt sich die Frage, inwieweit motivationale Faktoren für die gezeigte Leistung in einem Schülerwettbewerb ebenfalls relevant sind. Für Jugendliche wurden bislang Hinweise darauf gefunden, dass im Zusammenhang mit Schülerwettbewerben insbesondere die motivationalen Faktoren der Erfolgserwartung im Wettbewerb und die relativen Kosten den Erfolg in Schülerwettbewerben vorhersagen können (Stang, Urhahne, Nick, & Parchmann, 2014; Urhahne, Ho, Parchmann, & Nick, 2012). Die vorherige Leistung zeigte sich in diesen Studien ebenfalls prädiktiv für das erfolgreiche Absolvieren einer Wettbewerbsrunde (Stang et al., 2014). Im Vergleich zu Leistungsvorhersagen basierend auf dem Erwartungs-Wert-Modell scheinen die Leistung in Schülerwettbewerben damit auch unmittelbar von vorherigen Leistungen und motivationalen Faktoren abzuhängen.

Obwohl bereits in der Grundschule Kinder die Möglichkeit haben, sich in Wettbewerben zu vergleichen, wurde die Vorhersage der Leistung in einem solchen Wettbewerb durch personenspezifische Merkmale (z.B. Mathematikkompetenz, allgemeinem kognitive Fähigkeiten und motivationale Faktoren) bisher nicht untersucht. In diesem Beitrag wird daher der Frage nachgegangen, welche Faktoren den Erfolg in der Mathematik-Olympiade für Grundschulkindern vorhersagen.

Methode:

Zur Beantwortung der Fragestellung wird auf die Daten von zwei Untersuchungen – Quasi-experimentelles und experimentelles Warte-Kontrollgruppen-Design – zurückgegriffen.

Der erste Datensatz beinhaltet Informationen von 151 Dritt- und Viertklässlerinnen (84♂; Alter: M = 9.14 SD = 0.40), der zweite von 52 Kinder (32♂; Alter: M = 8.73 SD = 0.53).

In beiden Datensätzen wurde die Mathematikkompetenz mittels eines standardisierten Leistungstest (Studie 1: DEMAT 2+ Krajewski, Liehm, & Schneider, 2004; Roick, Göllitz, & Hasselhorn, 2004; Studie 2: Haffner, Baro, Parzer, & Resch, 2005), fluide kognitive Fähigkeiten (Studie 1 & 2, BEFKI-short, Schroeders, Schipolowski, Zettler, Golle, & Wilhelm, 2016) sowie motivationale Faktoren (Selbstkonzept und Interesse an Mathematik) vor der Teilnahme an der Mathematik-Olympiade erfasst. Nach ca. vier Monaten wurde die Leistung in der schwierigsten, dritten Stufe der Mathematik-Olympiade erhoben.

Ergebnisse:

Zur Vorhersage der Wettbewerbsleistung wurden für beide Datensätze multiple lineare Regressionen durchgeführt.

Die motivationalen Faktoren Selbstkonzept und Interesse an Mathematik waren in beiden Untersuchungen nicht prädiktiv für den Wettbewerbserfolg (Studie 1:  $\beta_{\text{Selbstkonzept}} = 0.07$ ,  $p = .640$ ;  $\beta_{\text{Interesse}} = -0.01$ ,  $p = .985$ ; Studie 2:  $\beta_{\text{Selbstkonzept}} = -0.52$ ,  $p = .131$ ;  $\beta_{\text{Interesse}} = 0.26$ ,  $p = .361$ ). Bezüglich der Leistungsparameter zeigte sich, dass in der ersten Untersuchung sowohl fluide kognitive Fähigkeiten als auch die Mathematikkompetenz den Wettbewerbserfolg signifikant vorhersagten (Studie 1:  $\beta_{\text{kognitive Fähigkeiten}} = 0.09$ ,  $p = .003$ ;  $\beta_{\text{Kompetenz}} = 0.37$ ,  $p < .001$ ). In der zweiten Studie zeigte sich dieser Zusammenhang jedoch nicht mehr (Studie 2:  $\beta_{\text{kognitive Fähigkeiten}} = 0.09$ ,  $p = .092$ ;  $\beta_{\text{Kompetenz}} = 0.01$ ,  $p = .192$ ).

Zusammengefasst deuten die Ergebnisse darauf hin, dass den Erfolg in einem mathematischen Schülerwettbewerb in der Grundschule vor allem kognitive Fähigkeiten und mathematische Vorkenntnisse vorhersagten, während motivationale Variablen darüber nicht signifikant zur Güte der Vorhersage beitragen.

## Welche Quellen nutzen Lehrkräfte zur professionellen Weiterentwicklung?

**Anabel Bach, Anja Böhnke, Felicitas Thiel**

Freie Universität Berlin, Deutschland; [anabel.bach@fu-berlin.de](mailto:anabel.bach@fu-berlin.de)

Als zentrale schulische Akteure sind Lehrpersonen maßgeblich verantwortlich für die Unterrichtsqualität, welche wiederum als einer der bedeutsamsten Prädiktoren für Schülerleistungen gilt. Dass Lehrkräfte ihre professionellen Kompetenzen im Laufe ihres Berufslebens stetig erweitern, ist demnach eine wichtige Voraussetzung für Schulentwicklung und -effektivität. Dementsprechend ist die stetige professionelle Weiterentwicklung der Kompetenzen auch ein zentraler Bestandteil in den Standards der Lehrerbildung der Kultusministerkonferenz (KMK, 2004, S. 12). Zahlreiche empirische Studien verweisen jedoch darauf, dass traditionelle Formen professioneller Weiterentwicklung (wie Fortbildungen) im Hinblick auf die Verbesserung von Unterrichts- und Schulqualität häufig wirkungslos bleiben und nur unter ganz bestimmten Voraussetzungen effektiv sind (Darling-Hammond, Wei, Andree, Richardson & Orphanos, 2009; Yoon, Duncan, Lee, Scarloss & Shapley, 2007). Auch zeigen Befunde aus der OECD-Studie TALIS (Teaching and Learning International Survey) darüber hinaus, dass Aktivitäten der Weiterentwicklung, welche von Lehrkräften als besonders hilfreich für die eigene Arbeit erachtet werden, in der Praxis häufig überhaupt nicht vorkommen (Schleicher, 2012, S. 75). Es ist daher nicht verwunderlich, dass das professionelle Wissen von Lehrkräften nach Abschluss des Studiums stagniert oder teilweise sogar abfällt (Kleickmann, Richter, Kunter, Elsner, Besser & Krauss, 2013, S. 98).

Die beschriebenen Befunde verdeutlichen die Notwendigkeit, sich mit weiteren Quellen der professionellen Weiterbildung von Lehrkräften zu beschäftigen. Zu diesem Zweck wurde, in Anlehnung an Kwakman (2003), ein Instrument entwickelt, in dem zwischen kollaborativen (z.B. „Ich führe mit Kollegen gemeinsame Unterrichtsstunden durch.“), individuellen (z.B. „Ich erstelle selber Unterrichtsmaterial, das ich im Unterricht ausprobieren.“) und unterrichtsbezogenen Aktivitäten (z.B. „Zur Nachbereitung meiner Unterrichtsstunden gleiche ich meinen Unterrichtsplan mit dem tatsächlichen Verlauf der Stunde ab.“) unterschieden wird (insgesamt 16 Items).

Folgende Forschungsfragen stehen für die vorliegende Untersuchung im Mittelpunkt: Lässt sich die theoretisch angenommene Faktorenstruktur des Instruments empirisch abbilden? Wie verbreitet sind die verschiedenen Formen professioneller Weiterentwicklung im Vergleich? Welche individuellen (z.B. Arbeitszufriedenheit, Selbstwirksamkeitserwartung, Fehlerorientierung, Geschlecht, Alter) und welche organisationalen Merkmale auf Schulebene (Schulkultur, Schulklima, Führungsverhalten der Schulleitung) stehen mit den drei verschiedenen Formen professioneller Weiterentwicklung im Zusammenhang?

Zur Beantwortung der Forschungsfragen stehen Daten einer in zwei deutschen Bundesländern durchgeführten Lehrkräftebefragung zur Verfügung (N= 658 Lehrkräfte aus 51 allgemeinbildenden Grund- und weiterführenden Schulen). Durchschnittlich haben 42 Prozent der Lehrkräfte einer Schule die Fragebögen beantwortet (Range von 2 bis 34 Lehrkräften bzw. 5 bis 90 Prozent), die durchschnittliche Clustergröße beträgt 13 Lehrkräfte. Im Hinblick auf ausgewählte Merkmale (Alter, Geschlecht, Berufserfahrung, Beschäftigungsverhältnis) entspricht die Stichprobenverteilung annähernd der Verteilung in der Population. Neben der Präsentation deskriptiver Ergebnisse werden unter Berücksichtigung der Mehrebenenstruktur der Daten eine Konfirmatorische Faktorenanalyse sowie Mehrebenenpfadmodelle berechnet.

Da die Analysen noch nicht vollständig abgeschlossen sind, können an dieser Stelle keine finalen Ergebnisse präsentiert werden. Die vermutete Faktorenstruktur kann nach ersten Ergebnissen bestätigt werden, kollaborative Aktivitäten (M = 3.93, SD = .91) werden von den Lehrkräften vergleichsweise seltener durchgeführt, und es zeigen sich auf Individualebene unter anderem signifikante Zusammenhänge mittlerer Effektstärke mit der Selbstwirksamkeitserwartung und der Arbeitszufriedenheit bei allen drei Aktivitäten. Auch die Fehlerorientierung (Böhnke & Thiel, 2016) steht offenbar in Zusammenhang mit den Aktivitäten der professionellen Weiterentwicklung, vor allem zur Vorhersage der unterrichtsbezogenen Aktivitäten scheint dieses Merkmal ein vergleichsweise wichtiger Prädiktor zu sein ( $\beta = .24^{***}$ ). Die Interkorrelationen zwischen den drei Formen professioneller Weiterentwicklung liegen im Bereich mittlerer Stärke ( $r = .35^{**}$  bis  $.45^{**}$ ).

## Vertrauen im Teampraktikum

**Jennifer Waber<sup>1</sup>, Gerda Hagenauer<sup>1</sup>, Lea de Zordo<sup>2</sup>**

<sup>1</sup>Universität Bern, Schweiz; <sup>2</sup>Pädagogische Hochschule Bern, Schweiz; [gerda.hagenauer@edu.unibe.ch](mailto:gerda.hagenauer@edu.unibe.ch)

Einleitung:

Angehende Lehrpersonen absolvieren während ihrer berufspraktischen Ausbildung verschiedene Schulpraktika, welche – aus organisatorischen und/oder didaktischen Gründen – häufig als Teampraktika angelegt sind. Damit die Studierenden die sich dadurch bietenden Gelegenheiten des voneinander und miteinander Lernens auch nutzen können, müssen einige Voraussetzungen erfüllt sein (Baeten & Simons, 2014). Gelingensbedingungen für eine erfolgreiche Kooperation im Praktikum umfassen neben strukturellen und organisatorischen Aspekten auch Handlungskompetenzen und zwischenmenschliche Bedingungen wie einen respektvollen Umgang, Sympathie und Vertrauen.

Vertrauen entsteht durch interpersonale Ereignisse und übernimmt eine handlungsentscheidende Funktion (Deutsch, 1958). So wirkt sich Vertrauen zwischen Lehrkräften positiv auf deren Kooperation und Wohlbefinden aus (Schweer, 1997, 2010). Die Effekte von Vertrauen auf das Erleben und Verhalten der Praktikant/innen sowie die Entstehung von Vertrauen im Schulpraktikum wurden bisher noch nicht untersucht und bilden den Fokus der vorliegenden Studie.

Fragestellung:

In der vorliegenden Studie werden die folgenden Hauptfragestellungen überprüft:

(1) Welche Effekte hat die wahrgenommene Vertrauensqualität zwischen der Praxislehrperson und der Praktikantin/des Praktikanten sowie zwischen den Praktikant/innen auf das Erleben (Emotion/Motivation) und die Anstrengungsbereitschaft im Schulpraktikum?

Es wird davon ausgegangen, dass eine hohe Vertrauensqualität (hohes spezielles interpersonelles Vertrauen zur Praktikumslehrkraft und zum/zur Praktikumpartner/in) mit positiveren Emotionen und einem positiveren Motivationsmuster (intrinsische und soziale Motivation) sowie mit einer erhöhten Anstrengungsbereitschaft assoziiert ist. Kontrolliert wird um das allgemeine interpersonelle Vertrauen.

Des Weiteren werden die wahrgenommenen konstitutiven Elemente einer „vertrauensvollen Beziehung“ sowie deren Aufbau vertieft untersucht:

(2) Wodurch zeichnet sich eine „vertrauensvolle“ Beziehung im Schulpraktikum aus?

(3) Welche Momente („relational turning points“; Docan-Morgan & Manusov, 2009) tragen zur Vertrauensbildung bzw. im negativen Sinne auch zum Vertrauensverlust bei?

Methode:

Es wurde eine Mixed-Methods-Studie unter Verwendung eines sequentiellen explanativen Designs (Creswell & Plano Clark, 2007) verwendet.

Zur Überprüfung von Fragestellung 1 wurden insgesamt 114 Studierende an einer Schweizer Pädagogischen Hochschule, die soeben ein Fachpraktikum absolviert hatten, mittels eines Online-Fragebogens befragt. Es wurden überwiegend standardisierte Skalen herangezogen, welche für den Praktikumskontext adaptiert wurden (z.B. allgemeines und spezielles interpersonelles Vertrauen; Dalbert, Donat und Dietrich, 2007-2009; Positive und negative Emotionen: Krohne et al., 1996; Motivation und Anstrengungsbereitschaft; Schiefele, Moschner und Husstege, 2002).

Zudem nahmen 25 Studierende derselben Pädagogischen Hochschule an einer Vertiefungsstudie teil, in der als Methode das Leitfadenterview eingesetzt wurde. Zur Identifikation von zentralen das Vertrauen beeinflussenden Situationen (= relational turning points) (Docan-Morgan & Manusov, 2009) wurde die „critical incident technique“ (Kain, 2004) im Interview angewendet. Die Vertiefungsstudie diente der Überprüfung der Fragestellungen 2 und 3. Zur Auswertung der Interviews kommt die inhaltlich strukturierende Inhaltsanalyse nach Mayring (2010) zum Einsatz.

Ergebnisse:

Die Ergebnisse von Strukturgleichungsmodellen bestätigen die erwarteten Zusammenhänge (Fragestellung 1): Umso höher das Vertrauen zur Praktikumslehrkraft und zur Praktikumpartnerin beschrieben wird, desto positiver ist auch das emotionale Erleben im Schulpraktikum und desto höher sind die intrinsische und soziale Motivation im Praktikum sowie die Anstrengungsbereitschaft ausgeprägt.

Die Auswertung der Interviews (Fragestellungen 2 und 3) befindet sich zum Zeitpunkt der Beitragseinreichung noch in Arbeit. Erste Analysen zeigen, dass es auf Basis des Konzepts der „relational turning points“ gut gelingt, zentrale Momente der Vertrauensbildung zu identifizieren und zu beschreiben. Die interviewten Praktikant/innen sind auch sehr gut in der Lage, konstitutive Elemente von Vertrauen zu explizieren. Zudem zeigt sich, dass die Praktikant/innen beschreiben können, wie sie sich zum Zeitpunkt solcher „relational turning points“ gefühlt haben und wie sich diese auf den allgemeinen Verlauf und insbesondere das weitere emotionale Erleben im Praktikum ausgewirkt haben. So deutet sich an, dass sich Praktikant/innen nach „relational turning points“, die sich positiv auf die Vertrauensbildung auswirkten, im Praktikum entspannter und zugleich motivierter gefühlt zu haben.

Die Ergebnisse werden im Hinblick auf die Gestaltung von Teampraktika unter dem Aspekt des „Vertrauens“ in Beziehungen diskutiert.

## **Kulturbezogene Aktivitätsmuster von Lehramtsstudierenden in ihrem Zusammenhang mit Facetten professioneller Kompetenz**

**Jana Costa, Barbara Drechsel**

Otto-Friedrich-Universität Bamberg, Deutschland; [jana.costa@uni-bamberg.de](mailto:jana.costa@uni-bamberg.de)

Lehrkräfte sind als professionell Handelnde mit vielfältigen Herausforderungen in ihrer alltäglichen Praxis konfrontiert. Vor dem Hintergrund fortschreitender Globalisierungs- und Pluralisierungstendenzen gewinnen dabei auch kulturbezogene Anforderungen zunehmend an Bedeutung (Costa 2017, S. 16 ff.; Lang-Wojtasik 2008, S. 71). Als implizite Vermittler und Multiplikatoren kultureller Deutungsmuster können Lehrende beispielsweise nicht nur die grundlegenden Einstellungen und Überzeugungen von SchülerInnen (mit-)beeinflussen, sondern darüber hinaus verschiedene Perspektiven auf Welt(en) ermöglichen und ihnen die Voraussetzungen für vielfältige Gestaltungs- und Interpretationsoptionen mit auf den Weg geben. Lehrkräfte sind somit nicht nur selbst kulturgeprägt, sondern auch kulturprägend. Umso wichtiger erscheint es bereits in der Lehrerbildung die Grundsteine für professionelles Handeln zu legen und angehende Lehrkräfte auf kulturbezogene Herausforderungen im schulischen Alltag vorzubereiten. Die Erforschung kulturbezogener Professionalisierungs- und Sensibilisierungsprozesse in der Lehramtsausbildung ist daher von grundlegender Bedeutung. Die Relevanz der Thematik spiegelt sich in der bisherigen Forschung zur kulturellen Bildung wie auch zur Professionalisierung allerdings kaum wider

(Scheunpflug und Prenzel 2013, 2 ff.). Insbesondere in der standardisierten Bildungsforschung spielen kulturbezogene Erfahrungen und Lernprozesse von Lehramtsstudierenden kaum eine Rolle. Auch mit Blick auf ein enges Verständnis von Kultur als Hochkultur finden sich keine Studien, in welchen explizit die Gruppe der Lehramtsstudierenden untersucht wird. Befunde zu Transfereffekten kultureller Bildung legen jedoch nahe, dass kulturbezogene Aktivitäten auch in Zusammenhang mit professionsbezogenen Merkmalen von Lehramtsstudierenden stehen könnten – eine differenzierte Erforschung von Aktivitätsprofilen und deren Zusammenhang mit professionsrelevanten Merkmalen scheint daher zentral für die Lehrerbildungsforschung. Ausgehend von einem weiten Kulturbegriff (vgl. z.B. Geertz 2010; Williams 1989) wird im BMBF-geförderten Forschungsprojekt „KulturLeBi“ an der Universität Bamberg deshalb die zentrale Bedeutung von Lehrkräften für kulturbezogene Bildungsprozesse fokussiert. Die theoretische Aufarbeitung des Forschungsfeldes und eine Evaluation der Analyseperspektiven verschiedener Datenquellen zeigt dabei, dass in existierenden Datenquellen wie der Studierendenkohorte des NEPS, der HISBUS-Befragung oder PaLea ein z.T. großes Potential bezüglich der Untersuchung von kulturbezogenen Aktivitäten und professionsrelevanten Merkmalen von Lehramtsstudierenden verborgen liegt und eine Vielzahl an relevanten Forschungsfragen auf Basis bestehender Datenquellen bearbeitet werden können (Costa 2017).

Das eingereichte Poster präsentiert erste Ergebnisse einer Sekundäranalyse zu kulturellen Präferenzen von Lehramtsstudierenden in ihrem Zusammenhang mit Facetten professioneller Kompetenz. Die Datengrundlage bildet die Studierendenkohorte des NEPS (Leibniz-Institut für Bildungsverläufe 2017). Das besondere Potential des Datensatzes liegt in der Studierendenkohorte mit einem Lehramtsstudierenden-Oversampling (n = 5500), welches differenzierte Analysen in Bezug auf Lehramtsstudierende zulässt. Basierend auf einem weiten Kulturverständnis können mit den Daten nicht nur kulturell-ästhetische Aktivitäten von Lehramtsstudierenden, sondern darüber hinaus auch weitere kulturbezogene Aktivitäten, wie (hochschul-)politisches und ehrenamtliches Engagement untersucht werden. Die Betrachtung der Aktivitätsprofile von Lehramtsstudierenden gibt dabei Hinweise über deren kulturelle Präferenzen und eröffnet einen Ausblick auf die kulturbezogene Prägung – in welcher Beziehung diese allerdings mit Facetten professioneller Kompetenz steht, bleibt zunächst ungeklärt. In dem Poster werden deshalb zudem mögliche Verbindungen einzelner kulturbezogener Aktivitäten mit Facetten professioneller Handlungskompetenz nach dem Modell von Baumert und Kunter diskutiert (Baumert und Kunter 2013).

Zusammenfassend werden folgende Fragen thematisiert:

- 1) Welchen kulturbezogenen Aktivitäten gehen angehende Lehrkräfte nach?
- 2) Stehen diese Aktivitäten in einem Zusammenhang mit Überzeugungen und Werthaltungen als Teilkomponente des Modells professioneller Handlungskompetenz (Baumert und Kunter 2013)?

Im Rahmen des Posters werden deskriptive Befunde zu kulturbezogenen Aktivitäten von Lehramtsstudierenden sowie erste Ergebnisse bezüglich theoretisch relevanter Zusammenhänge zwischen einzelnen kulturbezogenen Aktivitäten und den Überzeugungen in Bezug auf Unterrichten und Lernen von Lehramtsstudierenden vorgestellt. Dies ermöglicht es Erkenntnisse über kulturbezogene Bedingungs- und Einflussfaktoren auf einen Teilaspekt professionellen Handelns zu generieren und damit bildungspolitisch relevante Potentiale in der Lehrerbildung zu identifizieren. Auf Basis dieses grundlegenden Wissensbestandes kann dann darüber nachgedacht und diskutiert werden, welche normativen Zielvorstellungen durch die Lehrerbildung anvisiert werden sollen und welche Mittel hierfür geeignet scheinen.

### **„National Academics Panel Study (Nacaps)“ – Eine Panelstudie zum wissenschaftlichen Nachwuchs in Deutschland**

**Kolja Briedis, Uwe Ruß, Andreas Sarcletti, Ulrike Schwabe, Jacob Tesch**

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Deutschland; [schwabe@dzhw.eu](mailto:schwabe@dzhw.eu)

Die Ausbildung des wissenschaftlichen Nachwuchses hat gesamtgesellschaftlich eine essenzielle Bedeutung für die Leistungs-, Wettbewerbs- und Innovationsfähigkeit des Wissenschafts- und Wirtschaftsstandortes Deutschland. Für Individuen ist der Erwerb von Bildungszertifikaten auf höherem tertiärem Niveau mit vorteilhaften Arbeitsmarkchancen verbunden (Enders/Bornmann 2001, Heineck/Matthes 2012) und begünstigt speziell die Besetzung von Führungs- und Elitepositionen in unterschiedlichen Bereichen (Hartmann/Kopp 2001). Jedoch können zahlreiche aktuelle wissenschaftliche und wissenschaftspolitische Fragen zur Qualifizierung, Rekrutierung sowie den beruflichen Werdegängen des wissenschaftlichen Nachwuchses derzeit nicht adäquat empirisch beantwortet werden, da es an einer einheitlichen und belastbaren Datenbasis fehlt (BuWiN 2017).

Mit der „National Academics Panel Study (Nacaps)“, einer theoretisch fundierten Multi-Kohorten-Panelstudie, wird ein wichtiger Beitrag dazu geleistet, diese Datenlücke zu schließen. Für die Erklärung individueller Qualifizierungs-, Erwerbs- und Lebensverläufe von Hochqualifizierten werden zentrale theoretische Konstrukte wie

- Wissenserwerb und Kompetenzentwicklung
- Karrierestrategien und -entscheidungen
- Persönlichkeitsmerkmale und Selbstwirksamkeitskonzepte
- Soziale Netzwerke
- Monetäre und nicht-monetäre Bildungserträge
- Promotionsbedingungen an Hochschulen und außeruniversitären Forschungseinrichtungen
- Akademische Arbeitsmärkte
- Vereinbarkeit von Beruf und Familie

operationalisiert und in regelmäßigen Abständen überwiegend prospektiv erhoben.

Damit lassen sich dezidiert Fragen nach den Determinanten akademischer Karriereverläufe inner- und außerhalb der Wissenschaft auf Individual- und Kontextebene sowie den interdependenten Beziehungen beider Ebenen aus einer kausalanalytischen Perspektive beantworten. Gleichsam kann die relative Bedeutung asriptiver Merkmale (wie soziale Herkunft) und erworbener, leistungsbezogener Merkmale (wie vorgelagerte Bildungsabschlüsse und Promotionsnoten) analysiert werden.

Damit lassen sich exemplarisch folgende Forschungsfragen empirisch beantworten:



- Wie wirken sich unterschiedliche formelle und informelle Promotionskontexte, Förderbedingungen und Mentoring in der Promotions- und Postdoktorandenphase auf den Kompetenzerwerb und die Sozialisation des wissenschaftlichen Nachwuchses aus?
- Welche Determinanten beeinflussen den Promotionserfolg und den Karriereverlauf und in welchem Verhältnis stehen hierbei strukturelle, individuelle und askriptive Merkmale der Promovierenden und Promovierten?
- In welchem Verhältnis stehen individuelle Entscheidungsprozesse und institutionelle, systemische Selektions- und Auswahlprozesse?
- Wie stellt sich die berufliche Situation der Promovierten im Wissenschaftssystem im Vergleich zu den Promovierten dar, die das Wissenschaftssystem verlassen haben?
- Welche individuellen und gesellschaftlichen Renditen erbringen Promotionen innerhalb und außerhalb der Wissenschaft?

Nacaps ist als Online-Erhebung konzipiert und startet mit der Befragung einer ersten Promovierendenkohorte Anfang 2019; die Befragung einer ersten Promoviertenkohorte soll im Jahr 2021 folgen. Angestrebt wird eine Vollerhebung der Promovierenden und Promovierten aller promotionsberechtigten Hochschulen in Deutschland. Diese Start- und nachfolgende Kohorten werden über einen Zeitraum von mindestens 15 Jahren wiederholt befragt, um den Übergang von der Promotionsphase in nachgelagerte Arbeitsmarktpositionen sowie die weitere Karriereentwicklung erfassen zu können. Zugleich werden regelmäßig neue Promovierenden- und Promoviertenkohorten in das Panel aufgenommen.

Ziel des Projekts ist es, (a) regelmäßig deutschlandweit repräsentative Querschnittsdaten zur Qualifizierung des wissenschaftlichen Nachwuchses sowie (b) Längsschnittdaten zu den individuellen Karriereverläufen Promovierender und Promovierter innerhalb und außerhalb der Wissenschaft zu erheben und der Scientific Community zur Verfügung zu stellen. Damit leistet Nacaps eine wichtige Datengrundlage im Bereich theoriegeleiteter empirischer Bildungs- und Hochschulforschung. Die bestehende längsschnittliche Dateninfrastruktur in Deutschland (beispielsweise SOEP, PIACC-L und NEPS) wird somit um eine weitere Panelstudie, deren Fokus auf wissenschaftlichem Nachwuchs liegt, ergänzt.

Mit dem Poster-Beitrag werden das Erhebungsdesign und die zentralen theoretischen Konstrukte und Forschungsfragen dargestellt sowie ein Ausblick auf Analysepotenziale des Nacaps gegeben.

### **Einsatz digitaler Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe: Eine Metastudie zur Lernwirksamkeit**

**Delia Hillmayr, Lisa Ziernwald, Frank Reinhold, Kristina Reiss**

Technische Universität München, Deutschland; [delia.hillmayr@tum.de](mailto:delia.hillmayr@tum.de), [lisa.ziernwald@tum.de](mailto:lisa.ziernwald@tum.de)

Theoretischer Hintergrund

Die Verwendung digitaler Medien in Schule und Unterricht ist nicht neu. Insbesondere die naturwissenschaftlichen Unterrichtsfächer und Mathematik weisen durch techniknahe Themen hohes Potenzial für einen lernwirksamen Einsatz digitaler Medien auf. Unter welchen Bedingungen sich Effekte durch die Verwendung digitaler Medien im mathematisch-naturwissenschaftlichen Unterricht in Bezug auf das Lernen zeigen, wurde in zahlreichen Studien untersucht, wobei sich in den Ergebnissen nicht immer ein einheitliches Bild zeigt (vgl. Özyurt, Özyurt, Güven & Baki, 2014; Perry & Steck, 2015). In bisherigen Metaanalysen zum Einsatz digitaler Medien stehen zum Teil nur bestimmte Unterrichtsfächer oder einzelne Lernprogrammtypen (vgl. Ma, Adesope, Nesbit & Liu, 2014; Steenbergen-Hu & Cooper, 2014) im Fokus oder es wurden ausschließlich Studien aus bestimmten Ländern (vgl. Bayraktar, 2001) in die Analysen einbezogen. Der vorliegenden Metaanalyse liegt ein Datensatz zugrunde, der breiter angelegt ist, aus systematischer Recherche internationaler Literatur resultiert und sich durch hohe methodische Qualitätskriterien auszeichnet.

Fragestellung

In der vorliegenden Forschungssynthese geht es um die Frage, wie sich digitales Lehren und Lernen im mathematisch-naturwissenschaftlichen Unterricht (Mathematik, Biologie, Chemie und Physik) der Sekundarstufe auf die Lernleistung der Schülerinnen und Schüler sowie deren Einstellung zum jeweiligen Unterrichtsfach auswirkt und welche Bedingungen des Einsatzes digitaler Medien dabei begünstigend oder weniger begünstigend wirken.

Methode

Die Analysen basieren auf Daten aus insgesamt 79 Primärstudien, die seit dem Jahr 2000 publiziert wurden. Dabei wurde jeweils die Wirksamkeit des Einsatzes digitaler Medien bezüglich der Leistung von Schülerinnen und Schülern des Sekundarbereichs im Vergleich zu Kontrollgruppen untersucht, die ohne den Einsatz digitaler Medien unterrichtet wurden. Diese Studien wurden in drei einschlägigen Datenbanken (ERIC, Scopus und Web of Science) recherchiert. Bei der Auswahl der Studien wurde ein hoher Qualitätsstandard angesetzt: Es wurden nur Studien in die Analyse aufgenommen, denen ein experimentelles oder quasi-experimentelles Design (Prä-Post-Kontrollgruppen-Design) zugrunde liegt. Die Kodierung der Daten erfolgte durch zwei voneinander unabhängige Rater und anhand eines vorab detailliert ausgearbeiteten Leitfadens. Für die Berechnung der Effekte wurde ein Random-effects model verwendet.

Ergebnisse

Die Berechnung der Gesamteffekte zeigt, dass sich der Einsatz digitaler Medien im mathematisch-naturwissenschaftlichen Unterricht mit einem mittelstarken Effekt ( $g = .62$ ,  $p < .001$ ) positiv auf die Leistung der Schülerinnen und Schüler auswirkt. Auch bezüglich der Einstellung zum Fach lässt sich ein positiver, wenn auch kleinerer Gesamteffekt ( $g = .41$ ,  $p < .05$ ) nachweisen. Positive Effekte zeigen sich sowohl über alle untersuchten Unterrichtsfächer als auch über sämtliche Klassenstufen des Sekundarbereichs hinweg. Bei den Analysen konnten keine signifikanten Moderationseffekte festgestellt werden. Dennoch zeigen sich bei Betrachtung der Einzeleffekte die folgenden Tendenzen:

Lernprogramme, die adaptive Funktionen aufweisen, üben größere positive Effekte auf die Lernleistung aus als Lernprogramme, die keine adaptiven Funktionen aufweisen. Die Lernwirksamkeit ist außerdem höher, wenn die Schülerinnen und Schüler bei der Nutzung des entsprechenden Lernprogramms zusätzlich Unterstützung durch die Lehrkraft oder durch Mitschülerinnen und Mitschüler erhalten. Es zeigt sich auch, dass der positive Effekt auf die Lernleistung größer ist, wenn die Schülerinnen und Schüler das Medium in Paaren verwenden im Vergleich zu einer Nutzung allein. Wenn digitale Medien ergänzend zu anderen Unterrichtsmethoden eingesetzt wurden, zeigt sich ebenfalls ein größerer positiver Effekt als wenn sie traditionelle Methoden über ganze Lerneinheiten hinweg vollständig ersetzen. Hat vor der Intervention ein Lehrertraining bezüglich des eingesetzten Mediums stattgefunden, zeigen sich größere positive Effekte auf die Leistung der Schülerinnen und

Schüler. Da die Dauer der Intervention in den untersuchten Studien stark variiert, wurde auch dieser Faktor analysiert. Es zeigt sich folgendes Ergebnis: Je länger die Dauer der Intervention, desto niedriger die Lernleistung der Schülerinnen und Schüler.

Die gewonnenen Erkenntnisse werden in einer Praxisbroschüre für Lehrkräfte entsprechender Unterrichtsfächer aufbereitet und publiziert. Dadurch soll der Transfer in die Praxis gefördert werden.

## **Analyse der Diagnosekompetenzen von Studierenden des Grundschullehramtes Mathematik in simulationsbasierten Lernumgebungen**

**Angelika Wildgans<sup>1</sup>, Kristina Reiss<sup>1</sup>, Andreas Obersteiner<sup>2</sup>, Frank Fischer<sup>3</sup>**

<sup>1</sup>Technische Universität München - School of education, Deutschland; <sup>2</sup>Pädagogische Hochschule Freiburg, Deutschland;

<sup>3</sup>Ludwig-Maximilians-Universität München, Deutschland; [a.wildgans@tum.de](mailto:a.wildgans@tum.de)

Für eine adaptive Unterrichtsgestaltung sind diagnostische Kompetenzen von Lehrkräften wesentlich (Karst, 2012; KMK, 2004; Lorenz, 2011; Ostermann, Leuders & Nückles, 2015; Südkamp & Praetorius, 2017; Weinert, Schrader & Helmke, 1990). Ein Teil der diagnostischen Kompetenz im Fach Mathematik bezieht sich auf Fehler, da diese wichtige Indikatoren für Verständnisschwierigkeiten und Fehlvorstellungen sind (Beitlich, Moll, Nagel & Reiss, 2015; Brown & Burton, 1978; Heinrichs, 2015; Radatz, 1980). Ein anderer Aspekt der Diagnose im Unterricht ist die Leistungsbeurteilung. Derzeit wird vor allem anhand der sozialen Norm beurteilt. Dies zeigt sich darin, dass Lehrende sehr gut die Leistungen ihrer Schülerinnen und Schüler in eine Rangfolge bringen können, jedoch Schwierigkeiten haben, diese anhand kriterialer Bezugsrahmen zu bewerten (Südkamp, Möller & Pohlmann, 2008). Da eine möglichst präzise Leistungsdiagnose helfen kann, individuelle Schwierigkeiten zu identifizieren, ist es zusätzlich wünschenswert, die Fähigkeiten anhand gängiger Kompetenzstufenmodelle einzuschätzen (Beitlich et al., 2015; Reiss & Winkelmann, 2009; Reiss, Roppelt, Haag, Pant & Köller, 2012; Reiss & Obersteiner, 2017 (in Druck)).

Diese beiden Teilaspekte diagnostischer Kompetenz zeigen exemplarisch auf, weshalb der diagnostische Entscheidungsprozess bereits in der ersten Phase der Lehrerbildung theoriegeleitet vermittelt werden sollte. In der mathematikdidaktischen Forschung und der Pädagogik wurde dies bislang kaum diskutiert (Oser & Oelkers, 2001). Hingegen gibt es in anderen Disziplinen, beispielsweise der Medizin, bereits Untersuchungen zum diagnostischen Entscheidungsprozess (Bowen, 2006) und der Implementierung im Studium (Fischer, Aulinger & Kopp, 2005; Klopp, Stark, Kopp & Fischer, 2013). Zur Förderung der diagnostischen Fähigkeiten hat sich die DFG-geförderte Forschergruppe Cosima (FOR 2385) das Ziel gesetzt, die Erkenntnisse und Abläufe der Diagnose im Bereich der Medizin und der Lehramtsausbildung auszutauschen, zusammenzuführen, gegenseitig zu nutzen und in der Lehre an den Hochschulen zu verankern. Ziel sind spezifische (empirisch getestete) Interventionsmaßnahmen, die die einzelnen Schritte des Diagnoseprozesses bei den Studierenden fördern (Karing & Seidel, 2017).

Den Diagnoseprozess definiert die Forschergruppe anhand acht epistemisch-diagnostischer Aktivitäten (edAs) (Fischer et al., 2014), die einen prototypischen Verlauf einer wissenschaftlichen Argumentation beschreiben. Dieses Prozessmodell stimmt in den ersten sechs Schritten größtenteils mit dem Diagnoseprozessmodell des NeDiKo-Netzwerkes überein (Herppich et al., 2017), ergänzt dieses jedoch um die beiden Schritte (7) Schlussfolgerungen ziehen und (8) Prozesse und Ergebnisse kommunizieren und überprüfen. Die edAs stellen dabei einzelne, klar voneinander getrennte Prozessschritte dar, die in unterschiedlicher Reihenfolge und Häufigkeit auftreten können (Fischer et al., 2014).

Wird dieses Modell auf den Diagnoseprozess übertragen, stellt sich die Frage, welche der Prozessschritte genutzt werden, inwiefern Fehler und Fehlvorstellungen erkannt werden und ob das Diagnoseurteil akkurat, sowie der Prozess effizient ist. Interessant ist hierbei auch der Vergleich zwischen Studienanfängerinnen und fortgeschrittenen Studierenden.

Hierfür wurden „virtuelle“ Schülerinnen und Schüler für eine simulationsbasierte Lernumgebung konzipiert, die bestimmte mathematische Kompetenzstufen und Fehlvorstellungen aufweisen. Die originalen Schülerlösungen wurden dabei Pilotierungsdaten von VERA 3 aus den Jahren 2015-2017 entnommen. Im Folgenden wurden Studierende des Lehramts Grundschule mit Fach Mathematik aufgefordert, die Schülerinnen und Schüler zu diagnostizieren, wobei sie frei wählen konnten, welche Aufgaben sie zur Diagnose heranzogen. Zusätzlich wurde das Professionswissen der Studierenden erhoben.

Wir stellen die simulationsbasierte Lernumgebung vor und berichten die Ergebnisse der Datenerhebung. Aufgrund der fehlenden praktischen Erfahrungen der Studierenden wird erwartet, dass die Diagnosekompetenzen der Studienanfänger – ohne praktische Erfahrungen – weniger ausgeprägt sind als die fortgeschrittener Studierender. Diese fehlenden Erfahrungen könnten sich im Diagnoseprozess in der Akkuratheit, Effizienz und Angemessenheit der Begründungen zeigen. Hierbei ist auch von Interesse, welchen Einfluss profundes Professionswissen, Motivation oder Erfahrung auf den Diagnoseprozess und die Diagnosequalität haben.

Die simulierte Lernumgebung mit realen Schülerlösungen und virtuellen Kindern bietet somit eine dem Vorwissen von Studierenden angepasste Möglichkeit, bereits in der ersten Phase der Lehramtsausbildung den Grundstein für eine theoriegeleitete und praxisnahe Diagnose im Mathematikunterricht zu legen.

## **Implementation im schulischen Kontext: Die Bedeutung von Ressourcen, Barrieren und Akzeptanz für die Materialnutzung**

**Lina Neuber, Sarah-Ines Meudt, Nina Zeuch, Elmar Souvignier**

Westfälische Wilhelms-Universität Münster, Deutschland; [l.neuber@web.de](mailto:l.neuber@web.de)

Theoretischer Hintergrund

Die Implementation von Innovationen in den Schulalltag trifft auf eine Vielzahl von Herausforderungen. Eine besondere Rolle für die tatsächliche Nutzung neuer Materialien spielt dabei die Akzeptanz von innovativen Unterrichtskonzepten (Sterling-Turner & Watson, 2002). Wird ein Verfahren von der Person, die es implementieren soll, akzeptiert, wird die Implementation wahrscheinlich mit höherer Wiedergabetreue durchgeführt (Lakin & Shannon, 2015). Zusätzlich zu einer hohen Akzeptanz sind aber noch weitere Bedingungen für eine nachhaltige Implementation nötig (Shapiro & Eckert, 1994). Vermittelnde Einflussfaktoren können an einer Schule oder bei einer Person vorliegende Ressourcen und Barrieren sein, aus denen die Machbarkeit einer Maßnahme resultiert (vgl. Petermann, 2014). Für eine gelungene Implementation sind somit Akzeptanz und Machbarkeit zentrale Größen, die auf der Ebene der Lehrkräfte, der Schülerinnen und Schüler, der Schulleitung oder auch der Bildungsadministration erfasst werden können. Ob eine hohe Akzeptanz zu regelmäßiger Materialnutzung führt, sollte also unter anderem davon abhängig sein, welche konkreten Ressourcen und Barrieren für die Nutzung vorliegen.

## Fragestellung

Die vorliegende Studie stellt die Frage nach konkreten Ressourcen und Barrieren für Innovationen im schulischen Kontext sowie nach Zusammenhängen zwischen der Akzeptanz von neuen Fördermaterialien, Ressourcen und Barrieren für den Implementationsprozess und der Nutzung von Material im Schulalltag. Inwiefern wirkt sich die Machbarkeit (in Form von Implementationsbarrieren und -ressourcen) auf den Zusammenhang zwischen der Akzeptanz und dem Einsatz von Lesefördermaterialien aus? Wie wirken sich unterschiedliche Implementationsstrategien (top-down vs. bottom-up) auf die wahrgenommenen Barrieren und Ressourcen sowie auf den Materialeinsatz aus?

## Methode

Im Rahmen eines Forschungsprojektes zur Evaluation der Implementation von Konzepten zur Leseförderung in der Sekundarstufe (BiSS-EiLe) nahmen 15 Multiplikator/innen verschiedener Schulen in zwei deutschen Verbänden an einem leitfadengestützten Interview teil. In ca. 30-minütigen Telefonaten wurden Fragen zum Vorgehen der Schulen bei der Materialeinführung, der Nutzung und Akzeptanz des Fördermaterials sowie zu möglichen Ressourcen und Barrieren beantwortet.

Die Auswertung erfolgte anhand einer Kombination aus qualitativen und quantitativen Vorgehensweisen. A priori wurde ein Kategoriensystem auf Basis bisheriger theoretischer und empirischer Erkenntnisse erstellt, das a posteriori um neue Aspekte ergänzt wurde. Die Antworten wurden in die jeweiligen Kategorien eingeordnet und gewichtete Summenwerte für vorliegende Ressourcen und Barrieren erstellt. Akzeptanz und Nutzung wurden über eine Likert-Skala erfasst. Regressions- und Moderationsanalysen geben Einblick in die Zusammenhänge zwischen Akzeptanz, Ressourcen/Barrieren und Nutzung.

## Ergebnisse

Neben der Zeit (vgl. Langley et al., 2010) und der finanziellen Unterstützung (vgl. Forman et al., 2009) erweisen sich eine positive Einstellung der Schulleitung, der informelle Austausch, die Wirksamkeit und Akzeptanz auf Schülerschulebene und Maßnahmen zur Herstellung einer Verbindlichkeit über alle Schulen hinweg ebenfalls als konkrete bedeutsame Ressourcen. Als häufige Barrieren werden die unzureichende oder zu gering ausgeprägte schulinterne Kooperation, zu wenig bzw. unzureichende Verbindlichkeit, geringe Innovationsbereitschaft der Lehrkräfte und zu wenig Material genannt. Die Moderationsanalysen zeigen keine signifikanten Einflüsse der Ressourcen und Barrieren auf den Zusammenhang zwischen Akzeptanz und Materialnutzung, lassen aber zumindest die Tendenz erkennen, dass die Anzahl der Ressourcen einen größeren Einfluss auf den Zusammenhang von Akzeptanz und Einsatz hat als die Anzahl der Barrieren. Insbesondere die Bereitstellung von externen Unterstützungsangeboten sowie ein reger Austausch mit schulexternen Kooperationspartnern scheinen sich positiv auf den Zusammenhang von Akzeptanz und Einsatz auszuwirken. Eine Top-Down-Strategie geht nach den bisherigen Erkenntnissen eher mit einer höheren Verbindlichkeit der Materialnutzung, einem größeren schulinternen Unterstützungsangebot sowie einer geringeren Partizipation der Lehrkräfte bei der Materialentwicklung einher als eine Bottom-Up-Strategie.

Für die schulpraktische Umsetzung weisen diese Ergebnisse auf Möglichkeiten zur Unterstützung von Implementationsvorhaben hin. Insbesondere der Befund, dass die Bereitstellung von Ressourcen den Einsatz innovativer Maßnahmen wirksam unterstützt, stellt einen aussichtsreichen Ansatzpunkt dar. Zudem stellt die Kategorisierung der Lehreraussagen aus den Interviews eine empirische Grundlage zur Operationalisierung des Konstrukts der Machbarkeit im Rahmen der Implementationsforschung dar.

## **Openflux 1.0 – Entwicklung und Validierung eines frei verfügbaren, browserbasierten Instruments zur simulationsbasierten Diagnostik von Problemlösekompetenz**

**Alexander Christ<sup>1</sup>, Nicolas Becker<sup>2</sup>, Stephan Kröner<sup>1</sup>**

<sup>1</sup>Friedrich-Alexander-Universität Erlangen Nürnberg, Deutschland; <sup>2</sup>Universität des Saarlandes; [alexander.chrale.christ@fau.de](mailto:alexander.chrale.christ@fau.de)

Die Kompetenz zum Lösen komplexer Probleme wird als Schlüsselkompetenz für das 21. Jahrhunderts gehandelt. Sie wird als fächerübergreifende Kompetenz in PISA erhoben und gilt als mit fluider Intelligenz korreliert, aber nicht identisch. Weitere verwandte Konstrukte sind Selbstregulation und Arbeitsgedächtnis. In den meisten einschlägigen Studien wird die Problemlösekompetenz durch die Leistung bei Exploration und Steuerung eines unbekanntes Systems operationalisiert, wobei letztere mit Hilfe von Systemwissen oder basierend auf Informationen über die Systemzusammenhänge erfolgt. Nach anfangs aufgetretenen Problemen mit vorwissensbehafteten Szenarien werden in neuerer Zeit vor allem vorwissensneutrale Szenarien verwendet. Hier zeigte sich regelmäßig, dass eine Differenzierung von Problemlösekompetenz in Dimensionen wie „Exploration“, „Systemwissen“ und „Wissensanwendung“ den Daten besser gerecht wird als eine eindimensionale Betrachtung. Christ et al. (2017) haben erste Ergebnisse der Entwicklung des Problemlöseszenarios Openflux, einem frei verfügbaren, browserbasierten Problemlöseszenario nach dem Vorbild von Multiflux (Kröner, 2000) vorgestellt. Anliegen war neben der formativen Evaluation im Rahmen einer Cognitive-Lab-Studie die Replikation bisheriger, mit dem Instrument Multiflux gewonnener Erkenntnisse zum Zusammenhang von Intelligenz und Problemlösekompetenzmaßen (Kröner, Plass & Leutner, 2005) ebenso wie die Exploration von Zusammenhängen mit dem Problemlöse-Selbstkonzept als Indikator für die dispositionelle Selbstregulation, zu denen bei Rudolph et al. (2017) aufgrund geringer Reliabilitäten der dort verwendeten Selbstkonzeptskala keine Aussage getroffen werden konnte.

Die von Christ et al. (2017) berichtete Stichprobe war mit  $N = 72$  jedoch zu klein für die Berechnung differenzierter Strukturgleichungsmodelle. Für die vorliegende Arbeit wurde daher die Stichprobe auf  $N = 240$  Erwachsene (Alter:  $M = 25.87a$ ,  $SD = 3.71a$ ) erweitert, die Openflux gemeinsam mit dem Matrizenstest DESIGMA (Becker & Spinath, 2014) als Indikator für fluide Intelligenz und einem Fragebogen zum Problemlöse-Selbstkonzept bearbeiteten. Die Auswertung erfolgte über Strukturgleichungsmodelle. Es ergaben sich hohe interne Konsistenzen ( $\alpha > .80$ ;  $\omega_{1-3} > .90$ ;  $AVE > .60$ ) der Problemlösescores. Die CFA für das Problemlöseszenario mit den drei Subdimensionen Exploration, Systemwissen und Wissensanwendung wies eine sehr gute Passung auf ( $CFI = .99$ ;  $TLI = .98$ ;  $RMSEA = .05$ ;  $SRMR = .05$ ). Diese fiel zudem statistisch besser aus als für ein Modell, in dem die Problemlösekompetenz über einen Generalfaktor modelliert wurde ( $\chi^2 [3] = 197.11$ ;  $p = .00$ ). Für die Konstruktvalidität sprechen substantielle bivariate Korrelationen der Problemlösescores sowohl mit fluider Intelligenz ( $.29 < r < .51$ ) als auch mit dem Problemlöse-Selbstkonzept ( $.32 < r < .53$ ). Ein SEM mit Intelligenz und Problemlöse-Selbstkonzept als Prädiktoren der Problemlöseleistung mit Exploration, Systemwissen und Wissensanwendung als deren Kriterien (wobei zusätzlich jeder erstgenannte Problemlösescore Prädiktor für den danach genannten war) wies eine gute Passung auf ( $CFI = .97$ ;  $TLI = .96$ ;  $RMSEA = .07$ ;  $SRMR = .06$ ). Abgesehen von einem nichtsignifikanten Effekt der VOTAT-Explorationsstrategie auf das Systemwissen entsprach das Modell – bei hoher Varianzaufklärung in den Problemlösekompetenzmaßen – den theoretischen Erwartungen. Dabei ergab sich ein inkrementeller Beitrag des Problemlöse-

Selbstkonzepts für die Varianzaufklärung der drei Problemlösescores. Insgesamt zeigen sich damit vielversprechende Ergebnisse: Die Faktorstruktur entspricht den theoretischen Erwartungen und die erwartungskonformen Zusammenhänge mit Intelligenz und Problemlöse-Selbstkonzept sind erste Hinweise auf eine theoriekonforme Einbettung des Verfahrens in ein nomologisches Netzwerk. In Folgestudien sollten die Zusammenhänge mit einem breiteren Spektrum an Außenkriterien überprüft werden. Nach erfolgreicher Validierung ist dann ein dem Open Science-Gedanken entsprechendes Verfahren zur Diagnostik von Problemlösekompetenzen für die künftige Forschung auf diesem Gebiet verfügbar.

## **„Liebe Lehrkraft, du bist kompetent.“ Leistungsbeurteilung in zentralen Abschlussprüfungen im Spannungsfeld zwischen Autonomie und Kontrolle**

**Alexander Groß, Svenja Mareike Kühn**

Universität Koblenz-Landau, Deutschland; [gross@uni-koblenz.de](mailto:gross@uni-koblenz.de)

### Theoretischer Hintergrund

Um die Objektivität der Leistungsbeurteilung sowie der Notenvergabe zu steigern und Schulabschlüsse vergleichbarer zu machen, werden in zentralen Abschlussprüfungen von der zuständigen Behörde (Ministerien bzw. Landesinstitute) landesweit einheitliche Korrektur- und Bewertungsvorgaben ausgegeben, anhand derer Lehrkräfte eigenverantwortlich und schulin-tern die erbrachten Prüfungsleistungen beurteilen sollen (vgl. Klein et al., 2009). Welche Intentionen mit einheitlichen Korrektur- und Bewertungsvorgaben verbunden sind und welche Rolle die Schulaufsicht im Verhältnis zu eigenverantwortlich beurteilenden Lehrkräften einnimmt, ist Gegenstand dieses Posterbeitrags.

Als „Eingriffsaufsicht“ (van Ackeren, Klemm & Kühn, 2015, S. 104) handeln vorgesetzte Behörden traditionell mittels Anordnung, Weisung und Kontrolle (vgl. Huber, 2016, S. 42), wobei die Praxis der Schulaufsicht in den letzten Jahren durch die Stärkung der Eigenständigkeit von Einzelschulen sowie verschiedenen Maßnahmen zur Sicherung und Weiterentwicklung der Qualität schulischer Bildung vielfältigen Veränderungen unterworfen war (vgl. Bogumil 2017, S. 13). Nach Huber (2016, S. 42) fungiert die Schulaufsicht nicht mehr nur als Kontrollinstanz, sondern auch als Einrichtung, die Unterstützung, Förderung und Beratung leistet. Erste Studien zur Rolle der Schulaufsicht in landesweiten Qualitätssicherungs-/entwicklungsmaßnahmen weisen allerdings auf eine diffuse Rolle der Schulaufsicht im Spannungsfeld zwischen Schulautonomie und der Kontrolle schulischer Prozesse hin (vgl. Kuper et al., 2016 sowie Ramsteck & Maier, 2015 im Kontext von VERA). Im Forschungsfeld ‚zentrale Abschlussprüfungen‘ und speziell für den schulinternen Prozess der Leistungsbeurteilung sind Intentionen, Aktivitäten und Rolle der Schulaufsicht bisher nicht in den Blick genommen worden.

### Fragestellung

Im Posterbeitrag werden die Ergebnisse eines Forschungsprojektes vorgestellt, in dem die Perspektive der Schulaufsicht auf die Leistungsbeurteilung in zentralen Prüfungsverfahren zum Erwerb des Mittleren Schulabschlusses nach Klasse 10 untersucht wurde. Folgende Forschungsfragen sollen beantwortet werden:

Welche Intentionen werden von der Schulaufsicht mit der Einführung landesweit einheitlicher Korrektur- und Bewertungsvorgaben verfolgt?

Welche Rolle nimmt die Schulaufsicht hinsichtlich des Umgangs der Lehrkräfte mit den einheitlichen Korrektur- und Bewertungsvorgaben ein?

### Methode

In dieser Studie wurden leitfadengestützte Einzel- bzw. Gruppeninterviews mit den Vertreter/-innen der zuständigen Ministerien bzw. nachgeordneten Behörden durchgeführt, in deren Aufgabengebiet die Verantwortung für die zentralen Prüfungen zum Erwerb des Mittleren Schulabschlusses liegt. Insgesamt wurden 34 Personen aus allen deutschen Bundesländern interviewt, in denen es zentrale Abschlussprüfungen gibt. Der Interviewleitfaden umfasste verschiedene Themenblöcke zum Forschungsfeld ‚zentrale Abschlussprüfungen‘. Mit Blick auf die für diesen Beitrag relevanten Forschungsfragen wurde erfragt, welche Intentionen seitens der Schulaufsicht mit den Korrektur- und Bewertungsvorgaben verbunden sind und wie sich diese Intentionen in der Gestaltung der Vorgaben, aber auch in Unterstützungs- und Kontrollaktivitäten niederschlagen. Die Interviewdaten wurden inhaltsanalytisch ausgewertet. Grundlage dafür war ein deduktiv-induktiv entwickeltes Kategoriensystem (vgl. Kuckartz, 2012). Die Qualität des Kategoriensystems wurde über einen mehrstufigen Prozess der Erprobung, Diskussion, Ausdifferenzierung und Reduktion sichergestellt.

### Ergebnisse

Alle Befragten betonen, dass landesweit einheitliche Korrektur- und Bewertungsvorgaben unverzichtbarer Bestandteil in zentralen Prüfungsverfahren seien, um die mit der Einführung zentraler Prüfungen verbundenen Ziele (Transparenz, Objektivität, Vergleichbarkeit bei der Leistungsbeurteilung) zu erreichen. Aus dieser länderübergreifend formulierten Zielvorgabe werden gleichwohl unterschiedliche Maßnahmen zur Zielerreichung abgeleitet: Neben einer großen Bandbreite an Gestaltungsvarianten der Korrektur- und Bewertungsvorgaben (z.B. konkrete Musterlösungen mit eindeutiger Punktezuweisung und Punkte-Noten-Schlüssel vs. abstrakte Kriterien ohne Hinweise zur Notenvergabe) werden unterschiedliche Unterstützungsangebote für die eigenverantwortlich beurteilenden Lehrkräfte zum Umgang mit diesen Vorgaben gemacht (z.B. Korrekturberatungen, Fortbildungsveranstaltungen) sowie verschiedene Maßnahmen zur Kontrolle der Nutzung (z.B. verbindliche schulexterne Zweitkorrektur, Nachkontrollen durch die verantwortliche Behörde) durchgeführt. Einige Länder verzichten allerdings bewusst auf eine solche Unterstützung bzw. Kontrolle. Das Ausmaß des den beurteilenden Lehrkräften seitens der Schulaufsicht zugestandenen Grads an Autonomie sowie der Unterstützung und Kontrolle ist über unterschiedliche Umgangsweisen mit dem länderübergreifend formulierten Spannungsfeld zu erklären: Einerseits Objektivität und Vergleichbarkeit der Leistungsbeurteilung landesweit zu gewährleisten und gleichzeitig den Lehrkräften Beurteilungsfreiräume gewähren zu wollen, damit diese besser auf individuelle Schülerleistungen eingehen und den eigenen vorgelagerten Unterricht bei der Leistungsbeurteilung berücksichtigen können.

## **Examining the interplay of teachers' beliefs and motivational orientations for teaching gifted students**

**Svenja Matheis<sup>1</sup>, Leonie Kronborg<sup>2</sup>, Manfred Schmitt<sup>1</sup>, Franzis Preckel<sup>3</sup>**

<sup>1</sup>Universität Koblenz-Landau, Deutschland; <sup>2</sup>Monash University; <sup>3</sup>Universität Trier; [matheis@uni-landau.de](mailto:matheis@uni-landau.de)

We aim to make an empirical contribution to the discussion on the professionalization of teachers within gifted education by investigating preservice teachers' beliefs and motivational orientation as substantial aspects of their professional competences. Using an experimental vignette approach, we investigated the influence of student ability level on teachers' beliefs about these

students and on their motivational orientations for teaching them. Furthermore, it was examined how teachers' beliefs were related to their motivational orientations assuming motivation as precursor for actual behavior in classroom.

Motivational orientation as well as beliefs are substantial aspects of teachers' professional competences (Kunter et al., 2013; see also, Baumert & Kunter, 2006). Motivational orientation of teachers comprises cognitive components (i.e., self-efficacy) and affective components (i.e., enthusiasm) which can be related to actual classroom behavior (Kunter, Tsai, Klusmann, Brunner, Krauss, & Baumert, 2008; Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009). Teachers' beliefs comprise subjective theories of teaching and learning as well as subjective personality theories about students, which are assumed to play an important role for the interaction with students in school context (Voss, Kleickmann, Kunter, & Hachfeld, 2013). A case in point is teachers' personality theory about gifted students. Previous research revealed that teachers hold incorrect beliefs about gifted students combining high intellectual ability with social and behavioral difficulties (e.g., Preckel, Baudson, Krolak-Schwerdt, & Glock, 2015; Baudson & Preckel, 2013, 2016). In this study, we like to replicate those findings and to investigate the relation between teachers' beliefs and motivational orientation.

Using a between-subjects experimental vignette design, pre-service teachers (N=690) rated fictitious students varying in ability level (gifted/average) and gender (girl/boy). Pre-service teachers were randomly assigned to one of the four vignettes (girl/gifted, girl/average, boy/gifted, boy/average). We measured teachers' beliefs about the gifted and teachers' motivational orientation with the Teacher-Gifted-Questionnaire that assesses the dimensions of intellectual ability, social-emotional ability, maladjustment, teacher enthusiasm, and teacher self-efficacy for teaching students (Omega: .64 – .90).

After establishing scalar measurement invariance over vignettes, repeated-measures ANOVA (repeated measurement across five scales) with latent factor values showed that pre-service teachers considered gifted students superior regarding intellectual ability, but more maladjusted, compared to average-ability students. Whereas higher intellectual ability is in line with empirical findings about the gifted, higher maladjustment are not. Moreover, pre-service teachers felt less well prepared for teaching gifted students as they reported significantly lower self-efficacy for teaching the gifted as compared to average-ability students (no significant differences in enthusiasm).

For assessing the relation between students' ability level, teachers' beliefs, and motivational orientations, we used regression analysis within a multi-group structural equation model. The model with beliefs as predictors for teachers' motivational orientations was successful in explaining a significant amount of the variability of pre-service teachers' enthusiasm and self-efficacy. Ratings of intellectual ability and maladjustment significantly predicted pre-service teachers' motivational orientations. Their ratings of a student's intellectual ability had significantly positive regression weights, which indicates that pre-service teachers with high ratings of student intellect can be expected to exhibit a higher enthusiasm and higher self-efficacy for teaching this student. Rating of students' maladjustment, in turn, had significantly negative regression weights, which suggests that those pre-service teachers with higher ratings of maladjustment possessed a lower enthusiasm and self-efficacy. In summary, pre-service teachers' enthusiasm and self-efficacy is lowered by high maladjustment ratings, whereas high intelligence ratings had positive effects. In this manner, high intelligence did not appear to be detrimental to teachers' enthusiasm and self-efficacy, but its association with high maladjustment.

To conclude, we supposed that initial teacher training on giftedness should be devoted to both, beliefs about gifted students' characteristics and teachers' motivation.

### **Erweiterter Förderansatz zur Förderung selbstregulierten Lernens bei Vorschulkindern**

**Lisa Dörr, Sandra Dörrenbächer, Lisa Jacob, Franziska Perels**

Universität des Saarlandes, Deutschland; [lisa.doerr@mx.uni-saarland.de](mailto:lisa.doerr@mx.uni-saarland.de)

Ziel der vorliegenden DFG-geförderten Studie ist es, das selbstregulierte Lernen von Vorschulkindern durch eine Kombination von expliziten Maßnahmen zur Förderung des selbstregulierten Lernens und einem impliziten sprachbasierten Vorgehen zu unterstützen. Dazu sollen verschiedene Trainingsvarianten in Bezug auf ihre Effektivität evaluiert werden. Dabei meint Selbstregulation allgemein, dass Handlungsprozesse eigenständig initiiert, reguliert und reflektiert werden (Zimmerman, 1989, 2000). Wird diese Fähigkeit auf den akademischen Kontext übertragen, spricht man von selbstreguliertem Lernen (Landmann et al. 2015). Bei eher prozessbezogenen Konzeptionen wird selbstreguliertes Lernen definiert "[...] as the process by which a person generates thoughts, feelings and actions which are systematically oriented toward achieving one's goals" (Bembenutty & Karabenick, 2004). Diese Fähigkeit, das eigene Lernen selbständig, aktiv und strukturiert zu organisieren, wird insofern hoch eingeschätzt, als sie dem Lernenden erlaubt, sich lebenslang eigenständig Wissen anzueignen (vgl. Brunstein & Spörer, 2001; Zimmerman et al., 1996). Ein möglichst frühes Heranführen an Selbstregulationsstrategien fördert dabei die Etablierung günstigen Lernverhaltens von Anfang an und soll dazu beitragen, Kindern bereits im Vorschulalter wichtige Basisressourcen für künftige Lernerfahrungen zu vermitteln (Otto et al., 2011). In früheren Studien konnte gezeigt werden, dass positive Effekte auf das selbstregulierte Lernverhalten der Vorschulkinde erzielt werden konnten (Perels et al., 2009). Zusätzlich zu dieser bereits überprüften Methode der expliziten Förderung selbstregulierten Lernens soll als Erweiterung zusätzlich ein impliziter Ansatz zur Unterstützung des selbstregulierten Lernens verfolgt werden. Dazu wird das selbstregulierte Lernen durch gezielte Förderung der private speech (vgl. auch self-guided private speech, vgl. Foster et al., 1988) angeregt. Dieses Vorgehen scheint insofern vielversprechend, als im Vorschulalter die Selbstregulation aufgrund entwicklungspsychologischer Voraussetzungen noch eng an die Sprache bzw. Verbalisierung gekoppelt ist (Vygotsky, 1978; Bronson, 2000). Dabei wird angenommen, dass gerade das selbstregulierte (Lern-)Verhalten durch Selbstinstruktionen in Form der private speech gefördert wird: „Private Speech serves as a tool for thinking that facilitates problem solving and is also a tool that mediates people's self-regulatory capacity“ (Atencio 2008). Vor allem die dem selbstregulierten Lernen zugrundeliegenden metakognitiven Strategien bedürfen im Vorschulalter häufig noch einer Versprachlichung und werden erst schrittweise im Laufe der Entwicklung in innere Denkprozesse überführt. Daher kann angenommen werden, „dass zumindest ein Teil der kindlichen selbstbezogenen Sprache eine selbststeuernde Funktion hat und vom Kind als Hilfsmittel zur Strukturierung und Überwachung des eigenen Verhaltens genutzt wird“ (Dubowy, 2009).

Aufbauend auf diesen Erkenntnissen soll in der vorgestellten Studie mit einem 2 (explizite Förderung selbstregulierten Lernens ja/nein) x 2 (implizite Förderung durch Verbalisierung ja/nein)-Design überprüft werden, ob eine Kopplung der Maßnahmen wirksamer ist als ein reines Training selbstregulierten Lernens. Dabei erfolgt die Förderung in allen Bedingungen mittels einer Kombination aus direkter Interventionen mit den Kindern (10 Trainingssitzungen à 45 Minuten) und indirekter Interventionen durch ein Training der Erzieherinnen (ein Workshoptermin à 2 Stunden, Aushändigung eines „Handbuchs“ mit wichtigen Informationen, Beispielübungen und Anweisungen zur Umsetzung entsprechender Übungen im Kindergartenalltag sowie Fragen zu den Erfahrungen mit den Übungen im Sinne eines Manipulationchecks). Die eingesetzten Maßnahmen werden auf Ebene der Vorschulkinder als primäre Zielgruppe der Intervention evaluiert.

Zur Erfassung liegt ein multimethodaler Ansatz vor, da das selbstregulierte Lernen sowie die Sprach-/Verbalisierungsfähigkeit der Kinder mit Hilfe unterschiedlicher Erhebungsverfahren bzw. Informationsquellen erfasst wird (vgl. Seidenstücker & Baumann, 1978). So kommen zum einen ein handlungsnahes, kindgerechtes Instrument zur Erfassung der frühkindlichen Selbstregulationsfähigkeit (Handlungssequenztest), sowie eine Ratingskala für Erzieher als Beurteilungsinstrument für das selbstregulierte Lernen der Vorschulkinder zum Einsatz (vgl. z. B. Whitebread et al., 2009). Die Sprachfähigkeit der Vorschulkinder wird über die allgemeine Sprachskala des HAWIVA-III (Ricken et al., 2007) erfasst. Im Oktober dieses Jahres soll die Studie in etwa 15 Kindertageseinrichtungen im Saarland mit rund 220 Vorschulkindern durchgeführt werden. Die bis zum Zeitraum der Tagung gewonnenen Ergebnisse möchten wir mit dem Fachpublikum diskutieren.

## **Haben witzige Lehrkräfte motiviertere Schülerinnen und Schüler? Die Wirkung von Lehrkrachthumor auf die Interessantheit von Unterricht und intrinsische Motivation**

**Sonja Bieg<sup>1</sup>, Robert Grassinger<sup>2</sup>**

<sup>1</sup>Universität Augsburg, Deutschland; <sup>2</sup>Pädagogische Hochschule Weingarten; [robert.grassinger@phil.uni-augsburg.de](mailto:robert.grassinger@phil.uni-augsburg.de)

### Theoretischer Hintergrund

Unter Humor wird die Kommunikation von inkongruenten Bedeutungen verstanden, die vom Rezipienten wahrgenommen und aufgelöst werden müssen und dann zu Lachen führen (Martin, 2007). Diese Inkongruenz von Informationen vermag situationales Interesse zu fördern (z.B. Bertyne, 1963; Hidi & Renninger, 2006) und sich somit positiv auf die Unterrichtsqualität auswirken (z.B. Banas, Dunbar, Rodriguez & Liu, 2011). Empirische Evidenz für diese Argumentation erbrachte eine Studie von Dresel et al. (2014), in welcher der Lehrkrachthumor moderat bis stark mit der Interessantheit des Unterrichts assoziiert war. Allerdings wurde in der Studie Humor als ein eindimensionales Konstrukt operationalisiert, in der Literatur jedoch verschiedene Humorformen unterschieden (Bieg & Dresel, 2016; Frymier, Wanzer, & Wojtaszczyk, 2008; Wanzer & Frymier, 1999; Wanzer, Frymier, Wojtaszczyk, & Smith, 2006; Wanzer, Frymier & Irwin, 2010). So differenzieren Bieg und Dresel (2016) zwischen lerngegenstandsbezogenem Humor, Humor ohne Bezug zum Lerngegenstand, selbstabwertendem Humor und aggressivem Humor und fanden bei dieser querschnittlich angelegten Studie differenzielle Zusammenhänge der Lehrkrachthumorformen mit der Interessantheit des Unterrichts. Nicht untersucht ist, inwieweit sich dieser Zusammenhang kausal interpretieren lässt und Humor von Lehrkräften als instruktionales Element zur Förderung der Interessantheit von Unterricht angesehen werden kann. Jedoch verweisen querschnittlich angelegte empirische Arbeiten, dass Lehrkrachthumor mediiert über die Interessantheit des Unterrichts, einen Zusammenhang zur intrinsischen Motivation der Lernenden zeigt (Bieg & Dresel, 2013; Wanzer et al., 2010). Zu untersuchen inwieweit die verschiedenen Formen des Lehrkrachthumors unter Berücksichtigung der Interessantheit des Unterrichts, positiv auf die Motivation der Schülerinnen und Schüler wirken, ist Ziel dieses Beitrags. Dies erscheint plausibel, da ein interessanter Unterricht mit höherer Lernmotivation der Schülerinnen und Schüler einhergeht.

### Hypothesen

Entsprechend zielt der Beitrag darauf ab, folgende zwei Hypothesen zu prüfen: (1) Je mehr lerngegenstandsbezogenen Humor Lehrkräfte im Unterricht verwenden, desto interessanter erleben Schülerinnen und Schüler den Unterricht und desto höher ist deren intrinsische Motivation. (2) Der Zusammenhang zwischen lerngegenstandsbezogenem Lehrkrachthumor und intrinsischer Schülermotivation wird vermittelt über die Interessantheit des Unterrichts.

### Methodik

Zur Prüfung der Hypothesen wurden 609 Schülerinnen und Schüler aus 38 Klassen der Jahrgangsstufen 5 bis 10 zweimal in einem Schuljahr über Online-Fragebögen befragt. Zu Beginn des Schuljahres schätzten sie den Humor ihrer Lehrkraft ein und zur Mitte des Schuljahres gaben sie Auskunft über die Interessantheit des Unterrichts und ihre intrinsische Motivation. Fehlende Werte wurden mit Full Information Maximum Likelihood (FIML) geschätzt. Die Prüfung der postulierten Zusammenhänge erfolgte mit einem 2-Ebenen Strukturgleichungsmodell in Mplus Version 6.

### Ergebnisse

Die Lernenden erlebten den Unterricht zum zweiten Messzeitpunkt umso interessanter, je mehr ihre Lehrkräfte zu Beginn des Schuljahres lerngegenstandsbezogenen Humor einsetzten ( $\beta = 0.82$ , S.E. 0.04) und je weniger diese Humor ohne Bezug zum Lerngegenstand ( $\beta = -0.17$ , S.E. 0.06) verwendeten. Zudem wirkten diese beiden Humorformen vermittelt über die Interessantheit des Unterrichts indirekt auf die intrinsische Motivation (lerngegenstandsbezogener Humor  $\beta = 0.30$ , S.E. 0.07, Humor ohne Bezug zum Lerngegenstand  $\beta = -0.06$ , S.E. 0.03). Selbstabwertender Humor sowie aggressiver Humor hingegen erwiesen sich weder direkt noch indirekt als bedeutsam für die Interessantheit des Unterrichts und die intrinsische Motivation der Schülerinnen und Schüler.

Diese längsschnittlichen Ergebnisse zeigen, dass durch gezielt eingesetzten Humor die Unterrichtsqualität im Sinne von Interessantheit erhöht werden kann und sich die intrinsische Motivation der Lernenden dadurch erhöhen lässt. Lehrkräfte sollten in ihre Unterrichtsplanungen Überlegungen einbeziehen, an welchen Stellen lerngegenstandsbezogener Humor den Unterrichtsinhalt bereichern könnte.

## **Selektive Zusammenhänge zwischen selbstregulativer Planungsfähigkeit und proaktiver Kontrolle im Vorschulalter**

**Sandra Dörrenbächer, Efsevia Kapsali, Lisa Dörr, Lisa Jacob, Franziska Perels**

Universität des Saarlandes, Deutschland; [sandra.doerrenbaecher@uni-saarland.de](mailto:sandra.doerrenbaecher@uni-saarland.de), [sevi\\_kapsali@hotmail.com](mailto:sevi_kapsali@hotmail.com)

Ein Erfolgsfaktor beim Übergang vom offenen Kindergartenalltag zu engeren schulischen Regelstrukturen ist die Ausreifung planvoller, strategischer und selbstgesteuerter Handlungsformen. Diese Fähigkeit zur Selbstregulation erlaubt es Kindern im Vorschulalter, neue komplexe Aufgabenanforderungen in präaktionale (Planungs-), aktionale (Ausführungs-) und postaktionale (Reflexions-) Anteile zu untergliedern, in denen phasenoptimierte Handlungsstrategien eingesetzt werden können (Zimmerman, 2000). Eine ähnliche sequenzielle Organisation wird für exekutive Kontrollmechanismen angenommen (Banich, 2009), die eine Optimierung grundlegender Prozesse der Handlungsvorbereitung, -überwachung und -korrektur ermöglichen. Tatsächlich werden exekutive Kontrollprozesse als Voraussetzung der Implementierung selbstregulativer Strategien diskutiert (z. B. Hofmann et al., 2012; aber siehe auch beispielsweise Diamond (2013) für alternative Interpretationen der Konstruktzusammenhänge). Innerhalb der Selbstregulation kommt planungsbezogenen Strategien ein besonderer Stellenwert zu, da „alle Parameter der präaktionalen Phase (...) Einfluss nehmen auf die weiteren Phasen“ (Schmitz & Schmidt, 2007, S. 14). Ähnlich stellt der Einsatz proaktiver, antizipativer exekutiver Kontrolle, also einer früh initiierten, top-down gesteuerten

Selektion zielrelevanter Prozesse, in herausfordernden Situationen einen relativen Vorteil gegenüber reaktiver Kontrolle, also einer späten, bottom-up getriggerten Handlungskorrektur, dar (Braver & Barch, 2002). Das Vorschulalter erweist sich hierbei als kritische Phase sowohl der Entwicklung selbstregulativer Planungsfähigkeiten (Bronson, 2000) als auch proaktiver Kontrollmechanismen (Lucenet & Blaye, 2014), jedoch mit einer hohen interindividuellen Streubreite der Entwicklungstrajektorien (von Suchodoletz et al., 2014). Ziel der vorliegenden Arbeit war es daher, im Rahmen eines Pilotprojekts zu untersuchen, ob solche interindividuellen Unterschiede in proaktiver exekutiver Kontrolle im Vorschulalter selektiv mit interindividuellen Unterschieden in selbstregulativem Strategiewissen zur Handlungsvorbereitung, jedoch weniger zur Handlungsausführung, zusammenhängen.

Die Pilotstichprobe bestand aus n=15 Vorschulkindern im Alter von 5 bis 6 Jahren.

Zur direkten Erfassung phasenspezifischen selbstregulativen Strategiewissens auf Kindebene wurde ein Selbstregulations-Strategietest für das Vorschulalter entwickelt (ähnlich zu Lockl et al., 2016). Hierzu wurden, eingebettet in eine kindgerechte Bildergeschichte, verschiedene Problemstellungen entlang der Phasen des Selbstregulationsprozesses präsentiert, zu deren Lösung jeweils vier alternative Strategiehandlungen zur Verfügung gestellt wurden. Das Kind sollte jede Alternative auf einer dichotomen Skala als „toll“ oder „schlecht“ bewerten. Zwei der dargebotenen Antwort-Alternativen beinhalteten eine strategisch sinnvolle Lösung des Problems (Target-Lösungen), zwei bildeten Distraktoren. Aus der Differenz der Gesamtzahl richtiger und falscher Antworten (unter Berücksichtigung korrekter Ablehnungen und falscher Alarme) wurden phasen- und strategie-spezifische Selbstregulations-Indizes berechnet.

Zur Extraktion proaktiver versus reaktiver exekutiver Kontrollstile wurde eine computergestützte, für das Vorschulalter adaptierte Variante der AX-continuous-performance-Task mit Bild-Stimuli herangezogen (AX-CPT; vgl. Doebel et al., 2016). Bei dieser Aufgabe wurden Cue-Probe-Bild-Paarungen präsentiert und die Kinder instruiert, auf eine valide Cue-Probe-Paarung (Target-Sequenz) mit einer bestimmten Zielreize Taste zu antworten, während bei Paarungen mit invalidem Cue oder invalidem Probe eine andere Taste gedrückt werden musste. Die Targetsequenz wurde in 70% der Fälle präsentiert, um einen proaktiven Reaktionsbias zu induzieren, d. h. die antizipative Nutzung der Cue-Information ermöglichte eine effektivere Aufgabebearbeitung. Relativ höhere Antwortlatenzen in Durchgängen mit a) validem Cue und invalidem Probe indizieren den Einsatz proaktiver Kontrolle, b) invalidem Cue und validem Probe indizieren den Einsatz reaktiver Kontrolle.

Erste korrelative Analysen auf Basis nonparametrischer Spearman-Rangkorrelationen ergaben erwartungsgemäß einen hochsignifikanten positiven Zusammenhang zwischen der korrekten Identifizierung präaktionaler Selbstregulationsstrategien im Strategiewissenstest und dem Proaktivitätsindex in der AX-CPT ( $\rho=.75$ ,  $p<.001$ ). Kinder mit besseren Planungsfähigkeiten machten also auch stärker Gebrauch von proaktiven exekutiven Ressourcen. Ergänzend zeigte sich auf Ebene einzelner Strategien eine starke monotone Beziehung zwischen der selbstregulativen Strategie „Planung und Aktivierung von Vorwissen“ und Proaktivität ( $\rho=.91$ ,  $p<.001$ ). Im Sinne diskriminanter Validität korrelierte der Proaktivitätsindex nicht substantiell mit aktionalem Strategiewissen ( $\rho=.11$ ), und der Reaktivitätsindex mit keinem der Selbstregulationsindizes (alle  $p's>.14$ ). Zusammenfassend liefern die Befunde, wenn auch noch mit geringer Teststärke, erste Hinweise auf eine selektive Überlappung von Planungsfähigkeiten und proaktiven Kontrollprozessen im Vorschulalter. Diese geteilte Varianz in elementaren Basisressourcen gilt es für die weitere Bildungsentwicklung der Kinder auszuschöpfen und zum Gegenstand schulvorbereitender kombinierter Fördermaßnahmen zu machen.

## **Die Effekte einer Intervention zur Autonomieförderung auf die Überzeugungen sowie das Professionswissen von Lehramtsstudierenden im Fach Biologie**

**Nadine Großmann, Stefan Fries, Matthias Wilde**

Universität Bielefeld, Deutschland; [ngrossmann@uni-bielefeld.de](mailto:ngrossmann@uni-bielefeld.de)

### Theoretischer Hintergrund und Forschungsstand

Das Abnehmen zentraler motivationaler Variablen über die Schullaufbahn wurde vielfach gezeigt (z.B. Jacobs et al., 2002). Lehrpersonen nehmen diese Abnahme wahr, jedoch fehlen ihnen oftmals die didaktisch-methodischen Fertigkeiten, um mit Motivierungsproblemen umzugehen (Reeve et al., 2004). Die bereits evaluierten Konzepte zur Motivationsförderung scheinen ihren Weg in die Praxis bislang nicht zu finden. Ein prominentes Beispiel ist die Autonomieförderung im Sinne der Selbstbestimmungstheorie (Reeve, 2002; Ryan & Deci, 2017).

In Interventionsstudien hat sich gezeigt, dass motivationsförderliche Verhaltensweisen im Unterricht formbar sind, wenn Lehrpersonen die Relevanz, die einfache Umsetzung und die Effektivität dieser Maßnahmen erkennen (Aelterman et al., 2014; Reeve & Cheon, 2016). Es gibt zudem Hinweise darauf, dass derartige Interventionen bereits in der Lehrerbildung eingesetzt werden sollten (Su & Reeve, 2011). Lehramtsstudierende sind noch flexibel in ihrem Lehrerverhalten und greifen im Unterricht auf motivational ungünstige kontrollierende Verhaltensweisen zurück (Martinek, 2010; Tessier et al., 2010). Autonomieförderliche Maßnahmen können die motivationale Qualität von Schülerinnen und Schülern im Unterricht positiv beeinflussen (z.B. Hofferber et al., 2017; Reeve, 2002). In der vorliegenden Studie wurde daher auf Grundlage der Selbstbestimmungstheorie eine Maßnahme für die erste Phase der Lehrerbildung konzipiert und pilotiert, die den Studierenden Kenntnisse und Fertigkeiten zur Autonomieförderung vermittelt.

### Fragestellungen

- 1) Ist die Intervention lernwirksam bezüglich des Theorie- und Handlungswissens der Lehramtsstudierenden?
- 2) Kann die Intervention die Überzeugungen der Lehramtsstudierenden bezüglich der einfachen Implementation und der Effektivität autonomieförderlicher Maßnahmen beeinflussen?
- 3) Kann die Intervention die zukünftigen Intentionen der Lehramtsstudierenden, autonomieförderliche Maßnahmen einzusetzen, beeinflussen?

### Methode

An der Studie nahmen im Wintersemester 2016/17 sowie im Sommersemester 2017 58 Studierende des Lehramts Biologie (MAlter=25.18±3.79 Jahre; MSemester=7.78±1.23; 65% weiblich) teil. 35 Studierende erhielten eine Intervention zur Autonomieförderung im Unterricht (Experimentalgruppe), während 23 Studierende die Kontrollgruppe ohne diese Intervention bildeten. Die Intervention war in Vorbereitungsseminare zum Praxissemester eingebunden und umfasste eine Theoriesitzung sowie zwei Trainingssitzungen. In der Theoriesitzung wurden die in der Selbstbestimmungstheorie verankerten psychologischen Grundbedürfnisse sowie das Motivationskontinuum der Internalisierung extrinsischer Motivation vermittelt. In den Trainingssitzungen wurden autonomieförderliche Maßnahmen theoretisch fundiert und am Beispiel des Biologieunterrichts geübt.

Die Wirksamkeit der Intervention wurde anhand dreier abhängiger Variablen überprüft. Zur Erfassung des Theorie- und Handlungswissens zur Autonomieförderung wurden 15 Wissensitems mit offenem Antwortformat erstellt (Cronbachs  $\alpha$ =.67-.81; Cohens  $\kappa$ =.91-.93). Zudem wurden die Überzeugungen der Studierenden bezüglich der Implementierbarkeit und der Effektivität autonomieförderlicher Maßnahmen sowie ihre Intention zum zukünftigen Einsatz der Maßnahmen erfasst. Hierfür wurde eine übersetzte Version des Teaching Scenarios Measure (Reeve et al., 2014; Cronbachs  $\alpha$ =.73-.94) genutzt. Alle Instrumente wurden im Prä- sowie Posttest eingesetzt. Die statistische Auswertung erfolgte anhand von Varianzanalysen mit Messwiederholung.

#### Ergebnisse

Berichtet werden die Interaktionseffekte der Faktoren Zeit und Treatment. Im Vergleich zur Kontrollgruppe erreichten die Studierenden der Experimentalgruppe höhere Punktzahlen im abschließenden Theorie- und Handlungswissenstest ( $F(1,56)=40.19, p=.000, \eta^2=.42$ ;  $F(1,56)=82.73, p=.000, \eta^2=.60$ ). Zudem erachteten die Studierenden der Experimentalgruppe die autonomieförderlichen Maßnahmen nach der Intervention als leichter umsetzbar ( $F(1,56)=16.83, p=.000, \eta^2=.23$ ) und effektiver ( $F(1,56)=3.17, p=.081, \eta^2=.05$ ). Die Intention zum zukünftigen Einsatz der vermittelten Maßnahmen war in der Experimentalgruppe nach der Intervention höher als in der Kontrollgruppe ausgeprägt ( $F(1,56)=7.01, p=.011, \eta^2=.11$ ).

Die Ergebnisse dieser ersten Pilotierung zeigen, dass die konzipierte Maßnahme das Wissen, die Überzeugungen sowie die Intentionen der Studierenden, eine Autonomieförderung im Biologieunterricht umzusetzen, positiv beeinflussen kann. Um die Wirksamkeit der Maßnahme weiter zu überprüfen, finden im laufenden Wintersemester die Follow-Up Erhebungen der beschriebenen Studie sowie weitere Erhebungen im Rahmen einer Folgestudie statt. Diese Ergebnisse werden auf der Tagung ebenfalls präsentiert.

In Zukunft sollte untersucht werden, ob und wie sich die Intervention auf das Unterrichtsverhalten der Studierenden auswirkt und welche motivationalen Effekte es schülerseitig erzielt. Ebenfalls stellt sich die Frage, ob die Maßnahme auch im Rahmen von Lehrerfortbildungen zur Anwendung kommen kann.

### **Gemeinsam lernen, miteinander teilen - Beeinflusst eine inklusive Beschulung das prosoziale Verhalten von SchülerInnen ohne sonderpädagogischen Förderbedarf?**

**Amelie Hacker, Lena Nusser**

Leibniz-Institut für Bildungsverläufe, Deutschland; [amelie.hacker@lifbi.de](mailto:amelie.hacker@lifbi.de)

Durch die Ratifizierung der UN-Behindertenrechtskonvention hat sich Deutschland 2009 zu einem inklusiven Schulsystem verpflichtet, in dem „alle[n] denkbaren Adressaten“ zu einer besseren Entwicklung verholfen werden soll (Tenorth, 2013, S. 13). Daher werden SchülerInnen mit und ohne Förderbedarf in inklusiven Bildungssettings gemeinsam und nicht mehr nach personenbezogenen Merkmalen differenziert in verschiedenen Lerngruppen unterrichtet (Tenorth, 2013; Ainscow, 1991; Lipsky & Gartner, 1999). Folglich besteht ein hohes Anliegen darin, die Entwicklung von SchülerInnen in inklusiven Settings auch empirisch zu untersuchen. Allerdings liegt dabei der Fokus vornehmlich auf der Entwicklung von SchülerInnen, die sonderpädagogisch gefördert werden. Die Entwicklung ihrer MitschülerInnen hingegen wird, wenn überhaupt, überwiegend in Bezug auf ihre akademischen Leistungen beleuchtet, weshalb bisher nicht fundiert belegt ist, wie sie sich innerhalb ihres Schulkontextes emotional und sozial entwickeln (Avramidis, 2009, S. 14; Wild, Schwinger, Lütje-Klose, Yotodying & Gorges, 2015). Um diese Forschungslücke näher zu beleuchten, wird in dem Posterbeitrag die Entwicklung des Sozialverhaltens von SchülerInnen ohne sonderpädagogischen Förderbedarf (SPF), die gemeinsam mit SchülerInnen mit SPF unterrichtet werden, betrachtet.

Da Kinder mit SPF häufig dysfunktionale Verhaltensweisen aufzeigen (Heimlich, Kahlert, Lelgemann & Fischer, 2016), werden besonders aus Elternperspektive Befürchtungen geäußert, dass durch eine gemeinsame Beschulung von Kindern mit und ohne Förderbedarf die emotional-soziale Entwicklung ihrer Kinder ohne SPF negativ beeinflusst wird und sie entsprechende Verhaltensweisen übernehmen könnten (Müller, 2008, S. 67-69). Folglich wird aufgrund des Einflusses der Peergroup und insbesondere der SchülerInnen mit SPF angenommen, dass sich das Sozialverhalten der SchülerInnen ohne SPF verschlechtern könnte (Prinstein & Wang, 2005). Weiter wird davon ausgegangen, dass ein schlechtes Sozialverhalten mit einer reduzierten Aufmerksamkeit gegenüber dem Unterrichtsgeschehen einhergeht, was letztlich auch negative Auswirkungen auf die schulischen Leistungen von SchülerInnen ohne SPF, die gemeinsam mit SchülerInnen mit SPF unterrichtet werden, haben könnte (Wischnowski et al., 2004). In Anlehnung an die Kontakttheorie (Allport, 1954) kann allerdings dagegen argumentiert werden, dass der Kontakt zwischen Personengruppen positive Auswirkungen auf Verhaltensweisen und Einstellungen gegenüber „anderen“ Menschen hat. In Bezug auf inklusive Beschulungssettings kann daher angenommen werden, dass sich häufiger Kontakt zwischen SchülerInnen mit und ohne Förderbedarf und dessen Intensität entgegen der elterlichen Sorge positiv auf die Einstellungen und Verhaltensweisen von SchülerInnen ohne SPF gegenüber ihren MitschülerInnen mit SPF auswirkt. Entsprechend ist zu erwarten, dass sie im Gegensatz zu SchülerInnen, die nicht in inklusiven Settings unterrichtet werden, verstärktes prosoziales Verhalten aufzeigen, indem sie stärker für die Bedürfnisse und Anforderungen von Menschen mit Beeinträchtigung sensibilisiert werden (Himes, 1960; Maras & Brown, 2000; Brinker, 1985; Scheepstra, Nakken & Pijl, 1999). Dabei scheinen strukturiert-kooperierende Unterrichtsformen positive Auswirkungen auf das prosoziale Verhalten zu haben (Acton & Zarbatany, 1988), während die Anzahl der integrierten SchülerInnen mit SPF das Sozialverhalten negativ beeinflusst.

Da der Forschungsstand diesbezüglich bisher heterogene Ergebnisse zeigt, gilt es zu untersuchen, ob SchülerInnen ohne SPF, die in inklusiven Settings unterrichtet werden, höheres prosoziales Verhalten zeigen als SchülerInnen ohne SPF, die nicht inklusiv beschult werden. Zur Analyse der Forschungsfrage werden Daten des Nationalen Bildungspanels (NEPS; Blossfeld, Roßbach & von Maurice, 2011) herangezogen. Hierbei werden die subjektiven Einschätzungen der SchülerInnen ohne SPF ( $n=1119$  in 78 inklusiven und 28 nicht-inkluisiven Klassen) zu ihrem prosozialem Verhalten in der sechsten Jahrgangsstufe mehrbenenanalytisch unter Kontrolle der Unterrichtsgestaltung und der Klassenzusammensetzung innerhalb der inklusiven Settings analysiert. Erste Ergebnisse zeigen, dass SchülerInnen ohne SPF, die gemeinsam mit SchülerInnen mit SPF unterrichtet werden, höhere prosoziale Verhaltensweisen aufzeigen, als diejenigen, die nicht in inklusiven Settings unterrichtet werden. Ebenso zeigt sich, dass sich strukturiert-kooperierende Unterrichtsformen in inklusiven Settings positiv auf das Sozialverhalten der SchülerInnen ohne SPF auswirken. Weitere Analysen sollen zeigen, ob dieser Zusammenhang von der Anzahl der SchülerInnen mit SPF in einer Klasse moderiert wird.

### **Je freier, desto mehr Beziehungsziele? Effekte von ersten Praxisphasen auf das Bedürfniserleben von Lehramtsstudierenden und die Folgen für Lern- und Beziehungszielorientierungen**



**Lia Oberhauser, Silke Hertel**

Universität Heidelberg, Deutschland; [oberhauser@heiedu.uni-heidelberg.de](mailto:oberhauser@heiedu.uni-heidelberg.de)

Frühe Praxisphasen im Lehramtsstudium (Orientierungspraktika von drei Wochen Dauer) sollen Studierende in Ihrer Berufsorientierung unterstützen und können damit die motivationale Grundlage für deren Professionalisierung legen. Motivation ist Teil der professionellen Handlungskompetenz von Lehrkräften (Baumert & Kunter, 2011) und umfasst verschiedene motivationale Konstrukte wie zum Beispiel Zielorientierungen. Als habituelle Präferenz für eine bestimmte Art von Zielen sind insbesondere Lernzielorientierungen und Beziehungszielorientierungen für das lebenslange Lernen und die Unterrichtsqualität von Lehrkräften nachweislich von Bedeutung: Eine starke Lernzielorientierung zeigt Zusammenhänge mit dem Lesen von Fachzeitschriften und dem Besuch von Weiterbildungen (Nitsche, Dickhäuser, Dresel, & Fasching, 2013), der Annahme von Hilfe bei Problemen (Dickhäuser, Butler, & Tönjes, 2007) und wenig emotionaler Erschöpfung (Tönjes & Dickhäuser, 2009); Beziehungsziele zeigen Zusammenhänge mit konstruktiv unterstützendem Unterricht (Butler & Shibaz, 2014). Zielorientierungen gelten zwar als relativ stabil, Lernzielorientierungen weisen jedoch Schwankungen auf (vgl. Fasching, Dresel, Dickhäuser, & Nitsche, 2010; Malmberg, 2008; Praetorius et al., 2014; Senko & Harackiewicz, 2005) und zeigen einen Zusammenhang mit der Befriedigung grundlegender psychologischer Bedürfnisse wie Autonomie, Kompetenz und Zugehörigkeit (Janke, Nitsche, & Dickhäuser, 2015). Längere Praxisphasen im Studium tragen eher zu einem Anstieg der Lernzielorientierung bei (Malmberg, 2008), im Vorbereitungsdienst kommt es eher zu einer Abnahme (Fasching et al., 2010). Hingegen sind Effekte früher und kurzer Praxisphasen, in denen es zum ersten Kontakt mit beruflichen Anforderungen kommt, kaum befocht. Die vorliegende Studie nimmt kurze Praxisphasen in den Blick und betrachtet, ob sich Lern- und Beziehungszielorientierungen durch eine solche Praxisphase und abhängig vom Bedürfniserleben der Studierenden in dieser Praxisphase verändern. Studien lassen vermuten, dass die Bedürfnisseinschränkung bei ersten Unterrichtserfahrungen von Studierenden erhöht ist (Evelein, Korhagen, & Brekelmans, 2008). Nach dem Two Process Model (Sheldon, 2011) könnte dies jedoch adaptiv zu einem Anstieg von Zielorientierungen führen. Untersucht wurden daher die Hypothesen, dass es 1) zu einem Anstieg von Beziehungszielorientierung und Lernzielorientierung durch die Praxisphase kommt und dass 2) dieser Anstieg durch Bedürfnisseinschränkung im Praktikum vorhergesagt werden kann. Zur Prüfung der Hypothesen wurden anhand eines quasi-experimentellen Designs mit zwei Messzeitpunkten Fragebogendaten von N=52 Studierenden vor und nach dem Praktikum erhoben. Zur Operationalisierung der Zielorientierungen dienten die Fragebögen von Nitsche, Dickhäuser, Fasching, und Dresel (2011) und Butler und Shibaz (2014), zur Operationalisierung des Bedürfniserlebens die Skala von Neubauer und Voss (2016). Robuste lineare Modelle (Field & Wilcox, 2017) ließen einen Anstieg der Lernzielorientierungen ( $F_t=6.12$ ,  $FCrit=3.97$ ,  $p<0.05$ ) und der Beziehungszielorientierungen ( $F_t=14.06$ ,  $FCrit=4.14$ ,  $p<0.05$ ) vermuten, Bedürfnisseinschränkung hingegen schien mit Ausnahme der Zugehörigkeitseinschränkung im Praktikum erwartungswidrig eher geringer ausgeprägt zu sein als davor ( $p<0.05$  für alle Komponenten außer Kompetenzbefriedigung). Zusätzliche Varianz in den Beziehungszielwerten zum Messzeitpunkt 2 über die Beziehungszielwerte zum Messzeitpunkt 1 hinaus konnte durch niedrige Autonomieeinschränkung im Praktikum erklärt werden ( $b=-0.19$ ,  $p<.001$ ,  $\Delta R^2=.16$ ). Studierende fühlen sich in einem kurzen frühen Praktikum in ihren Bedürfnissen eher wenig eingeschränkt und werden eher darin bestärkt, Lernziele und Beziehungsziele zu verfolgen. Grenzen der Studie sind die Stichprobengröße, die fehlende Kontrollgruppe, nichtkontrollierbare Einflussfaktoren im Praktikum sowie die mangelnde Generalisierbarkeit aufgrund der Beschränkung auf einen Hochschulstandort. Dennoch liefert sie erste wichtige Hinweise auf die motivationalen Effekte kurzer Praxisphasen und die mögliche Bedeutung von Autonomie für die Beziehungszielorientierung.

### **Dialogisches Lehren im Unterricht: Qualitätskriterien zur Implementation einer Lehrerfortbildung aus Schülersicht**

**Ann-Kathrin Schindler<sup>1</sup>, Ricardo Boenheim<sup>1</sup>, Tina Seidel<sup>1</sup>, Maralena Pielmeier<sup>1</sup>, Alexander Gröschner<sup>2</sup>**

<sup>1</sup>Technische Universität München, Deutschland; <sup>2</sup>Friedrich-Schiller-Universität Jena, Deutschland; [ann-kathrin.schindler@tum.de](mailto:ann-kathrin.schindler@tum.de)

#### Theoretischer Hintergrund

Das Unterrichtsgespräch und seine Veränderbarkeit sind Gegenstand der aktuellen Unterrichtsqualitätsforschung (Drechsel & Schindler, in Druck). Walshaw und Anthony (2008) betonen die Wichtigkeit, das Unterrichtsgespräch schüleraktivierend zu gestalten, indem (1) Lernende Regeln und Verantwortlichkeiten für die Teilnahme kennen. Schüler/innen wissen um die Bedeutung des aktiven Zuhörens sowie die Anforderung der eigenen verbalen Beteiligung. Sind die Schüler/innen mit den Erwartungen vertraut, muss ihnen im Sinne von Aktivierung (2) die Möglichkeit für eigene Beiträge und Ideen gegeben werden. Verschiedene Schüler/innen können sich aktiv mit Beiträgen beteiligen und gelten als gleichberechtigte Gesprächspartner. Offene Fragen (Oliveira, 2010) und ein lernprozessbegleitendes Feedback (Hattie & Timperely, 2007) gelten dabei als wichtige Aktivierungs- und Unterstützungselemente der Lehrkraft. Andere Forschergruppen betonen zudem, dass eine gute Strukturierung und Orientierung an einem Gesprächsziel als Grundlage für das schüleraktivierende Unterrichtsgespräch gegeben sein muss (Resnick et al., 2008).

Diese Forderung zur Schüleraktivierung konfundiert mit Ergebnissen, dass Lehrpersonen die aufgezeigten Qualitätskriterien nicht notwendigerweise in ihrem Klassenzimmer implementiert haben. Schüler/innen nehmen in Gesprächssituationen z.B. häufig die Rolle von Stichwortgebern ein (Jurik et al., 2013) und wissen unzureichend über Lernziele im Unterrichtsgespräch (Seidel, Rimmelme & Prenzel, 2005).

Die aufgezeigte Diskrepanz zwischen Ist- und Soll-Zustand lässt das Thema Unterrichtsgespräch in den Fokus der Fortbildungslandschaft rücken (z.B. Reznitskaya et al., 2016). Die von Schüler/innen wahrgenommene Unterrichtsqualität ist in Untersuchung entsprechender Angebote ein zentrales Anliegen (Pehmer et al., 2015). Um die von Schüler/innen wahrgenommene Implementation des schüleraktivierenden Unterrichtsgesprächs zu verstehen, bedarf es einem dafür ausgerichteten Erhebungsinstrument. Folgende Fragestellung wird diesbezüglich verfolgt:

Inwiefern lässt sich ein Fragebogeninstrument entwickeln, das zentrale Faktoren von Schüleraktivierung im Unterrichtsgespräch reliabel und valide erfasst?

#### Methode

Aus der dargestellten Theorie wurden Fragebogenitems (4-stufig von „trifft nicht zu“ bis „trifft zu“) zu den angenommenen Qualitätskriterien des Unterrichtsgesprächs entwickelt:

- (1) Orientierung an einem Gesprächsziel
- (2) Transparenz von Regeln und Verantwortlichkeiten

- (3) Möglichkeiten für eigene Beteiligung
- (4) Möglichkeit für Beteiligung vieler Schüler/innen
- (5) Beziehung zum Gesprächspartner Lehrperson
- (6) Empfinden gegenüber dem schüleraktivierenden Unterrichtsgespräch
- (7) Fragen und Begleitung der Lehrperson

Das Fragebogeninstrument wurde im Rahmen einer Fortbildungsstudie zum Thema „Schüleraktivierung im Unterrichtsgespräch“ im Schuljahr 2016/17 eingesetzt. 16 Lehrpersonen nahmen mit ihren N = 343 Schüler/innen an einem einjährigen Fortbildungsprogramm teil. Die Schülerinnen und Schüler besuchten die 8. bis 10. Klasse des Gymnasiums. Zur Prüfung der vorgelegten Fragestellung wurden eine exploratorische (Hauptkomponentenanalyse mit Varimax-Rotation) und anschließend eine konfirmatorische Faktorenanalyse durchgeführt. In die Analysen wurden nur Fälle ohne fehlende Werte aus dem ersten Erhebungszeitpunkt (PräFortbildung) eingeschlossen (N = 304).

#### Ergebnisse

In der Hauptkomponentenanalyse konnten insgesamt 11 Faktoren (Eigenwert > 1) extrahiert werden, die zusammengenommen 63.30% der Varianz erklärten (KMO = .71, Bartlett-Test =  $\chi^2(406) = 1707.73$ ,  $p < .001$ ; Ladungshöhe Items  $\alpha > .40$ ). Für die anschließende konfirmatorische Faktoranalyse wurden fünf Faktoren und die dazugehörigen Items (N = 10) ausgeschlossen, da auf diese Faktoren jeweils weniger als drei Items luden. Resultierend war Qualitätsdimension (7) nicht abbildbar.

Die Ergebnisse der konfirmatorischen Faktorenanalyse mit den resultierenden sechs Faktoren (Qualitätsdimension (1)-(6)) zeigten eine gute Modellpassung:  $\chi^2(137) = 175.169$ ,  $p < 0.05$ ,  $\chi^2/df = 1.278$ , CFI = .967, TLI = .958, RMSEA = .030, SRMR = .053\*. Die Modellparameter wurden mittels der Maximum-Likelihood-Methode geschätzt. Alle 19 Items zeigten signifikante Ladungshöhen ( $p < .001$ ). Für die sechs Subskalen des Unterrichtsgesprächs wurden akzeptable Reliabilitätskoeffizienten ( $\alpha(1) = .57$ ,  $\alpha(2) = .62$ ,  $\alpha(3) = .75$ ,  $\alpha(4) = .69$ ,  $\alpha(5) = .72$ ,  $\alpha(6) = .65$ ) erreicht.

#### Diskussion

Für einige Qualitätsdimensionen des Unterrichtsgesprächs konnten reliable Skalen bestätigt werden. Zur Diskussion steht, ob bestimmte für Schüler/innen beobachtbare Qualitätsaspekte wie offene Lehrerfragen bevorzugt auf Einzelitembasis abgebildet werden. Die reliablen Skalen kommen zum Einsatz um die Implementation des Unterrichtsgesprächs sowie dessen Veränderbarkeit aus Schülerperspektive zu untersuchen. Dabei wird der Vorteil spezifisch auf das Fortbildungsangebot zugeschnittener Instrumente diskutiert.

### **Einheitliches Entscheidungskalkül? Die Rolle stereotypischer Denkweisen bei der Wahl der weiterführenden Schule**

**Melanie Trypke, Katja Serova, Ferdinand Stebner, Joachim Wirth**

Ruhr-Universität Bochum, Deutschland; [melanie.trypke@rub.de](mailto:melanie.trypke@rub.de)

Die wachsende lern- und bildungsbiografische Eigenverantwortung (Bittlingmayer, 2005) zeigt sich in der Bildungsentscheidung der Eltern bei der Wahl der weiterführenden Schule ihres Kindes. Gegenstand des Forschungsinteresses ist vorrangig die soziale Herkunft der Eltern im Zusammenhang mit der Schulwahl (z.B. Schneider, 2004). Vergleichsstudien wie IGLU und PISA belegen immer wieder die Tendenzen eines schichtspezifischen Entscheidungsverhaltens. Zurückgeführt werden diese Unterschiede zumeist auf sozioökonomische und soziokulturelle Merkmale, die in unterschiedliche Bildungsaspirationen münden (z.B. Paulus & Blossfeld, 2007). Aus Sicht der Rational-Choice Theory sind die Bildungsentscheidungen Resultate individueller Kosten-Nutzenkalkulationen der Eltern (Kristen, 1999).

In diesem Forschungsbereich bleibt ein Gegenstand weitestgehend unberücksichtigt: die Vorstellungen der Eltern über die einzelnen Schulformen. Auf Basis der in der Öffentlichkeit vermittelten Stereotypisierung der Hauptschule (Knigge, 2009) besteht die Frage, ob sich auch stereotypische Denkweisen über die anderen Schulformen in den Köpfen der Eltern implementiert haben und einzelne Merkmale (z.B. sozioökonomische Situation) von Familien mit einer bestimmten Schulform assoziiert werden. Für die Untersuchung wurden folgende Unterfragen formuliert:

- 1) Werden die einzelnen Merkmale unterschiedlich stark gewichtet?
- 2) Beeinflussen die soziodemografischen Daten der Eltern die Wichtigkeit der einzelnen Merkmale?
- 3) In welchem Maße verändert die Ausprägung der Merkmale die Präferenz, eine bestimmte Schulform zu wählen?

Um diese Fragen zu untersuchen, wurde im Rahmen des Kooperationsprojektes „Übergangsgestaltung“ der Stadt Herten vom Februar bis März 2017 eine Elternbefragung (N=243; 78% weiblich) der fünften Klassen durchgeführt. Schwerpunkt dieser zweiteiligen Befragung bildet der Einsatz des Verfahrens „Conjoint-Measurement“.

Hierbei handelt es sich um ein dekompositionelles multivariates Analyseverfahren, bei dem sich der Gesamtnutzen eines Produktes additiv aus den Teilnutzwerten seiner Merkmale zusammensetzt (Backhaus et al., 2015). Das Ergebnis besteht in der Ausgabe sogenannter Teilnutzenwerte für die einzelnen Ausprägungen jedes Merkmals und ihrer relativen Wichtigkeit. Als Maß für die Effektstärke dient die Spannweite der Teilnutzenwerte, die Aufschluss über die Wahrscheinlichkeit einer Präferenzveränderung gibt (Backhaus et al., 2015).

Die relevanten Merkmale der Familien wurden aus dem bisherigen Forschungsstand abgeleitet und mit dichotomen Ausprägungen dargestellt (z.B. niedrig/hoch): Menge an Informationen über die Schulformen (z.B. Bell, 2009), Leistungsstärke des Kindes (z.B. Schwippert et al., 2011), finanzielle Situation der Familie und Bildungsniveau der Eltern (z.B. Schneider, 2004). Die Eltern bekamen acht fiktive Familien mit unterschiedlichen Merkmalskombinationen vorgelegt (zufälliges orthogonales Design). Die Aufgabe bestand darin, anhand einer Rating-Skala (0-100%) die Wahrscheinlichkeit einzuschätzen, mit der sich die vorgestellten Familien für die jeweilige Schulform (Gymnasium, Realschule, Gesamtschule, Sekundarschule) entscheiden.

Die Analyse (N=112) bezieht sich, aufgrund einer hohen Drop-out Rate, auf einen reduzierten Stichprobenumfang. Eine Dropout-Analyse zeigt, dass dieser die Gesamtstichprobe gut repräsentiert.

Die Analysen zeigen folgende Ergebnisse:

- 1) Über alle Schulformen hinweg wird die Leistungsstärke des Kindes als wichtigstes und die finanzielle Situation der Familie als unwichtigstes Merkmal bewertet.

2) Die Wichtigkeit der einzelnen Merkmale wird nicht von den soziodemografischen Daten der Eltern beeinflusst.

3) Die Teilnutzenwerte der einzelnen Merkmale zeigen, dass eine hohe Leistungsstärke, eine gute finanzielle Situation, ein hohes Bildungsniveau der Eltern und viele Informationen über die Schulformen die Präferenz für das Gymnasium stärken. Interessant ist, dass genau die gleichen Ausprägungen auch die Präferenz für die Realschule erhöhen. Hingegen zeigen die Teilnutzenwerte der Eltern für die Gesamt- und Sekundarschule genau das entgegengesetzte Bild.

Es kann zusammengefasst werden, dass die Leistungsstärke des Kindes als das entscheidende Merkmal für die Wahl einer Schulform eingeschätzt wird; sozioökonomische und soziokulturelle Merkmale spielen eine untergeordnete Rolle. Dementsprechend kann bezüglich der Präferenzurteile von einem einheitlichen Entscheidungskalkül und nicht von stereotypischen Denkweisen gesprochen werden.

Allerdings haben, analog zu den IGLU-Ergebnissen, von diesen Eltern paradoxerweise größtenteils Akademikereltern ihre Kinder auf das Gymnasium geschickt ( $\chi^2(1, N=104)=15.6, p=.00$ ). Wie diese Ergebnisse erklärt werden können und wie die Bildungsforschung darauf reagieren kann, wird auf der Tagung diskutiert.

## **"Man kann immer klüger und klüger werden" - Zusammenhang von impliziten Theorien und Selbstregulationsprozessen im Kindergarten**

**Miriam Compagnoni<sup>1</sup>, Yves Karlen<sup>2</sup>**

<sup>1</sup>Universität Zürich, Schweiz; <sup>2</sup>PH FHNW, Schweiz; [mcompagnoni@ife.uzh.ch](mailto:mcompagnoni@ife.uzh.ch)

Theoretischer Hintergrund

Selbstregulationskompetenzen gehören zu den wichtigsten Prädiktoren erfolgreicher Schul- und Lebensverläufe (Moffitt, Poulton, & Caspi, 2013). Bereits im Kindergarten können sowohl exekutive Funktionen (z.B. Impulskontrolle, Arbeitsgedächtnis, kognitive Flexibilität) als auch kontextspezifisches selbstreguliertes Lernen (z.B. Planung von Arbeitsschritten, Monitoring) schulische Fähigkeiten vorhersagen (Blair & Raver, 2015; Gestsdotir et al., 2014; Neuenschwander, Rothlisberger, Cimeli, & Roebbers, 2012). Implizite Theorien über die Veränderbarkeit von Fähigkeiten werden als Überzeugungssystem beschrieben, das Lern- und Leistungszielorientierungen zu Grunde liegt und einen wesentlichen Einfluss auf die Entwicklung von Selbstregulationskompetenzen hat (Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013; Dweck & Leggett, 1988). Wer überzeugt ist, dass die Fähigkeiten veränderbar sind, zeigt eher eine Lernzielorientierung, nimmt Herausforderungen an, zeigt Persistenz und eine bessere Selbstregulation als Personen mit einer Sicht auf Fähigkeiten als eher stabiles Set prädestinierter Stärken und Schwächen. Ob implizite Theorien schon im Kindergartenalter entwickelt sind und sich die beschriebenen Zusammenhänge mit Selbstregulationskompetenzen zeigen, bleibt auf Grund von entwicklungspsychologischen Besonderheiten, wie der überoptimistischen Einschätzung der eigenen Fähigkeiten (e.g. Lipko, Dunlosky, Lipowski & Merriman, 2012) und fehlender empirischer Grundlagen unklar. Zudem steht kein Messinstrument zur Erfassung impliziter Theorien für diese Altersgruppe zur Verfügung.

Ziele

Dieser Studie liegen zwei Ziele zum differenzierteren Verständnis von Lern- und Entwicklungsprozessen zu Grunde: Erstens wird die Frage nach der Struktur impliziter Theorien von Kindern unter 7 Jahren untersucht. Es soll geklärt werden, ob schon im Kindergartenalter implizite Theorien reliabel und valide erfassbar sind und sich individuelle Unterschiede bezüglich Geschlecht, Leistung und Alter zeigen. Zweitens wird die Rolle von impliziten Theorien zur Erklärung von Unterschieden in exekutiven Funktionen (EF) und selbstreguliertem Lernen im Kindergarten (SRL) beleuchtet sowie angenommene differentielle Effekte bezüglich den zwei Selbstregulationsmassen im Zusammenhang mit vorakademischen Leistungen untersucht.

Methode

Eine Stichprobe von 143 Kindergartenkindern im zweiten Kindergartenjahr (52.4% weiblich; Alter,  $M = 6.4$  Jahre,  $SD = .39$ ) wurde in städtischen und ländlichen Regionen der Deutschschweiz mittels individueller Interviews befragt. Implizite Theorien wurden mit einem neu entwickelten Messinstrument basierend auf dem "Berkeley Puppet Interview" erfasst. Die Kinder hören die Antworten von zwei Plüschtieren auf Fragen zu impliziten Theorien (fix vs. veränderbar) und zeigen auf dem Touchscreen, mit welchem der Plüschtiere sie sich eher identifizieren können. Selbstregulationsprozesse wurden multimethodisch erfasst: EF mittels eines behavioralen Tests (HTKS, McClelland & Cameron Ponitz, 2012,  $\alpha = .90$ ), SRL im Kindergarten durch Lehrpersoneneinschätzungen (CBRS, Bronson, Goodson, Layzer, & Love, 1990,  $\alpha = .92$ ). Die vorakademische Leistung ( $\alpha = .84$ ) wurde von Lehrpersonen eingeschätzt.

Ergebnisse und Diskussion

Die faktorenanalytischen Ergebnisse zeigen, dass sich in dieser Altersgruppe implizite Theorien über die Veränderbarkeit von Intelligenz von Lern- und Leistungszielorientierungen unterscheiden. Außerdem wurden bedeutsame differentielle Zusammenhänge zwischen impliziten Theorien und Selbstregulationsprozessen gefunden. Kinder mit einer veränderbaren impliziten Intelligenztheorie weisen unter Kontrolle von Alter, Geschlecht und SES besseres SRL auf ( $F(4,134) = 6.51, p = .012$ , partial  $\eta^2 = .046$ ) als Kinder mit fixen impliziten Intelligenztheorie. In Bezug auf EF waren keinen signifikanten Unterschied feststellbar. Kinder mit einer stärkeren Lernzielorientierung weisen unter Kontrolle von Alter, Geschlecht und SES bessere EF ( $F(4,138) = 6.54, p = .012$ , partial  $\eta^2 = .045$ ) und SRL ( $F(4,137) = 8.35, p = .004$ , partial  $\eta^2 = .057$ ) als Kinder mit Leistungszielorientierung. Des Weiteren konnten in einem latent modellierten Strukturgleichungsmodell die angenommenen differentielle Zusammenhänge zwischen impliziten Intelligenztheorien, Zielorientierungen und Selbstregulationsprozessen modelliert werden sowie indirekte Effekte von impliziten Theorien über Selbstregulationsprozesse auf Leistung aufgezeigt werden (RMSEA = .061, CFI = .950,  $\chi^2 = 33.93, DF = 22, p < .049, n = 146$ ). Gesamthaft weisen die Ergebnisse darauf hin, dass sich implizite Theorien bereits im Kindergartenalter erfassen lassen und in einem Zusammenhang zu Selbstregulationsprozessen und Leistung stehen. Die Resultate verdeutlichen die Bedeutung einer frühen Förderung lernzielorientierter Überzeugungen, veränderbarer impliziter Theorien und Selbstregulationskompetenzen im Kindergarten.

## **Elterliche Bildungsaspiration und soziokulturelle Herkunft in der PISA-Studie. Eine Sekundäranalyse des Zusammenhangs zweier Konstrukte im Kontext sozialer Ungleichheit**

**Judith C. Horlboge, Ariane S. Willems**

Georg-August-Universität Göttingen, Institut für Erziehungswissenschaft, Deutschland; [judith.horlboge@sowi.uni-goettingen.de](mailto:judith.horlboge@sowi.uni-goettingen.de)

Seit der ersten PISA-Erhebung 2000 werden nicht nur die Kompetenzen von 15-jährigen Schüler\*innen in den Bereichen Lesen, Mathematik und Naturwissenschaften getestet, sondern u.a. die familiären Verhältnisse der Schüler\*innen analysiert.

Ein Bestandteil davon ist die elterliche Bildungsaspiration. Konzeptuell werden bereits seit Lewin et al. (1944) zwei Formen der Bildungsaspiration unterschieden: die idealistische und die realistische. Während die idealistische Bildungsaspiration der Eltern den gewünschten Schulabschluss für das Kind beschreibt, bezieht sich die realistische Aspiration auf den erwarteten Bildungsabschluss für das Kind. Für Deutschland werden in PISA kontinuierlich signifikante Kompetenzunterschiede in Abhängigkeit von der sozialen und kulturellen Herkunft der Schüler\*innen berichtet (Müller & Ehmke, 2016; Ehmke & Jude, 2010). Ergänzend zu diesen Befunden weisen aktuelle Studien nicht nur auf den Zusammenhang von Herkunft und Leistung hin, sondern zeigen auch, dass die elterliche Bildungsaspiration in Abhängigkeit des sozialen Status variiert (Rohlf, 2011) – allerdings ist die Befundlage hier noch uneindeutig (Deppe, 2013). Die theoretische Basis der vorliegenden Untersuchung ist eine Weiterentwicklung des Modells von Becker (2010) zu den Determinanten von Bildungsaspirationen. Als besondere Einflussfaktoren auf die Bildungsaspiration von Eltern werden dabei soziokulturelle Faktoren in Anlehnung an den Kapitalansatz von Bourdieu (1983) berücksichtigt.

Mit dem vorliegenden Beitrag gehen wir der Frage nach, wie die beiden Konstrukte elterliche Bildungsaspiration und soziale Herkunft in Deutschland zusammenhängen. Dabei gehen wir auch der Frage nach, inwieweit sich spezifische Zusammenhangsmuster für idealistische bzw. realistische Bildungsaspirationen und soziokultureller Herkunftsmerkmale nachweisen lassen.

Die Analysegrundlage zur Bearbeitung der Fragestellung sind die Schüler- und Elterndaten der PISA-2009 Erhebung, in der systematisch zwischen der idealistischen und realistischen Bildungsaspiration von Eltern unterschieden wird. Es lagen Daten von  $N = 1739$  Schüler\*innen und  $N = 1360$  dazugehörigen Eltern zur Untersuchung vor. Die elterliche Bildungsaspiration wurde dabei anhand von zwei Items operationalisiert, die auf Ordinalskalenniveau nach den gewünschten bzw. erwarteten Schulabschlüssen für ihr Kind fragen und dementsprechend die idealistische bzw. realistische Aspiration abbilden. Die Antwortskala reicht von 1 „Sonder-/Förderschule“ bis 8 „Universitäts-/Fachhochschulabschluss“ und wurde für die Durchführung von Regressionsanalysen in niedrige und hohe Bildungsaspiration dichotomisiert. Das Konstrukt der sozialen Herkunft wurde über zwei einzelne Items, den höchsten Schulabschluss der Eltern und die Anzahl der Bücher im Haushalt, sowie zwei Indizes, HISEI ( $M = 47.94$ ,  $SD = 15.1$ ) und ESCS ( $M = 0.11$ ,  $SD = 0.88$ ), gemessen. Darüber hinaus wurden die schulische Leistung – operationalisiert über die Lesekompetenz ( $M = 486.52$ ,  $SD = 86.01$ ) – das Geschlecht des Kindes sowie der familiäre Migrationshintergrund miteinbezogen. Zur Überprüfung des Zusammenhangs wurden Chi-Quadrat-Tests sowie logistische Regressionen durchgeführt. Die Kriteriumsvariable war dabei jeweils die idealistische bzw. realistische Bildungsaspiration der Eltern.

Die bivariaten Befunde zeigen erwartungsgemäß hoch signifikante Zusammenhänge der elterlichen Bildungsaspiration (realistisch sowie idealistisch) mit den ausgewählten Indikatoren der soziokulturellen Herkunft. Vertiefende logistische Regressionen zeigen, dass die soziokulturelle Herkunft der Familien – unter Kontrolle des Geschlechts und der Lesekompetenz der Schüler\*innen – stärker zur Vorhersage der realistischen ( $R^2 = .40$ ) als der idealistischen Bildungsaspiration ( $R^2 = .29$ ) beitragen. In Bezug auf die relative Vorhersagekraft der einbezogenen Prädiktoren für die Bildungsaspiration der Eltern zeigen sich signifikante Differenzen für beide Formen der Bildungsaspiration insbesondere hinsichtlich der Leistung ( $OR_i = 1.01$ ,  $pi = .00$ ;  $OR_r = 1.02$ ,  $pr = .00$ ), des Migrationshintergrundes ( $OR_i = 3.27$ ,  $pi = .00$ ;  $OR_r = 2.74$ ,  $pr = .00$ ) und des sozioökonomischen Status ( $OR_i = 1.04$ ,  $pi = .00$ ;  $OR_r = 1.04$ ,  $pr = .01$ ). Insgesamt weisen unsere Befunde darauf hin, dass sowohl die realistischen als auch die idealistischen Bildungsaspirationen primär durch das Kompetenzniveau der Schüler\*innen vorhergesagt werden, darüber hinaus allerdings auch durch soziokulturelle Merkmale der Familien. Substanzielle Unterschiede in der Vorhersage zwischen der idealistischen und realistischen Bildungsaspiration sind nur hinsichtlich des Migrationshintergrundes zu erkennen.

## **Entwicklung einer adaptiven Lehrkompetenz im Umgang mit Heterogenität in der zweiten Phase der Lehrerbildung**

**Thomas Rey**

Pädagogische Hochschule Heidelberg, Deutschland; [rey@ph-heidelberg.de](mailto:rey@ph-heidelberg.de)

Das Forschungsprojekt EKoL 02 untersucht die Entwicklung einer adaptiven Lehrkompetenz im Umgang mit Heterogenität von Referendarinnen und Referendaren während ihres Vorbereitungsdienstes. Adaptive Lehrkompetenz wird im Projekt als eine Kompetenz modelliert, welche die beiden Facetten Handlungs- und Planungskompetenz beinhaltet. Durch eine längsschnittliche Messung soll einerseits eine theoretisch konzeptualisierte Kompetenzentwicklung innerhalb des Vorbereitungsdienstes empirisch geprüft und förderliche Hintergrundfaktoren identifiziert werden. Darüber hinaus werden andererseits schriftliche Unterrichtsentwürfe von Lehramtsanwärterinnen und -anwärtinnen hinsichtlich definierter Kriterien analysiert, um eine postulierte adaptive Planungskompetenz der Stichprobe im Zeitverlauf quantifizieren zu können. Zusammenhangsmaße zwischen den beiden Kompetenzfacetten sollen identifiziert und das theoretisch postulierte Kompetenzmodell empirisch überprüft werden. Die Frage des Zusammenhangs zwischen adaptiver Planungskompetenz und einer kontext- und situationsbezogenen Handlungskompetenz stellt bis dato ein Forschungsdesiderat dar. Auf Basis der vorliegenden Messergebnisse werden Zusammenhangsmaße zwischen den beiden Kompetenzfacetten untersucht und intra-individuelle Entwicklungsverläufe abgebildet.

Theoretischer Hintergrund. Das Forschungsvorhaben ist dem kompetenztheoretischen Ansatz verpflichtet und gründet auf dem Modell professioneller Handlungskompetenz im Lehrberuf nach COACTIV (Baumert & Kunter 2006): Neben Motivation, Überzeugungen und selbstregulativen Fähigkeiten von Lehrkräften wird das Professionswissen fokussiert. Dies gründet auf Shulmans (1987) postuliertes ‚pedagogical knowledge‘. Umgang mit Heterogenität wird im Projekt als Wissen über effektive Klassenführung, über Unterrichtsmethoden, über Leistungsbeurteilung sowie über individuelle Unterschiede und die daraus resultierenden Anforderungen an Lehrkräfte modelliert (Voss & Kunter 2011). Ein in diesem Kontext bedeutender Ansatz ist der des ‚adaptive teaching‘ (Beck et al. 2010; Brühwiler 2014). Dabei unterscheiden die Verfasser zwischen adaptiver Planungs- und adaptiver Handlungskompetenz als Bestandteile des Konstrukts der adaptiven Lehrkompetenz.

Fragestellungen. Wie sich eine adaptive Handlungskompetenz im Umgang mit heterogenen Gruppen im Verlauf der zweiten Phase der Lehrerbildung entwickelt, ist die in diesem Forschungsprojekt verfolgte zentrale Fragestellung. Vor diesem Hintergrund eruiert das Forschungsvorhaben, (1) ob sich Veränderungen zwischen den beiden Messzeitpunkten aufzeigen lassen, (2) diesbezüglich förderliche Hintergrundfaktoren (Kovariate) identifiziert werden können und (3) sich Entwicklungszusammenhänge zwischen der adaptiven Planungs- und Handlungskompetenz bei Referendarinnen und Referendaren zeigen?

Methode. Im Rahmen einer ersten Projektphase wurde ein text- und videobasierter Vignettest zur Messung der ‚Adaptiven Handlungskompetenz in heterogenen Gruppen‘ entwickelt, validiert und querschnittlich bei Lehramtsstudierenden eingesetzt ( $n=706$ ). Basierend auf den Ergebnissen aus der ersten Projektphase wird das vignettengestützte Testinstrument zu Beginn

und zum Ende des Referendariats an Staatlichen Seminaren für Didaktik und Lehrerbildung in Baden-Württemberg eingesetzt. Die geplante Stichprobengröße liegt bei  $n > 150$ , um eine Skalierung der Daten nach der Item-Response-Theorie zu ermöglichen. Zudem lassen sich förderliche Hintergrundfaktoren wie Berufswahlmotive (Pohlmann & Möller 2010) und Einstellung zu Inklusion (Seifried & Heyl 2016) der Lehramtsanwärterinnen und -anwärter identifizieren sowie Dimensionen des beruflichen Selbstkonzeptes (Retelsdorf et al. 2014) lokalisieren.

Ferner wird versucht, den Zuwachs an gemessener Planungskompetenz im Vorbereitungsdienst längsschnittlich abzubilden. Hierzu werden Unterrichtsentwürfe von Lehramtsanwärterinnen und -anwärtinnen ( $n > 150$ ) hinsichtlich spezifischer Kriterien analysiert, um Planungsentscheidungen zu quantifizieren. Die insgesamt 15 Items zur Erfassung der adaptiven Planungskompetenz sind ebenfalls in Anlehnung an Beck et al. (2008) an die drei theoretisch begründeten Subfacetten des professionellen Handelns im Umgang mit Heterogenität sowie an evidenzbasierten Forschungsmethoden (König, Buchholtz & Dohnen 2015) und den Ausbildungsstandards der Staatlichen Seminare für Didaktik und Lehrerbildung in Baden-Württemberg (Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2015) konzeptualisiert worden.

Erwartete Ergebnisse und Diskussion. Es wird ein Zuwachs hinsichtlich der adaptiven Handlungs- als auch der Planungskompetenz im Umgang mit Heterogenität in der zweiten Phase der Lehrerbildung angenommen. Zudem ist davon auszugehen, dass bestimmte Hintergrundfaktoren (Kovariate) einen positiven Effekt auf die Entwicklung der Handlungskompetenz bedingen. Ferner wird erwartet, dass ein Zusammenhang in der Genese der Handlungs- und Planungskompetenz bei den Probandinnen und Probanden aufzufinden ist. Die projektspezifischen Erkenntnisse vermögen Aufschluss über den Entwicklungsprozess einer adaptiven Lehrkompetenz im Vorbereitungsdienst zu geben und tragen zu einer evidenzbasierten Lehrerbildung in der zweiten Phase bei.

## **Mittendrin statt nur dabei? Effekte der Klassenkomposition auf die soziale Integration in inklusiven Schulklassen**

**Hannah Decker, Sebastian Röhl, Katja Scharenberg**

Pädagogische Hochschule Freiburg, Deutschland; [hannah.decker@stud.ph-freiburg.de](mailto:hannah.decker@stud.ph-freiburg.de)

### Theoretischer Hintergrund

Seit der Ratifizierung der UN-Behindertenrechtskonvention (United Nations, 2006) wird in Deutschland verstärkt über gleichberechtigte gesellschaftliche Teilhabe von Menschen mit und ohne Behinderung diskutiert (Werning, 2014). Bei der Umsetzung von Inklusion im schulischen Bereich ist neben gleichberechtigten Bildungschancen auch soziale Teilhabe ein wichtiges Zielkriterium des gemeinsamen Unterrichts (KMK, 2011). Trotz Leistungsvorteilen in Regelklassen (z.B. Cole, Waldron & Majd, 2004; Kocaj et al., 2014; Ruijs & Peetsma, 2009) scheinen Kinder mit sonderpädagogischem Förderbedarf (SPF) in inklusiven Settings hinsichtlich ihrer sozialen Integration jedoch häufig benachteiligt zu sein (Frostad & Pijl, 2007; Koster, Nakken, Pijl & van Houten, 2010). Zur Bedeutung der Zusammensetzung (Komposition) von Schulklassen als unmittelbarer Lern- und Entwicklungsumwelt sowie zu Auswirkungen solcher Kompositionseffekte liegt mittlerweile eine Reihe empirischer Befunde vor (z.B. Dumont, Neumann, Maaz & Trautwein, 2013; Harker & Tymms, 2004; Scharenberg, 2012; Van der Slik, Driessen & De Bot, 2006; Van Ewijk & Slegers, 2010; Zurbriggen, 2016): Nationale und internationale Studien haben wiederholt gezeigt, dass die Zusammensetzung der Schülerschaft unabhängig von und zusätzlich zu individuellen Lernvoraussetzungen der Schülerinnen und Schüler eigenständig bedeutsam für die erzielten schulischen Kompetenzen ist. Diese Studien legen nahe, dass unterschiedliche Lernkontexte differenzielle Lern- und Entwicklungsmilieus darstellen, die die individuelle Entwicklung von Schülerinnen und Schülern beeinflussen können (Baumert, Stanat & Watermann, 2006). Für die soziale Integration als sozioemotionale Kriteriumsvariable im Kontext schulischer Inklusion liegen bislang jedoch kaum Forschungsbefunde zur Effektivität unterschiedlicher Schulklassen vor.

### Fragestellung

Unser Beitrag geht daher der Frage nach, ob Klassenkompositionsmerkmale über Individualmerkmale hinaus einen Effekt auf die subjektive Wahrnehmung von sozialer Integration bei Schülerinnen und Schülern in inklusiven Settings ausüben. Es werden die Instrumente, verschiedene Möglichkeiten und Grenzen der Operationalisierung von sozialer Integration sowie erste Ergebnisse aus der Pilotierungsstudie vorgestellt und diskutiert.

### Methode

Erhebungsfeld sind Gemeinschaftsschulen in Baden-Württemberg. Neben Schülerinnen und Schülern (Jahrgangsstufen 5, 6 und 7) werden auch ihre Klassenlehrkräfte befragt, um Wahrnehmungen und Kontextbedingungen sozialer Integration multiperspektivisch und mehrdimensional abzubilden. Zur quantitativen Erfassung der Selbsteinschätzung der sozialen Integration wurde ein Instrument in Anlehnung an Bossaert et al. (2013) und Koster et al. (2009) entwickelt, das vier Dimensionen abbildet: 1.) Peerakzeptanz (5 Items, z.B.: Ich bin bei meinen Mitschülern beliebt), 2.) Freundschaft (5 Items, z.B.: Ich habe wenige Freunde in meiner Klasse), 3.) soziale Selbstsicht (5 Items, z.B. Ich fühle mich einsam), 4.) Interaktionen (5 Items, z.B.: Ich spiele in der Pause mit Mitschülern). Darüber hinaus wurden auch soziometrische Nominierungs- und Ratingverfahren eingesetzt. Über den Schülerfragebogen wurden soziodemografische Merkmale (z.B. Geschlecht, soziale Herkunft, Migrationshintergrund) erhoben, die auf aggregierter Ebene (Aggregation über Mittelwert und Streuungsmaße) zur Operationalisierung der Kompositionsmerkmale dienen. Der Fragebogen für die Klassenlehrkräfte erfasste u.a. Einstellungen zur schulischen Inklusion, Kontextbedingungen sowie Unterrichtsmerkmale in den jeweiligen Schulklassen (z.B. Sozialformen, Binnendifferenzierung, Individualisierung).

### Ergebnisse und Diskussion

Die Ergebnisse aus der Pilotierungsstudie (55 Schülerinnen und Schüler, 3 Schulklassen, Jahrgangsstufe 5 und 6) weisen überwiegend eine hinreichende Reliabilität der vier Subskalen auf (Peerakzeptanz:  $\alpha = .687$ , Freundschaft:  $\alpha = .802$ , Selbstsicht:  $\alpha = .837$ , Interaktion:  $\alpha = .576$ ). Die Korrelationen zwischen den vier Subdimensionen fallen bei Schülerinnen und Schülern mit SPF geringer aus ( $.473 \leq r \leq .777$ ,  $p < .05$ ) als bei jenen ohne SPF ( $.718 \leq r \leq .901$ ,  $p < .01$ ).

Die Analysen deuten unterschiedliche Wahrnehmungen von sozialer Integration in inklusiven Schulklassen bei Schülerinnen und Schülern mit und ohne SPF an. Unser Forschungsvorhaben leistet somit einen wichtigen Beitrag zur Frage, welches Integrationspotenzial unterschiedliche Klassenkontexte besitzen, die sich in ihrer soziodemografischen und leistungsbezogenen Zusammensetzung der Schülerschaft unterscheiden. Es dient dazu, jene Strukturmerkmale und Barrieren zu identifizieren, die der sozialen Integration in inklusiven Schulklassen potenziell entgegenstehen können. Die zu erwartenden Ergebnisse des Forschungsprojekts können genutzt werden, um sozialen Ausgrenzungsmechanismen durch inklusive Schul- und Unterrichtsentwicklung gezielt entgegenwirken zu können.

## **Förderung der professionellen Wahrnehmung von Unterrichtsstörungen. Überprüfung der Effektivität einer für Noviz\*innen optimierten Lehr-Lerngelegenheit.**

**Victoria L. Barth<sup>1</sup>, Valentina Piwowar<sup>1</sup>, Diemut Ophardt<sup>2</sup>, Katharina Krysmanski<sup>1</sup>, Irina Kumschick<sup>1</sup>, Felicitas Thiel<sup>1</sup>**

<sup>1</sup>Freie Universität Berlin, Deutschland; <sup>2</sup>Technische Universität Berlin, Deutschland; [victoria.barth@fu-berlin.de](mailto:victoria.barth@fu-berlin.de)

### Theoretischer Hintergrund

Der Umgang mit Störungen im Unterricht zählt zum Kompetenzbereich des Klassenmanagements (Ophardt & Thiel, 2013) und wird von Lehrpersonen als zentrale Herausforderung wahrgenommen (Gallagher, 2009). Ein wesentlicher Erfolgsfaktor für einen effektiven Umgang mit Störungen ist eine professionelle Problemwahrnehmung (Wettstein, 2013). Unter dem Begriff der „professionellen Wahrnehmung“ wird verstanden, dass in der hochkomplexen Situation des Unterrichts die relevanten Merkmale identifiziert (Erkennenskompetenz) und aus diesen Schlussfolgerungen gezogen werden (Beurteilungskompetenz) (Sherin, 2001). Eine professionelle Unterrichtswahrnehmung ist damit eine zentrale Voraussetzung für erfolgreiches Lehrer\*innenhandeln (Grossman et al., 2009; Sherin, 2001). Besonders in der ersten Phase der Lehrkräftebildung fehlt es Studierenden jedoch an notwendigen Kenntnissen und Fertigkeiten, um relevante Aspekte aus dem Unterrichtsgeschehen herauszufiltern, diese theoriegeleitet zu beurteilen und darauf aufbauend alternative Handlungsstrategien zu entwickeln und auszuführen (Berliner, 1994, 2004; Emmer & Gerwels, 2006). Damit Noviz\*innen lernen, auf (plötzlich auftretende, störende) Ereignisse adäquat zu reagieren, ist es notwendig, sie bereits im Studium beim Aufbau dieser Kompetenz durch die Entwicklung, Durchführung und Evaluation von gezielten Lehr-Lerngelegenheiten zu unterstützen (Barth, 2017).

### Fragestellung

Während bisherige Studien bereits zeigen konnten, dass Kompetenzen der professionellen Wahrnehmung durch videobasierte Fallarbeit gefördert werden können (Goeze, Zottmann, Vogel, Fischer & Schrader, 2014; Krammer & Reusser, 2005; Piwowar, Thiel & Ophardt, 2013), liegen kaum Befunde vor, die die Effektivität der Lehr-Lerngelegenheit, in die die Fallarbeit eingebunden ist, untersuchen. Die vorliegende Arbeit widmet sich daher zwei Fragestellungen:

1. Unterscheiden sich Lehramtsstudierende in ihrer professionellen Wahrnehmung, wenn ein problembasiertes Lernarrangement durch eine direkte Instruktion ergänzt wird?
2. Können durch diese adaptierte Lehr-Lerngelegenheit die Kompetenzen der professionellen Wahrnehmung und des professionellen Wissens gesteigert werden?

### Methode & Ergebnisse

Zwei quasi-experimentelle Studien wurden durchgeführt: In Teilstudie 1 wurden Lehramtsstudierende des Masters (N=198) zwei Versuchsbedingungen zugeordnet: In einer Gruppe wurden gemäß des klassischen problembasierten Lernens (PBL) zentrale theoretische Konzepte zum effektiven Umgang mit Störungen im Unterricht selbstständig in Kleingruppen erarbeitet (n=108). In der anderen Gruppe fand ein instruiertes problembasiertes Lernen statt (I-PBL), bei dem den Studierenden die theoretischen Konzepte vermittelt wurden (n=90). Beide Gruppen arbeiteten anschließend in Kleingruppen mit videobasierten Fällen. Die Erfassung der professionellen Wahrnehmung erfolgte zur Posttestung durch eine videobasierte Fallanalyse mit offenem Antwortformat (Barth, 2017). Die Untersuchung zeigte, dass sich die Gruppen nach Absolvieren der Lehr-Lerngelegenheiten nicht hinsichtlich der erkannten störungskritischen Unterrichtsmerkmale unterscheiden ( $F=3.085$ ,  $df=1$ ,  $p=.081$ ,  $d=.25$ ), jedoch Studierende der I-PBL-Gruppe störungsanfällige Unterrichtssituationen elaborierter beurteilen ( $F=6.484$ ,  $df=1$ ,  $p=.012$ ,  $d=.36$ ) als Studierende der PBL-Gruppe (Piwowar et al., in Begutachtung).

Bezüglich der zweiten Fragestellung wurde in einer weiteren quasi-experimentellen Interventionsstudie (Piwowar et al., in Begutachtung) untersucht, inwieweit durch das I-PBL-Lernsetting sowohl die Kompetenzen der professionellen Wahrnehmung als auch das Wissen zum Klassenmanagement der Studierenden verbessert werden können. 180 Lehramtsstudierende des Masters wurden zwei Versuchsbedingungen zugeordnet: die Gruppe A wurde Prä-Post befragt (n=59) und die Gruppe B nur Post (n=121). Dieses Studiendesign erlaubte, a) Lernzuwächse zu messen bei b) gleichzeitiger Kontrolle potentieller Lerneffekte durch die bereits zum ersten Messzeitpunkt eingesetzten Instrumente: eine videobasierte Fallanalyse zur professionellen Wahrnehmung (Barth, 2017) und ein objektives Leistungsmaß zum Wissen im Klassenmanagement (Multiple-Choice, Krysmanski & Piwowar, in Vorbereitung). Es konnte gezeigt werden, dass sich die Ergebnisse der Posttestungen von Gruppe A und Gruppe B hinsichtlich des Wissens im Klassenmanagement ( $F=1.478$ ,  $df=1$ ,  $p=.226$ ) als auch hinsichtlich der professionellen Wahrnehmung (für Erkennen:  $F=.004$ ,  $df=1$ ,  $p=.949$ ; für Beurteilen:  $F=.723$ ,  $df=1$ ,  $p=.396$ ) nicht signifikant unterscheiden, weswegen kein Lerneffekt durch die Instrumente angenommen wird. Varianzanalysen mit Messwiederholung zeigten über die Zeit eine signifikante Verbesserung des Wissens im Klassenmanagement ( $F=56.456$ ,  $df=1$ ,  $p<.001$ ,  $\eta^2=.489$ ) sowie der Beurteilungskompetenz ( $F=15.989$ ,  $df=1$ ,  $p<.001$ ,  $\eta^2=.216$ ). Die Erkennenskompetenz blieb über die Zeit stabil ( $F=.114$ ,  $df=1$ ,  $p=.737$ ,  $\eta^2=.002$ ). Die gefundenen Ergebnisse werden integrierend diskutiert.

## **Modelling and predicting non-linear changes in educational trajectories: The multilevel latent growth components approach**

**Christoph Kiefer, Axel Mayer**

RWTH Aachen University, Deutschland; [christoph.kiefer@psych.rwth-aachen.de](mailto:christoph.kiefer@psych.rwth-aachen.de)

The investigation of developmental trajectories is a central goal of educational science. However, investigating a research question like “Does a change of major after the first year of studies improve students’ satisfaction with their academic success sustainably?” already requires a statistical model that considers a) the potentially non-linear trajectories of satisfaction as well as b) the multiple levels of analysis (i.e. the individual level, the university level etc.) and c) a measurement model as satisfaction is a non-observable, latent variable. In this poster, we use NEPS data to develop and illustrate a new statistical approach that can adequately address all three challenges simultaneously. The new approach is termed multilevel latent growth components model (ML-LGCoM) and can be used to model the trajectory of an outcome-of-interest in its full complexity. A key feature of this approach is, that the functional form of the trajectories does not have to be pre-specified – in contrast to latent growth curve models. In our NEPS example, we model a) the (non-linear) development of students’ satisfaction with their academic success over four years, b) cluster- and individual-level trajectories and c) measurement error.

Our initial findings show a sustainable change in satisfaction after the first year both on cluster- (satisfaction increases on average) and on individual-level (much variability in individual trajectories), but later on satisfaction appears very stable on both levels. These findings indicate that during the first years of studies satisfaction fluctuates, but converges to a stable level

afterwards. In addition, we examine individual-specific predictors (e.g. change of major) and cluster-level predictors (e.g. difficulty of study program) of this process of stabilisation.

## **Beobachtung, Gestaltung und Reflexion von entwicklungsförderlicher sprachlicher Interaktion: Ein Seminarkonzept für Studierende des Lehramts an beruflichen Schulen/ Fachrichtung Sozialpädagogik**

**Anita Knöferle, Dorothea Dornheim**

Universität Bamberg, Deutschland; [anita.knoeferle@uni-bamberg.de](mailto:anita.knoeferle@uni-bamberg.de)

Im Prozessmodell zur professionellen Handlungskompetenz von Lehrpersonen (1) werden Kognitionen (z.B. pädagogisch-psychologisches Wissen) sowie motivational-affektive Einstellungen/Überzeugungen über vermittelnde situationspezifische Fertigkeiten als relevant für unterrichtliches Handeln angesehen (2). Für pädagogisch-psychologische Wissensbestände zum Lernen/Lehren konnte die Relevanz für erfolgreiches Unterrichten empirisch gezeigt werden (3). Wenig untersucht sind im deutschsprachigen Raum Zusammenhänge von entwicklungspsychologischen Wissensbeständen und professioneller Handlungskompetenz (4). Allerdings zeigen Trainingsstudien aus dem englischsprachigen Raum, dass Wissen über kognitiv-sprachliche Entwicklungsveränderungen und Einüben sprachförderlicher Interaktionen mit Zuwächsen in sprachlichen Kompetenzen bei Kindern einhergeht (5,6).

In diesem Zusammenhang stellt sich die Frage, wie Wissensbestände und förderliche Überzeugungen zu kognitiv-sprachlicher Entwicklung und Förderung Lehramtsstudierenden in einem Seminar so vermittelt werden können, dass sie Kompetenzen im Beobachten und Gestalten von entwicklungsförderlichen Interaktionen mit dem entsprechenden theoretischen Wissen in Verbindung bringen und reflektieren können und damit erste professionelle Handlungskompetenzen angebahnt werden.

Konkret: Kann Wissen zu kognitiv-sprachlicher Entwicklung und Förderung allein durch den Theorieteil eines Seminars gesteigert werden und davon ausgehend Überzeugungen zur Bedeutsamkeit von Sprachförderung positiv verändert werden? Lässt sich durch die Kombination von Wissensvermittlung und Beobachtungsübung die Beobachtungskompetenz steigern und können diese Effekte durch den Praxisteil intensiviert werden? Kann durch die Kombination von theoretischem Wissen und Praxis (intensiviert durch Selbst- und Fremdbeobachtung) Handlungswissen generiert und durch die Kombination von theoretischem Wissen und Praxis das Handlungswissen besser auf die theoretischen Grundlagen rückbezogen werden als in einer Kontrollgruppe ohne Praxiserfahrung?

Mit 30 Lehramtsstudierenden, vorwiegend Berufliche Schulen (Unterrichtsfach Pädagogik/Psychologie für fröhpädagogische Fachkräfte), wurde ein Seminarkonzept „Entwicklung und entwicklungsförderliche sprachliche Interaktion“ mit einer Theoriephase an der Universität und einer mehrwöchigen Praxisphase im Kindergarten mit anschließender Reflexion umgesetzt.

Der Theorieteil vermittelte Grundlagen zur kognitiven Entwicklung, Sprachentwicklung, Entwicklungstheorie von Vygotski und Befunde zur Bedeutung/Erfassung von Prozessqualität (Erwachsenen-Kind-Interaktion) mit Blick auf die weitere Kompetenzentwicklung von Kindern. Alltagsintegrierte Förderprinzipien (sprachliche Förderung durch dialogisches Bilderbuchlesen, Förderung mathematischer Basiskompetenzen in Projekten, Sustained Shared Thinking bei naturwissenschaftlichen Experimenten) bildeten den Abschluss.

Der Praxisteil bestand aus vier Fördersitzungen, in denen zwei Studierende mit zwei Kindergartenkindern nacheinander in einer 1:1-Situation Interaktionen in einem gewählten Förderbereich gestalteten. Die Tandempartner beobachteten sich wechselseitig mit einer verkürzten Beobachtungsskala in Anlehnung an das Classroom Assessment Scoring System (CLASS (7)) hinsichtlich der Qualität ihrer sozio-emotionalen und instruktionalen Unterstützung. Videoaufnahmen der Sitzungen, der Abgleich von Selbst- und Fremdwahrnehmung sowie die Diskussion im Tandem und abschließend im Seminar sollten die Beobachtungs-, Interaktions- und Reflexionskompetenzen der Studierenden verbessern (vgl. Evidenzbasierte Methoden der Unterrichtsdiagnostik und -entwicklung (EMU (8,9))).

Folgende Maße wurden auf der Grundlage eines Prä-Post-Designs mit drei Messzeitpunkten (vor/nach Theorieteil, nach Praxisphase) eingesetzt: Wissen über kognitiv-sprachliche Entwicklung sowie entwicklungsförderliche sprachliche Interaktion (Single-Choice-Test), Überzeugungen zum Thema Sprachförderung (Fragebogen), Erfassung der Beobachtungskompetenz bei Erwachsenen-Kind-Interaktionen (Videorating) sowie Erfassung von praxisbezogenem Handlungswissen, Reflexionsfähigkeit und Verknüpfung mit Theoriewissen (Fallvignetten). Dieselben Variablen werden in einer Kontrollgruppe mit gleichem Seminarinhalt, in klassischer Seminarkonzeption ohne Praxisphase, im laufenden Semester erhoben.

Erste Auswertungen zeigen, dass sich durch den Theorieinput Wissen und Überzeugungen zum Thema Sprachförderung signifikant positiv veränderten. Während die subjektive Bedeutsamkeit von Sprachförderung kontinuierlich über die Messzeitpunkte anstieg, wurde das (intuitive) theoretische Wissen über Sprachförderung (T1) offenbar durch die Theoriephase zunächst in Frage gestellt, was sich in einer kurzzeitigen signifikanten Verschlechterung der Leistung (T2) manifestierte. Zum letzten Messzeitpunkt (T3) wurden im Vergleich zu T1 sogar leicht höhere Werte im Wissen über Sprachförderung gemessen. Dies könnte dadurch erklärt werden, dass das anfängliche intuitive Wissen durch die Praxiserfahrung, gekoppelt mit stabilerem theoretischem Wissen, gestärkt und stabilisiert wurde. Analysen des Videoratings der Studierenden im Vergleich mit dem von Experten sowie erste Auswertungen der Fallvignetten verdeutlichen die wachsende Beobachtungskompetenz sowie die Fähigkeit, praktisches Handlungswissen zu generieren und auf theoretische Wissensbestände rückzubeziehen. Der Vergleich mit Ergebnissen der Kontrollgruppe ohne Praxisphase soll weiteren Aufschluss bringen.

## **Erfassung von Sprachkompetenzen: Die Validitäten verschiedener Items zur Selbsteinschätzung**

**Anike Dröscher**

Leibniz-Institut für Bildungsverläufe e.V., Deutschland; [anike.droescher@ifbi.de](mailto:anike.droescher@ifbi.de)

Korrelationen zwischen objektiven Kompetenzmaßen und selbsteingeschätzten Kompetenzen haben durchweg kleine bis moderate Effektstärken (s. Freund & Kasten, 2012; Mabe & West, 1982; Zell & Krizan, 2014). Edele, Seuring, Kristen und Stanat (2015) haben mithilfe der Daten des Nationalen Bildungspanels gezeigt, dass multivariate Analysen zu unterschiedlichen Ergebnissen führen, abhängig davon, ob man Selbsteinschätzungen oder Testergebnisse als Indikator für Sprachkompetenzen verwendet. Die Eignung von Selbsteinschätzungen als Maß für Sprachkompetenzen ist dementsprechend fraglich. Zur validen Erfassung von Sprachkompetenzen sind Sprachtests bislang die beste Wahl. Jedoch ist es nicht in allen Studien möglich, Kompetenzmessungen durchzuführen. Deshalb sind ökonomische, aber valide Befragungsinstrumente zur Erfassung von Sprachkenntnissen auch weiterhin unverzichtbar. Nimmt man also an, dass Sprachkompetenzen auch zukünftig mittels Selbsteinschätzungen erfasst werden müssen, muss es das Ziel sein, die Validität von Selbsteinschätzungen durch die

Wahl geeigneter Frageformulierungen und Antwortskalen zu steigern. Freund und Kasten (2012) haben z.B. metaanalytisch gezeigt, dass die Validität von Selbsteinschätzungen höher ist, wenn relative Skalen mit klar definierten Vergleichsgruppen eingesetzt werden.

Um das Potenzial verschiedener Selbsteinschätzungsitems als Kompetenzmaße zu ergründen, setzen wir in der ersten Welle der Panelstudie ReGES (Refugees in the German Educational System) am IIfBi (Leibniz-Institut für Bildungsverläufe e.V.) vier verschiedene Arten von Items zur Selbsteinschätzung der Deutschkompetenzen ein. Neben den vor allem in Large-Scale Assessments eingesetzten Standarditems, wie z.B. „Wie gut verstehen Sie Deutsch?“ mit einer fünfstufigen Antwortskala von „gar nicht“ bis „sehr gut“ werden wir diese Fragen auch mit einer ebenfalls fünfstufigen Antwortskala, die deutsche Muttersprachler als Vergleichsgruppe vorgibt, einsetzen. Darüber hinaus werden wir diese Fragen auf einer Endpunktgelabelten 10er-Skala von „sehr schlecht“ bis „sehr gut“ bewerten lassen. Zuletzt werden wir ein Multiple Choice-Item einsetzen, in dem der Befragte von 14 Kompetenzen diejenigen auswählen soll, die er auf Deutsch kann, wie z.B. „etwas zum Essen und Trinken bestellen“. Als Kriterium dienen die vierte Version des deutschen Peabody Picture Vocabulary Tests (PPVT-4; Lenhard, Lenhard, Segerer & Suggate, 2015) sowie die deutsche Version des Tests zum Grammatikverständnis (TROG-D; Fox-Boyer, 2016). In der ReGES-Studie werden 2400 jugendliche Flüchtlinge befragt und von diesen 1200 zusätzlich getestet.

Ziel der Arbeit ist es, die Validität der vier verschiedenen Arten von Selbsteinschätzungsitems und deren Eignung zum Einsatz in multivariaten Analysen als Maß der Sprachkompetenz theoretisch zu diskutieren und später empirisch zu untersuchen.

Die Methodik wird an die Studie von Edele et al. (2015) angelehnt sein. Es ist geplant, sowohl Korrelationen zwischen Selbsteinschätzungen und Testleistungen zu berechnen, als auch verschiedene Regressionsmodelle, die Selbsteinschätzungen oder Testleistungen als Indikator für Sprachkompetenzen verwenden, jeweils miteinander zu vergleichen. Diese Regressionsmodelle sollen sich auf aktuelle Forschungsfragen zu Determinanten und Auswirkungen von Sprachkompetenzen beziehen. In ein Modell mit Deutschkompetenzen als Outcome könnten beispielsweise kognitive Fähigkeiten, fluchtspezifische motivationale Faktoren (z.B. Aufenthaltsstatus), Teilnahme an Sprachfördermaßnahmen und der Kontakt zur Aufnahmegesellschaft als unabhängige Variablen aufgenommen werden. So kann untersucht werden, ob sich die Ergebnisse von multivariaten Analysen zu aktuellen Themen unterscheiden, in Abhängigkeit davon, welcher Indikator für Deutschkompetenzen verwendet wird.

Die erste Feldphase der ReGES-Studie startet im Winter 2017. Die Ergebnisse für diese Arbeit sind erst gegen Ende nächsten Jahres zu erwarten.

## **Prädiktion der Lese- und Rechtschreibleistung durch spezifisches und generelles orthographisches Wissen**

**Jelena Zaric<sup>1,2</sup>, Telse Nagler<sup>1,2</sup>, Marcus Hasselhorn<sup>1,2,3</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung; <sup>2</sup>Center for Individual Development and Adaptive Education of Children at Risk (IdeA); <sup>3</sup>Johann Wolfgang Goethe-Universität Frankfurt, Fachbereich Psychologie; [jelena.zaric@dipf.de](mailto:jelena.zaric@dipf.de)

Theoretischer Hintergrund: Neben der phonologischen Bewusstheit, die als wichtiger Prädiktor für den Schriftspracherwerb gilt (Vellutino, Fletcher, Snowling & Scanlon, 2004), wird das orthographische Wissen zunehmend mit Lese- und Rechtschreibleistungen in Verbindung gebracht (Deacon, Benere & Castles, 2012). Das orthographische Wissen ist ein multidimensionales Konstrukt, welches das Wissen über erlaubte Buchstabenmuster sowie Regelmäßigkeiten der visuellen und orthographischen Aspekte der Schriftsprache umfasst (Roman, Kirby, Parrila, Wade-Woolley & Deacon, 2009). Es besteht aus dem wortspezifischen (Wissen über die korrekte Schreibweise eines spezifischen Wortes) und dem generellen orthographischen Wissen (Wissen über zulässige Buchstabenmuster; Rothe, Cornell, Ise & Schulte-Körne, 2014). Verschiedene Studien konnten zeigen, dass beide Komponenten des orthographischen Wissens als Prädiktoren für die Lese- und Rechtschreibentwicklung genutzt werden können (z. B. Georgiou, Parrila & Papadopoulos, 2008; Bekebrede, van der Leij & Share, 2009). Es ist jedoch bislang nicht geklärt, inwieweit sich die zwei Komponenten des orthographischen Wissens in der Vorhersagekraft unterscheiden und ob der Anteil aufgeklärter Varianz für die Lese- und der Rechtschreibleistung unterschiedlich groß ist. Ziel dieser Untersuchung ist es daher die Rolle des wortspezifischen und des generellen orthographischen Wissens als Prädiktoren in Bezug auf die Varianzaufklärung von Lese- und Rechtschreibleistungen zu überprüfen.

Methode: Die Daten von 79 Kindern (Alter  $M = 9.46$ ,  $SD = 0.65$ ; weiblich = 36) mit durchschnittlichem IQ ( $M = 107$ ,  $SD = 12.30$ ; ermittelt durch den Zahlen-Verbindungs-Test, ZVT; Oswald & Roth, 1987) wurden in die Analysen miteinbezogen. Die Leseleistungen wurden mit dem Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6; Lenhard & Schneider, 2006), die Rechtschreibleistungen durch den Salzburger Lese- und Rechtschreibtest (SLRT-II; Moll & Landerl, 2010) erfasst. Zur Erfassung des wortspezifischen orthographischen Wissens wurde eine orthographische Wahlaufgabe eingesetzt, bei der zwischen einem bestehenden Wort und einem Pseudohomophon das richtige Wort ausgewählt werden sollte. Das generelle orthographische Wissen wurde anhand von zwei Aufgaben mit generierten Non-word-Paaren in Bezug auf (1) das Wissen über die Häufigkeit von Doppelkonsonanten sowie (2) das Wissen über die legale Position von Doppelkonsonanten erhoben.

Ergebnisse: Die Aufgabe zur Erfassung des wortspezifischen orthographischen Wissens sowie die Aufgabe 1 zur Erfassung des generellen orthographischen Wissens korrelierten signifikant mit den Lese- und Rechtschreibleistungen. Die Aufgabe 2 zur Erfassung des generellen orthographischen Wissens korrelierte signifikant mit den Rechtschreibleistungen. Für die Ermittlung der Vorhersagekraft der Variablen in Bezug auf die Lese- und Rechtschreibleistung wurde eine multiple Regressionsanalyse durchgeführt. Das wortspezifische orthographische Wissen konnte als signifikanter Prädiktor sowohl zur Vorhersage der Leseleistung,  $T = 4.88$ ,  $p \leq .01$ , als auch für die Vorhersage der Rechtschreibleistung,  $T = 5.29$ ,  $p \leq .01$ , identifiziert werden. Ebenfalls zeigte sich die Aufgabe 2 zur Erfassung des generellen orthographischen Wissens als signifikanter Prädiktor für die Rechtschreibleistung,  $T = -2.16$ ,  $p \leq .05$ , allerdings nicht für die Leseleistung. Die Aufgabe 1 zur Erfassung des generellen orthographischen Wissens ergab keine signifikante Varianzaufklärung. Die quadrierten Semipartial-Korrelationen ergaben weiterhin, dass der eigenständige Anteil des wortspezifischen orthographischen Wissens für die Rechtschreibleistung höher ist (21.16%) als für die Leseleistung (19.36%). Ebenfalls zeigte sich bezüglich der Aufgabe 2 zur Erfassung des generellen orthographischen Wissens ein höherer Anteil an erklärter Varianz für die Rechtschreibleistung (3.61%) als für die Leseleistung (1.21%).

Entgegen bisheriger Befunde (z. B. Arab-Moghaddam & Sénéchal, 2001; Rothe et al., 2015) sprechen die vorliegenden Ergebnisse dafür, dass sowohl auf wortspezifischer als auch genereller Ebene des orthographischen Wissens ein größerer Anteil der Varianz für die Rechtschreibleistung als für die Leseleistung erklärt werden kann, jedoch scheint das wortspezifische orthographische Wissen hierbei aussagekräftiger zu sein. Ausgebildete wortspezifische Repräsentationen sowie das Wissen



über legale Buchstabenmuster scheinen daher besonders für den Rechtschreiberwerb wichtige Prädiktoren zu sein und könnten zukünftig als Maße zur frühzeitigen Identifikation von Risikokinder herangezogen werden.

## **Diktionärsbasierte Textanalysen zur Darstellung von Gender, Care & MINT in Schulbüchern**

**Michael Heilemann, Sigrun Schirner, Heidrun Stöger**

Universität Regensburg, Lehrstuhl für Schulpädagogik, Deutschland; [michael.heilemann@ur.de](mailto:michael.heilemann@ur.de)

Theoretischer Hintergrund:

Mediale Präsentationen von weiblichen und männlichen Rollenmodellen haben einen Einfluss auf die Übernahme von Geschlechtsrollenstereotypen und darauf, was als typisch männlich bzw. als typisch weiblich eingeschätzt wird (Oppliger, 2006). Studien weisen darauf hin, dass in unserer Gesellschaft nach wie vor Geschlechtsrollenstereotype von der sozial engagierten, emphatischen Frau sowie vom technisch-naturwissenschaftlich begabten Mann vorherrschen (Rose & Stibane, 2013); entsprechend wird der Care-Bereich als typisch weibliche, der naturwissenschaftliche Bereich als typisch männliche Domäne wahrgenommen (Keller, 2001). Diese Vorstellungen über Care und MINT (Mathematik, Informatik, Naturwissenschaft, Technik) tragen zur geringen Beteiligung von Jungen und Männern im Care-Bereich bzw. von Mädchen und Frauen im MINT-Bereich bei (z.B. Leslie, Cimpian, Meyer & Freeland, 2015).

Schulbücher stellen in einer prägenden Entwicklungsphase nahezu täglich Identifikationsfiguren zur Verfügung (Moser, 2016). Bisherige Studien konnten vor allem auf Ebene der visuellen Darstellung zeigen, dass in Schulbüchern der Care-Bereich überwiegend als weiblich und der MINT-Bereich überwiegend als männlich dargestellt wird (Finsterwald & Ziegler, 2007). Entsprechende quantitative Auswertungen von Schulbüchern auf schriftsprachlicher Ebene fehlen bislang jedoch. Dies ist unter anderem vor dem Hintergrund bedeutsam, dass zunehmend geschlechtsneutrale Abbildungen von Jungen und Mädchen in Schulbüchern Einzug halten, deren Geschlecht erst eindeutig in Textpassagen identifiziert werden kann, die den Abbildungen zugeordnet sind (Pech, 2009).

Fragestellung:

Dieser Beitrag fokussiert die Frage, inwiefern auf lexikalischer Ebene in Schulbüchern ein Gender Bias für die Bereiche Care und MINT nachzuweisen ist. Es wird angenommen, dass Schulbücher (1) mehr Care-Wörter mit weiblicher als mit männlicher Konnotation aufweisen sowie (2) mehr MINT-Wörter mit männlicher als mit weiblicher Konnotation.

Methode:

Als Datengrundlage wurde ein digitales Textkorpus aus aktuell an bayerischen Schulen zugelassenen Schulbüchern erstellt. Konkret wurden 47 Schulbücher (insg. 12172 Seiten) ausgewertet, die derzeit an bayerischen Mittelschulen im Fach Deutsch zugelassen sind.

Die Auswertung des Textkorpus erfolgte mit LIWC (Linguistic Inquiry and Word Count), einem Programm zur automatischen Textanalyse, das in der psychologischen Forschung breiten Einsatz findet (Pennebaker, Boyd, Jordan & Blackburn, 2015). Für die Analyse wurden systematisch umfangreiche Wortlisten erstellt, die Care- und MINT-Wörter mit weiblicher (z.B. Krankenschwester, Informatikerin) bzw. männlicher (z.B. Krankenpfleger, Informatiker) Konnotation umfassen.

Ergebnisse:

Die Ergebnisse zeigen, dass MINT-Wörter mit männlicher Konnotation (91.53%) häufiger auftreten als MINT-Wörter mit weiblicher Konnotation (8.47%);  $\chi^2(1, N=236)=162.78, p<.001$ . Der naturwissenschaftlich-technische Bereich wird demzufolge in Schulbüchern nach wie vor sehr stark als männliche Domäne dargestellt.

Die Ergebnisse zeigen außerdem, dass Care-Wörter mit weiblicher Konnotation (52.44%) häufiger in den Schulbüchern auftreten als Care-Wörter mit männlicher Konnotation (47.56%);  $\chi^2(1, N=3261)=7.75, p<.05$ . Eine differenzierte Auswertung der Ergebnisse für den Care-Bereich zeigt, dass im Bereich Familie und Haushalt mehr weibliche Care-Rollenmodelle dargestellt werden (52.82%);  $\chi^2(1, N=3158)=10.03, p<.05$ , während bei der Darstellung von Care-Berufen mehr männliche Care-Rollenmodelle zu finden sind (59.22%), wobei der Unterschied allerdings nur marginal signifikant ist ( $\chi^2(1, N=103)=3.51, p<.10$ ). Sobald es also um die Erwerbsarbeit geht, scheinen die Männer in ihrer medialen Darstellung in Schulbüchern auch im Care-Bereich wieder in den Vordergrund zu treten.

Die Ergebnisse liefern erste Hinweise auf den Ausprägungsgrad des Gender Bias in der schriftsprachlichen Darstellung von Care und MINT in aktuellen Schulbüchern. In weiteren Analysen soll überprüft werden, ob Männer und Frauen in den verschiedenen Domänen überwiegend in als positiv bzw. als negativ zu bewertenden Kontexten dargestellt werden.

## **Moderiert die Arbeitsbelastung den Zusammenhang zwischen Lernzielen und Lernerfolg von Wissenschaftler(inne)n?**

**Julia Hein<sup>1</sup>, Martin Daumiller<sup>2</sup>, Stefan Janke<sup>1</sup>, Markus Dresel<sup>2</sup>, Oliver Dickhäuser<sup>1</sup>**

<sup>1</sup>Universität Mannheim, Deutschland; <sup>2</sup>Universität Augsburg, Deutschland; [julia.hein@uni-mannheim.de](mailto:julia.hein@uni-mannheim.de)

Theoretischer Hintergrund:

Zur professionellen Bewältigung der vielfältigen Aufgaben in Lehre und Forschung, sind Wissenschaftler(innen) auf fachliche und didaktische Kompetenzen angewiesen. Die kontinuierliche Weiterentwicklung der Kompetenzen von Wissenschaftler(inne)n ist sowohl als Voraussetzung für den Lernerfolg von Studierenden durch qualitativ hochwertige Lehre (Biggs & Tang, 2011) als auch weitreichender für den gesellschaftlichen Fortschritt durch gute Forschung von hoher Bedeutung. Dabei ist davon auszugehen, dass motivationale Variablen auf Seite der Wissenschaftler(innen) maßgeblich die Entwicklung des Kompetenzerwerbs beeinflussen (Pintrich, 2004). Studien im schulischen Kontext zeigen, dass selbstbezogene Ziele von Lehrkräften mit dem persönlichen Lernzuwachs in Zusammenhang stehen. Dabei sind insbesondere Lernziele, also das persönliche Streben nach Kompetenzerweiterung, positiv mit der Inanspruchnahme von kompetenzförderlichen Fortbildungen assoziiert (Nitsche, Dickhäuser, Fasching & Dresel, 2013). Gleichzeitig zeigt allerdings die Metaanalyse von Payne, Youngcourt und Beaubien (2007) beachtliche Schwankungen des Zusammenhangs zwischen selbstbezogenen Zielen und dem Lernverhalten Erwachsener. Moderatoren, welche die Enge des Zusammenhangs erklären, wurden in der bisherigen Forschung bisher noch nicht ausreichend untersucht. Wir gehen davon aus, dass solche Moderatorvariablen gerade im Bereich der Hochschullehre eine wichtige Rolle spielen. Das Tätigkeitsprofil von Wissenschaftler(inne)n ist vielfältig und zeichnet sich dadurch aus, dass Aufgaben in Lehre, Forschung und Verwaltung im Arbeitsalltag miteinander konkurrieren. Hierdurch entsteht ein Ressourcenkonflikt, welcher besonders dann zu Tage treten

sollte, wenn die wahrgenommene Arbeitsbelastung und die in der Folge eintretende emotionale Erschöpfung (Metz & Rothe, 2001) hoch sind. In Einklang mit dem Modell der Selbstregulation (Schmitz & Schmidt, 2007) ist davon auszugehen, dass die eigenen Ziele bei verringerten emotionalen und energetischen Ressourcen seltener verfolgt werden und sich weniger stark auf das Lernverhalten und Lernergebnis auswirken können. Es ist daher anzunehmen, dass die emotionale Erschöpfung auch beeinflusst, in welchem Ausmaß Wissenschaftler(innen) ihren Lernzielen nachgehen können. In Folge würden Lernziele in einem Zustand starker emotionaler Erschöpfung auch weniger verhaltenswirksam als bei niedrig ausgeprägter emotionaler Erschöpfung. Entsprechend postulieren wir, dass der positive Zusammenhang zwischen Lernzielen und dem persönlichen Lernzuwachs bei einer höheren emotionalen Erschöpfung abgeschwächt sein sollte. Dieser Moderationseffekt der emotionalen Erschöpfung wird kontextübergreifend in der Lehre und Forschung erwartet und in der vorliegenden Studie überprüft.

Methode.

Die Fragestellung wurde anhand der Befragung einer repräsentativen Zufallsstichprobe von Wissenschaftler(inne)n (N = 609) in Deutschland untersucht. Zur Erfassung der Lernziele in Lehre und Forschung wurde ein gut validiertes Messinstrument zur Erfassung von selbstbezogenen Zielen bei Wissenschaftler(inne)n eingesetzt (Daumiller & Dresel, 2017). Um das Belastungserleben der Wissenschaftler(innen) zu erfassen, wurde die Skala emotionale Erschöpfung einer leicht modifizierten Version des MBI-GS-D (Büssing & Glaser, 1998) verwendet. Der Lernzuwachs wurde jeweils für Lehre und Forschung anhand eines, in einer vorherigen Studie entwickelten und validierten, Selbstberichtsverfahren erhoben. Die Probanden schätzten ein, in welchem Umfang sie ihre Fach- und Methodenkompetenzen in den letzten sechs Monaten weiterentwickelt haben. Zur messfehlerbereinigten Überprüfung der Fragestellung wurden latente Interaktionsanalysen nach Wang und Wang (2012) durchgeführt.

Ergebnisse.

Es wurde, wie erwartet, sowohl im Lehr- als auch im Forschungskontext ein signifikanter positiver Effekt der Lernziele auf das Lernausmaß gefunden (Lehre:  $\beta = 0.27$ , 90% CI [0.20, 0.34],  $p < .001$ ; Forschung:  $\beta = 0.27$ , 90% CI [0.20, 0.33],  $p < .001$ ). Latente Moderationsanalysen zeigten, dass es im Lehrkontext ( $\beta = -0.07$ , 90% CI [-0.13, -0.01],  $p = .029$ ) und Forschungskontext ( $\beta = -0.09$ , 90% CI [-0.17, -0.01],  $p = .034$ ) signifikante Interaktionseffekte der emotionalen Erschöpfung und Lernziele der Wissenschaftler(innen) auf das Lernausmaß gibt. Die Richtungen der konditionalen Effekte waren in beiden Kontexten hypothesenkonform. Bei höheren Werten auf der Skala zur emotionalen Erschöpfung fiel der positive Zusammenhang zwischen den Lernzielen der Wissenschaftler(innen) und dem Lernausmaß schwächer aus als bei niedrigen Werten.

## Wie begründen Vorschulkinder Ihre Entscheidungen beim Planen von naturwissenschaftlichen Experimenten?

Heidi Haslbeck<sup>1</sup>, Eva-Maria Lankes<sup>1</sup>, Lucia Kohlhauf<sup>2</sup>, Eva S. Fritzsche<sup>1</sup>, Birgit Jana Neuhaus<sup>2</sup>

<sup>1</sup>Technische Universität München, Deutschland; <sup>2</sup>Ludwig-Maximilians-Universität München, Deutschland; [heidi.haslbeck@tum.de](mailto:heidi.haslbeck@tum.de)

Theoretischer Hintergrund

Die Förderung des wissenschaftlichen Denkens ist ein Ziel des naturwissenschaftlichen Unterrichts. Zum wissenschaftlichen Denken gehört u.a. Wissen über die Schritte des wissenschaftlichen Erkenntnisprozesses und ihre Anwendung (Arnold, 2015). Mayer (2007) unterscheidet vier Teilkompetenzen in der Anwendung des wissenschaftlichen Erkenntnisprozesses: Fragestellungen formulieren, Hypothesen generieren, Experimente planen/durchführen und Daten auswerten. In der vorliegenden Studie befassen wir uns mit der Teilkompetenz „Experimente planen/durchführen“ und untersuchen, wie die Variablenkontrollstrategie als Teil des wissenschaftlichen Denkens bei der Planung von Experimenten berücksichtigt wird.

Entwicklungspsychologische Studien zeigen, dass bereits Kinder im Vor- und Grundschulalter über erste Fähigkeiten zum wissenschaftlichen Denken verfügen und zwischen einfachen konfundierten und unkonfundierten Experimenten unterscheiden können (Sodian & Mayer, 2013; Saffran, 2016). Während bei konfundierten Experimenten mehrere Variablen gleichzeitig variiert werden, wird bei unkonfundierten Experimenten immer nur eine Variable verändert. Wenig wissen wir aber darüber, ob die Kinder bei ihrer Entscheidung das Argumentationsschema des wissenschaftlichen Denkens bzw. der Variablenkontrollstrategien anwenden oder auf andere Argumentationsschemata zurückgreifen (Walton, Reed & Macagno, 2008). Theoretisch ist es möglich, dass die Kinder mit dem richtigen Argumentationsschema der Variablenkontrollstrategie zur richtigen Lösung kommen. Denkbar ist aber auch, dass sie die richtige Lösung aus anderen Gründen oder durch Zufall wählen.

Die allgemeinen kognitiven Fähigkeiten ebenso wie das individuelle Interesse der Kinder spielen eine Rolle für Lernmotivation und Lernleistung (Schiefele, 2008). Sie beeinflussen damit auch die Entwicklung der Argumentationsschemata.

Fragestellungen

1. Welchen Zusammenhang haben kognitive Fähigkeiten und das Interesse an Naturwissenschaften mit der „Fähigkeit Experimente zu planen“?
2. Inwiefern erklären kognitive Fähigkeiten, Interesse und die Verwendung verschiedener Argumentationsschemata die Richtigkeit der Antworten?

Methode

Insgesamt nahmen 17 Vorschulkinder aus 4 Kindergärten an der Studie teil. Die Kinder waren im Durchschnitt 5.5 (SD = 0.63) Jahre alt.

Auf Basis bereits vorhandener Testinstrumente wurde ein Interviewleitfaden entwickelt, mit dem die Kinder befragt wurden (Saffran, 2016; Schwichow, Zimmerman, Croker & Härtig, 2016). Den Kindern wurden 4 Aufgaben mit Auswahlmöglichkeiten zur Planung einfacher Experimente gestellt (z.B. zu der Frage, ob Eiswürfel im warmen Wasser schneller schmelzen als im kalten) und nach jeder Antwort wurde nach einer Begründung gefragt. Zur Bewertung der „Fähigkeit, Experimente zu planen“ wurde die Richtigkeit der Antworten codiert (21 Items,  $\alpha = .74$ ). Die Begründungen der Kinder befinden sich noch in der Auswertung. Sie werden auf der Grundlage eines Codierschemas unterschiedlichen Argumentationsschemata zugeordnet (Walton et al., 2008).

Die allgemeinen kognitiven Fähigkeiten der Kinder wurden mit der Skala Reihenfortsetzen des CFT 1-R (Weiß & Osterland, 2013) erhoben. Das Interesse der Kinder an den Naturwissenschaften wurde über die Erzieher und Erzieherinnen erfasst. Die Erzieher und Erzieherinnen füllten dazu für jedes Kind eine Skala mit 13 Items ( $\alpha = .95$ ) zum Interesse der Kinder an den Naturwissenschaften (veränderte Version des KOMPIK, 2012) aus.

Alle Kinder wurden von einer geschulten Interviewerin einzeln befragt. Das Interview dauerte insgesamt 15-25 Minuten und war für alle Kinder gut durchführbar.

#### Ergebnisse

Der nonparametrische positive Zusammenhang zwischen kognitiven Fähigkeiten ( $r = .45, p = .031$ ) und der Fähigkeit, Experimente zu planen sowie dem Interesse ( $r = .50, p = .029$ ) und der Fähigkeit, Experimente zu planen, zeigt, dass Kinder mit besseren Lernvoraussetzungen Vorteile haben, zumindest, wenn es um die richtige Planung von Experimenten geht. Dennoch bleibt ein Teil der Fähigkeit, Experimente zu planen, der sich nicht durch allgemeine kognitive Fähigkeiten erklären lässt. Fraglich ist, inwiefern Kinder die richtige Planung von Experimenten tatsächlich wissenschaftlich mit der Variablenkontrollstrategie begründen können oder ob die richtige Planung der Experimente beispielsweise auf Nachahmung beruht, Kindern aber keine wissenschaftliche Argumentation dazu gelingt. Aktuell werden die Begründungen der Kinder zu ihrer Planung von Experimenten unterschiedlichen Argumentationsschemata zugeordnet. Auf der GEBF werden die Ergebnisse hierzu im Detail vorgestellt.

## **Entwicklung und Validierung eines Kurzfragebogens zur Erfassung der motivationsrelevanten Aspekte im Unterricht**

**Madana Eiche, Gerlinde Lenke**

Universität Koblenz-Landau, Deutschland; [eiche@uni-landau.de](mailto:eiche@uni-landau.de)

#### Theoretischer Hintergrund

Motivation ist für den Lernprozess und die Lernentwicklung von großer Bedeutung (Deci & Ryan, 1993; Schiefele & Streblov, 2006; Wild & Möller, 2009) und stellt einen Prädiktor für die Schulleistung dar (Steinmayr & Spinath, 2007). Da die Motivation im Verlauf der Schullaufbahn (Freiberger & Spinath, 2014; Gottfried et al., 2001) und im Vergleich zur letzten Dekade (Spinath et al., 2016) stetig abnimmt, hat dieses Forschungsfeld weiterhin eine hohe Relevanz für die Bildungsforschung.

Zur Erfassung der Motivation von SchülerInnen im Unterricht werden verschiedene Fragebögen eingesetzt (u.a. Bürger, 2013; Rakoczy et al., 2005; Rheinberg et al., 2001; Spinath et al. 2012). Der Fokus dieser Instrumente liegt dabei häufig auf den Bereichen der Lern- und Leistungsziele (z.B. SELLMO) und der individuellen Einstellung oder weist in der Formulierung der Items eine hohe Aufgaben- oder Fachspezifität (z.B. FAM) auf. Ein Fragebogen zur differenzierten und ganzheitlichen Erfassung motivationsrelevanter Aspekte im Unterricht existiert nicht, dazu soll in der vorliegenden Arbeit ein Kurzfragebogen auf Grundlage der existierenden Skalen zum IBAS-Modell (Drexler et al., eingereicht) entwickelt und validiert werden. Das IBAS-Modell wurde in Anlehnung an das Target-Modell (Ames, 1992), die Selbstbestimmungstheorie (Deci & Ryan, 1993) und die aktuelle Literatur (z.B. Rakoczy, 2008) konzipiert und beschreibt explizit motivationsunterstützendes Handeln der Lehrkraft sowie dafür relevante Kontextfaktoren.

Die Gestaltung des Unterrichts durch die Lehrkraft steht in engem Zusammenhang mit der Motivation der SchülerInnen (Helmke, 2015; Kunter et al., 2011) und sollte demnach ebenso erhoben werden.

#### Fragestellung

Stellt der Kurzfragebogen ein geeignetes Messinstrument zur Erfassung motivationsrelevanter Aspekte im Unterricht dar?

- Ermöglichen die Skalen eine reliable Messung? (interne Konsistenz der Skalen, Urteilsübereinstimmung)
- Inwiefern lassen sich die theoretisch angenommenen motivationsrelevanten Aspekte empirisch für spezifische Unterrichtssituationen abbilden? (faktorielle Validität)

#### Methode

Zur Validierung des Fragebogens werden 280 SchülerInnen einbezogen. Die Erhebung erfolgt im Sekundarschulbereich in den Klassenstufen 5 bis 10. Der Fragebogen erfasst vier Dimensionen motivationsförderlichen Unterrichtshandelns: Inhalts-, Autonomie-, Bewertungsdimension sowie Soziale Dimension. Jede Dimension wird mit zwei bis neun Items aus den 11 Subskalen des IBAS-Modells (z.B. Adäquate Freiräume schaffen) über eine unipolare fünfstufige Likert-Skala (von 0 stimmt gar nicht bis 4 stimmt genau) abgefragt.

Erste Analysen (Reliabilitätsmessung, Standardabweichungen, Trennschärfe, Varianz, Itemschwierigkeit) dienen der inhaltlichen Überprüfung der Items. Redundante Items, zu leichte oder zu schwierige Items, Items mit mangelnder Homogenität und Reliabilität werden ausgeschlossen. Für die weiteren Analysen wird als Reliabilitätsmerkmal McDonald's Omega und als Anzeichen der Urteilerübereinstimmung die Intra-Klassenkorrelation ermittelt. Zur Validierung soll die faktorielle Struktur durch eine konfirmatorische Faktorenanalyse überprüft werden.

#### Ergebnisse

Nach entsprechender Itemanalyse werden neun Items entfernt. Bei der Überprüfung der internen Konsistenz zeigen sich für die Skalen Inhalt und Bewertung eine gute Reliabilität ( $\omega = .82 - .84$ ), während die Skalen für die Autonomie und Soziale Dimension unterhalb der akzeptablen Grenze ( $\omega = .55 - .56$ ) liegen. Deswegen sollen beide Skalen angepasst werden.

Im Antwortverhalten der SchülerInnen zeigen sich mittlere bis hohe Unterschiede auf Klassenebene ( $.07 \leq ICC1 \leq .40$ ). Die Genauigkeit der Urteile ist für die Dimension Autonomie im akzeptablen ( $ICC2 = .63$ ) und für die weiteren Dimensionen im guten Bereich ( $.81 \leq ICC2 \leq .92$ ).

Der Modellfit nach der Durchführung der CFA weist akzeptable bis gute Werte auf (ULS,  $\chi^2 = 489.105$ ,  $df = 203$ ,  $CFI = .976$ ,  $TLI = .973$ ,  $RMSEA = .078$ ,  $SRMR = .061$ ). Zur Überprüfung, ob das Modell für alle Altersstufen abgebildet werden kann, werden Subgruppen gebildet. Die Berechnung für die Daten im Sample Klasse 5-7 und Klasse 8-10 zeigen einen Modellfit, der nicht stark von dem der Gesamtstichprobe abweicht und damit auch für die Gruppierungen eine Passung darstellt.

Perspektivisch soll der Fragebogen ein geeignetes Instrument zur Evaluation einer Unterrichtsstunde hinsichtlich der motivationsrelevanten Aspekte aus Sicht der SchülerInnen darstellen. Durch diese Evaluation soll eine Weiterentwicklung des Unterrichts gewährleistet werden (Helmke et al. 2013).

## **Stehen Persönlichkeitsveränderungen in Zusammenhang mit Bildungsentscheidungen und -aspirationen von Auszubildenden und Studierenden?**

**Laura Braun, Nele Usslepp, Benjamin Nagengast, Marion Spengler**

Eine Vielzahl an Studien zeigte den Einfluss von Persönlichkeitseigenschaften im Bildungssystem – so hängen die Big Five (Extraversion, Neurotizismus, Gewissenhaftigkeit, Offenheit, und Verträglichkeit) beispielsweise mit dem Bildungserfolg (z.B., Poropat, 2009) und Bildungsentscheidungen (Usslepp et al., in prep.) zusammen. Während dem jungen Erwachsenenalter scheint die Entwicklung der Persönlichkeit vor allem dem Reifungsprinzip zu folgen (Roberts & Wood, 2006), wonach junge Menschen gewissenhafter, verträglicher und weniger neurotisch werden. Dieses Muster zeigte sich sowohl bei Studierenden, als auch bei Auszubildenden (Lüdtke, Roberts, Trautwein & Nagy, 2011). Nach der Social Investment Theorie liegen Gründe für Persönlichkeitsveränderungen vor allem in der Identifikation und Adaption neuer sozialer Rollen (Roberts, Wood & Smith, 2005). Dies spielt während der Ausbildung oder dem Studium eine besonders große Rolle, da Herausforderungen in gesellschaftlichen, akademischen, und identitären Dimensionen bewältigt werden müssen. Bisherige Forschung hat bisher nicht untersucht, ob die Anpassung an diese Herausforderungen mit Bildungsergebnissen zusammenhängt. Zwar lieferte eine Untersuchung von Bleidorn (2012) erste Hinweise darauf, dass Persönlichkeitsveränderungen mit einer Erhöhung der Leistungsanstrengung zusammenhängen, so fehlen jedoch Befunde die sich dem Einfluss auf Aspirationen und Entscheidungen des Bildungsverlaufs annehmen. Demnach ist es denkbar, dass eine Entwicklung in Richtung des Reifungsprinzips (z.B. Erhöhung von Gewissenhaftigkeit) während des Studiums oder der Ausbildung auch mit anspruchsvolleren Bildungsaspirationen und Bildungsentscheidungen zusammenhängt, unabhängig vom Startpunkt des Reifens. Die vorliegende Studie nimmt sich dieser Forschungslücke an und ergänzt damit bisherige Studien zur Persönlichkeitsveränderung von Auszubildenden und Studierenden (z.B., Lüdtke et al., 2011). Dabei untersuchen wir zwei Forschungsfragen:

- 1) Beeinflusst die Persönlichkeitsveränderung, ob Personen nach einer abgeschlossenen Ausbildung noch ein Studium anstreben bzw. noch studieren werden?
- 2) Beeinflusst die Persönlichkeitsveränderung die Berufsaspirationen von Studierenden?

Diese Fragestellungen wurden anhand der TOSCA (Transformation des Sekundarschulsystems und akademische Karrieren; Köller, Watermann, Trautwein & Lüdtke, 2004) untersucht, die Daten von Abiturientinnen und Abiturienten aus Baden-Württemberg in zweijährigem Abstand über 10 Jahre erfasst hat. In dieser Arbeit wurden die Daten von N=326 Auszubildenden und N=1.759 Studierenden zu den ersten drei Messzeitpunkten genutzt. Persönlichkeit wurde anhand des NEO-FFI erfasst (Borkenau & Ostendorf, 1993) und deren Veränderung vom ersten zum zweiten Messzeitpunkt (2 Jahre) anhand von Latent-Change Scores modelliert. Die abhängigen Variablen wurden zum dritten Messzeitpunkt erfasst, wobei auf Seiten der Auszubildenden die Studienaspiration sowie der tatsächliche Übergang in das Studium und in Bezug auf die Studierenden die Berufsaspiration bezüglich unterschiedlicher Subskalen (Leistung, extrinsische Orientierung, intrinsische Orientierung, Autonomie, Aufstiegsmöglichkeiten) herangezogen wurde (Eigenentwicklung in Anlehnung an Seifert & Bergmann, 1983). Im Rahmen der statistischen Analyse wurden lineare und logistische Regressionsmodelle durchgeführt, die jeweils für das Ausgangsniveau in der entsprechenden Persönlichkeitseigenschaft, den sozioökonomischen Status, das Geschlecht, sowie den IQ kontrollieren.

In Bezug auf die erste Forschungsfrage deuten erste Ergebnisse darauf hin, dass die Veränderung von Gewissenhaftigkeit und Offenheit, die Wahrscheinlichkeit erhöht, dass Auszubildende ein Studium anstreben oder beginnen. Im Rahmen der zweiten Forschungsfrage weisen die Ergebnisse darauf hin, dass bei Studierenden vor allem die Veränderung von Gewissenhaftigkeit mit der Aspiration nach leistungs- und prestigereichen Berufen zusammenhängt, wohingegen beispielweise die Veränderung in Extraversion und Offenheit in Zusammenhang mit dem Wunsch nach persönlicher Verwirklichung der Arbeit steht und die Veränderung in Verträglichkeit und Neurotizismus mit dem Ziel nach Autonomie zusammenhängt.

Die Ergebnisse legen nahe, dass die Persönlichkeitsveränderung von jungen Menschen während den ersten beiden Jahren nach dem Abitur in Zusammenhang mit weiteren Bildungsentscheidungen, sowie Berufsaspirationen steht. Vor allem die Veränderung von Gewissenhaftigkeit als Indikator für leistungsrelevante Entscheidungen, bestärkt das bisherige Befundmuster zur besonderen Rolle von Gewissenhaftigkeit für Bildungsergebnisse (siehe Poropat, 2009). Weitere Forschung ist notwendig, die den Einfluss der Persönlichkeitsentwicklung auf den tatsächlichen Übergang in den Beruf in den Fokus nimmt. Darüber hinaus werden die Ergebnisse vor dem Hintergrund der Besonderheit der Stichprobe und methodischer Herausforderungen diskutiert.

### **Absences of students during class: The effect of problematic computer game usage on students' cognitive and physical lesson avoidance – results of a latent path model**

**Arvid Nagel, Horst Biedermann**

Pädagogische Hochschule St.Gallen, Schweiz; [arvid.nagel@phsg.ch](mailto:arvid.nagel@phsg.ch)

Theoretical Framework and Aim:

In the past years, the phenomenon of lesson avoidance or even school absenteeism of students has aroused an increasing interest both in science and in media. Nevertheless, research has largely failed to keep up with these developments in offering theoretically sound models or empirical evidence for understanding students' lesson avoidance yet. This study aims to predict cognitive and physical lesson avoidance of students by problematic computer game usage.

Previous research on lesson avoidance focused predominantly on its antecedents, which are seen in factors of the family, school and peers (see Attwood & Croll, 2006; Claes, Hooghe & Reeskens, 2009; Dunkake, 2006, 2010; Sälzer & Lenski, 2016). Within that body of research, aspects such as day dreaming, turning off or disengaging mentally during lessons were not addressed – research did focus on physical avoidance exclusively (see Sälzer et al., 2012; Vaughn et al., 2013; Veenstra et al., 2010; Wagner & Dunkake, 2009). Furthermore, the effect of interactive media on students' lesson avoidance remains unanswered (Ricking, Schulze & Wittrock, 2009). Screen media such as game consoles, computer, smartphones and the internet play a very central and significant role in the lives of young people as many empirical studies report (see Feierabend, Plankenhorn, & Rathgeb, 2016). With the increasing availability of diverse kinds of media to youths, their problematic media use skyrocketed as well. Many empirical studies show the relevance of computer game addiction in adolescence (see Carbonell et al., 2009; Choo et al., 2014; King & Delfabbro, 2014; Kuss & Griffiths, 2012; Mößle & Rehbein, 2013; Rehbein & Baier, 2013; Rehbein, Kleimann & Mößle, 2010b; Rehbein et al., 2015; Rehbein & Mößle, 2013; Sim et al., 2012). Within this poster, we aim to reveal what effects (problematic) computer game use has on students' lesson avoidance. In contrast to previous studies, thereby we construe lesson avoidance in a broader scope to encompass not only physical but also cognitive avoidance (such as inattentiveness or daydreaming).

Research Questions:

Within the present study, the following research questions were addressed:

1. (How) Can we measure cognitive lessons avoidance behavior of students?
2. Can we identify effects of computer game use on cognitive and physical lessons avoidance behavior of students?

Sample and Methods:

The study was conducted based on a German student sample, attending four community schools in Schleswig-Holstein (federal state in Germany) from grades 7 to 10 (N=578 | M=14.3 years, SD=1.3). To assess cognitive lesson avoidance we developed a new four-item scale, and to measure physical lesson avoidance we developed two items. To assess the (problematic) use of computer game we used the "computer game dependence scale" of Rehbein, Kleimann and Mößle (2010a). Data analysis was performed using latent path analysis.

Findings:

The one-dimensionality of the scale students' cognitive lesson avoidance has been tested and confirmed empirically by means of a confirmatory factor analysis. The measurement model (general factor model) has a very good model to fit ( $\chi^2=1.76$ ; df=2, p=.42; CFI=1.000; RMSEA=.000; SRMR=.013). The measurement model has high factor loadings of  $\lambda=.65$  to  $\lambda=.84$  for the four indicators and the standard errors vary between .04 and .06. Internal consistency of the scale was high ( $\alpha=.87$ ).

We utilized structural equation modeling to test the effect of (problematic) computer game usage on physical lesson avoidance via cognitive lesson avoidance. Results show that computer game usage significantly predicts students' cognitive lesson avoidance ( $\beta_1=.33$ ,  $p<.001$ ), which in turn predicted physical lesson avoidance ( $\beta_2=.24$ ,  $p<.001$ ). The indirect effect was also significant (beta indirect=.08,  $p<.01$ ). When introducing the mediator (i.e. cognitive lesson avoidance), the direct effect from computer game usage on physical lesson avoidance vanished ( $\beta_3=.07$ ,  $p>.05$ ).

### **Professionelles Handeln im Unterricht: – Ein Konzept zur Förderung der Data Literacy in der Lehrkräftebildung**

**Katharina Thoren**

Freie Universität Berlin, Deutschland; [katharina.thoren@fu-berlin.de](mailto:katharina.thoren@fu-berlin.de)

Nach den Standards für die Lehrerbildung der Kultusministerkonferenz (2004) sollen Lehrkräfte ihre Entscheidungen nach Möglichkeit auf empirische Evidenz und eine nachgewiesene Wirksamkeit gründen. Im Kontext der Schul- und Unterrichtsentwicklung ist dieser Standard unmittelbar an die Data Literacy (Mandinach & Gummer, 2016), also die Erfassung, Analyse und das Verstehen von sowie das Entscheiden auf Grundlage empirischer Daten gekoppelt. Im Rahmen des Teilprojekts „Forschungskompetenzen für eine evidenzbasierte adaptive Unterrichtspraxis“ (Gesamtprojekt: K2teach) der BMBF Qualitätsoffensive Lehrerbildung wurde daher u.a. ein Seminarkonzept zur Förderung der Kompetenz der datengestützten Unterrichtsentwicklung entwickelt. Im Wintersemester 2016/17 wurde das Konzept in vier Seminaren begleitend zum Praxissemester erstmalig an der Freien Universität Berlin angeboten und von insgesamt 55 Studierenden besucht. Die praktische Auseinandersetzung mit Instrumenten zur datengestützten Schul- und Unterrichtsentwicklung (VERA-Rückmeldungen, Aufgabendatenbank, Selbstevaluationsportal) stellt den inhaltlichen Schwerpunkt dar. Diese Instrumente werden vom Institut für Schulqualität der Länder Berlin und Brandenburg (ISQ) angeboten und flächendeckend an Berliner Schulen eingesetzt. Somit bindet das Seminar die universitäre Ausbildung direkt an die Qualitätssicherung in der Schulpraxis an. Die Umsetzung des Seminarkonzepts wird dabei umfassend, sowohl formativ als auch summativ, evaluiert.

Zur Qualitätssicherung sowie zur gezielten (Weiter-)Entwicklung des Seminars wurde während der Konzeptionsphase das Modell der formativen Evaluation von Brown und Gerhardt (2002) angewendet. Das Vier-Phasen-Modell sieht einen Kanon interaktiver Feedbackschleifen unter Einbeziehung verschiedener Akteur\*innen vor. Dazu gehörte nicht nur die Kooperation mit verschiedenen Beteiligten der Lehrkräftebildung sowie Expert\*innen aus der Schulpraxis, sondern auch die Implementationskontrolle in Form von intensiven Vor- und Nachbesprechungen jeder Seminarsitzung mit allen Dozierenden. Weiterhin, als Teil der formativen Evaluation wurde für die Analyse intraindividuelle Veränderungen in ausgewählten motivationalen Merkmalen der Studierenden ein Messwiederholungs-Design implementiert, mit dem die Variation der Instruktionsevaluation aus Sicht der Studierenden sowie deren Lernmotivation im Seminarverlauf abgebildet werden. Zu diesem Zweck wurde ein Kurzfragebogen entwickelt und Skalen zu Interesse, Kontrollverhalten der/des Dozierenden, Unterstützung kognitiver Autonomie, autonomieunterstützendes Klima (Tsai, Kunter, Lüdtke, Trautwein & Ryan, 2008) sowie zur Einschätzung der Situation im Unterricht (Goetz, Lüdtke, Nett, Keller & Lipnevich, 2013) adaptiert.

Für die summative Evaluation wurde ein Prä-Post-Design wie folgt implementiert: Um motivationale Veränderungen der Studierenden zu messen, wurde ein Instrument konzipiert, welches die Einstellungen zur Datennutzung im Unterricht (in Anlehnung an das Modell von Mandinach & Gummer, 2016 unter Berücksichtigung der Wert-Erwartungs-Theorie von Eccles und Wigfield 2002) entwickelt. Die Einstellungen zum Nutzen der Vergleichsarbeiten (Richter & Böhme, 2014). Weiterhin enthielt der Fragebogen Fragen zu soziodemographischen Merkmalen, dem bisherigen Studienverlauf und zu Studienleistungen sowie dem Ausmaß berufspraktischer Erfahrungen

Erste Ergebnisse des Prä-Post-Designs zeigen, dass im Verlauf des Semesters die Selbstwirksamkeitserwartungen (expectation of success, Eccles & Wigfield, 2002) der Studierenden bezüglich konkreter Kompetenzfacetten der Data Literacy bedeutsam gesteigert, sowie Einstellungen bezüglich des Aufwandes (relative cost; ebd.) dieser Facetten bedeutsam gesenkt werden konnten (Cohen's d > .5). Weiterhin konnten die Einstellungen der Studierenden bezüglich des Nutzens der Vergleichsarbeiten zum Teil deutlich verbessert werden (Cohen's d > .8). Die Ergebnisse werden vor dem Hintergrund aktueller Erkenntnisse zur Data Literacy bei Lehrkräften diskutiert.

### **Hochbegabte im Beruf. Berufsbiographien, Positionen und Berufszufriedenheit.**

**Maren Schlegler<sup>1</sup>, Eveline Wuttke<sup>1</sup>, Kristina Kögler<sup>1,2</sup>**

<sup>1</sup>Goethe-Universität Frankfurt, Deutschland; <sup>2</sup>Universität Hohenheim, Deutschland; [maren.schlegler@gmx.de](mailto:maren.schlegler@gmx.de)

Theoretischer Hintergrund und Fragestellung:

Hochbegabung wird stereotyp als Synonym für absonderliches Verhalten verwendet (Horowitz, 1994). Eine allgemeingültige Definition von Hochbegabung existiert nicht (Baudson 2016), meist wird das Vorhandensein hoher Intelligenz betont (Preckel 2010). Der quantitative Ansatz bezeichnet Personen mit einem IQ von mindestens 130 als hochbegabt (Sparfeldt 2006). Studien über hochbegabte Erwachsene und ihre berufliche Situation sind rar und zeigen heterogene Befunde (Sparfeldt 2009).

Eine der wenigen Studien, die Längsschnittstudie von Terman aus den 1920er Jahren zeigt, dass sich hochbegabte Erwachsene hinsichtlich Erfolg, Gesundheit und Lebenszufriedenheit nicht von Normalbegabten unterscheiden (Freeman 2010). Hossiep et al. (2013) berichten, dass Hochbegabte eine schwächere Teamorientierung, Kontaktfähigkeit und Führungsmotivation, dafür eine höhere Gestaltungsmotivation aufweisen. An anderer Stelle wird auf nicht-lineare Bildungsbiographien und Schwierigkeiten Hochbegabter im sozialen Umfeld am Arbeitsplatz sowie die Bevorzugung flacher Hierarchien und selbstständiger Arbeitsumfelder mit hinreichenden Gestaltungsfreiräumen hingewiesen (Groll 2011; Hussla 2010). Die empirische Evidenz dafür ist bisher übersichtlich.

Ziel des Beitrags ist es, die berufliche Situation Hochbegabter zu beschreiben. Im ersten Schritt werden Berufsbiographien, berufliche Positionen und die Berufszufriedenheit betrachtet. Folgende Fragen stehen im Fokus: (1) Welche Bildungswege und Berufsbiographien verfolgen Hochbegabte? (2) Welche beruflichen Positionen nehmen Hochbegabte in welchen Berufsfeldern und Beschäftigungsformen ein? (3) Wie zufrieden sind Hochbegabte mit ihrer beruflichen Situation?

Methode:

Das Forschungsprojekt besteht aus drei Untersuchungsteilen: (1) quantitative Vorstudie (Fragebogen), (2) qualitative Hauptstudie (Interviews), (3) Instrumentenentwicklung zur Erfassung des subjektiven Erlebens der Berufstätigkeit von Hochbegabten in einer größeren Nachfolgestudie. Dieser Beitrag berichtet Befunde der Vorstudie.

Stichprobe: N=277 Mitglieder von Mensa in Deutschland e.V. (IQ von mindestens 130). Die Mitglieder wurden im Sommer 2016 angeschrieben und um ihre Teilnahme an der onlinebasierten Fragebogenerhebung gebeten (Rücklaufquote 28%).

Instrumente: Der Fragebogen besteht aus fünf Teilen zu soziodemografischen Daten, Berufsbildung, Berufsstatus, Karrierepfaden und Berufszufriedenheit. Die Skalen (Vorgesetzte, Tätigkeit, Bezahlung, Belastung, Kollegen, Allgemeine Berufszufriedenheit) zur Berufszufriedenheit (Merz 1979; Likert-Skala von 1 „trifft gar nicht zu“ bis 6 „trifft vollständig zu“) offenbaren gute Reliabilitäten zwischen .77 und .95. Die Teilnehmer (46,6% männlich) waren im Durchschnitt 42,9 Jahre alt (SD=9,82).

Ergebnisse:

(1) Ausbildung und beruflicher Werdegang: Die Stichprobe ist überdurchschnittlich gebildet (86,3% Abitur, 33,4% Universitätsabschluss, 23,5% andere Form des Hochschulabschlusses (z.B. FH), 10,4% promoviert). 22,7% durchliefen eine duale Ausbildung. Nicht-lineare Bildungswege zeigen sich darin, dass 14,4% mehr als einen Berufsabschluss und 11,9% drei und mehr Studienabschlüsse aufweisen.

(2) Berufliche Positionen, Beschäftigungsformen: Die meisten Teilnehmer arbeiten in den Bereichen IT und Naturwissenschaften (29%), Sozialwesen (13%) und Forschung (13%). 70% sind vollzeitbeschäftigt, 17,3% arbeiten Teilzeit und 17 Teilnehmer (6,1%) sind ohne Arbeit. 81,3% sind abhängig beschäftigt, 33,9% geben an, selbstständig tätig zu sein. In der Privatwirtschaft arbeiten 25,3% für Großunternehmen, 56,4% für kleine oder mittelständische Unternehmen. Durchschnittlich hatten die Teilnehmer bisher zwei Arbeitgeber, die Spanne reicht bis 20 (SD= 2,455). 26,8% (davon 52,8% männlich) arbeiten in Führungspositionen mit Verantwortung für bis zu 650 Mitarbeiter.

(3) Berufszufriedenheit: Hochbegabte Erwachsene sind im Mittel zufrieden mit ihrer Tätigkeit (MW=4,21, SD=1,37), empfinden sie aber nicht durchwegs als anregend (MW=3,3, SD=1,463). Sie kommen gut mit ihren Kollegen zurecht (MW=4,57, SD=1,039), haben aber kaum enge Freunde unter ihnen (MW=2,76, SD=1,306). Die allgemeine Berufszufriedenheit wird signifikant von der Tätigkeit ( $\beta=.882$ ,  $B=1.048$ ,  $p<.01$ ). und der Bezahlung ( $\beta=.109$ ,  $B=.098$ ,  $p<.01$ ) positiv und der Belastung ( $\beta=-.069$ ,  $B=-.086$ ,  $p<.05$ ) negativ beeinflusst. Kollegen ( $\beta=-.015$ ,  $B=-.023$ ,  $p>.05$ ) und Vorgesetzte ( $\beta=-.048$ ,  $B=-.051$ ,  $p>.05$ ) haben nur geringen Einfluss ( $R^2=.807$ ).

Diese deskriptiven Befunde bieten einen ersten Überblick in einer bislang überschaubaren Erkenntnislage und deuten - vorsichtig interpretiert - darauf hin, dass sich Hochbegabte nicht negativ von Normalbegabten unterscheiden. Um dies abzusichern, sollen die Ergebnisse im nächsten Schritt mit einer Vergleichsstichprobe Nicht-Hochbegabter verglichen werden. In der Interviewstudie sind typische Charakteristika der beruflichen Situation Hochbegabter herauszuarbeiten, insbesondere bei Teilnehmern, die erwartungswidrige Laufbahnen und Positionen einnehmen.

## **Schriftsprachliches Lernen mit literarischen Vorgaben in der Grundschule - Eine Interventionsstudie zur Förderung bildungssprachlicher Kompetenz in der 3. Klasse**

**Bettina Schartner-Ebelhäuser, Claudia Quaiser-Pohl**

Universität Koblenz-Landau, Deutschland; [schartner@uni-koblenz.de](mailto:schartner@uni-koblenz.de)

Der Beherrschung von Bildungssprache wird seit längerem eine Schlüsselfunktion für erfolgreiche Teilhabe am Bildungssystem zugeschrieben. Unterschiede in der bildungssprachlichen Kompetenz werden demnach als wichtiger Faktor für schulische Leistungsunterschiede gesehen (Baumert & Schümer, 2001; Schuth, Köhne & Weinert, 2017). Zwar scheint sich dies vor allem für Kinder mit nicht-deutscher Erstsprache nachteilig auszuwirken (Paetsch, Wolf, Stanat & Darsow, 2014), aber auch für Kinder mit Deutsch als Muttersprache stellt der Bildungsspracherwerb eine Herausforderung dar (Heppt, Stanat, Dragon, Berendes & Weinert, 2014). Inwiefern Kinder bei der Aneignung dieser Kompetenz wirkungsvoll unterstützt werden können, ist noch weitestgehend ungeklärt (Heppt, 2016). Abgeleitet aus den Befunden der internationalen Schulleistungsstudien, liegen Forschungsbemühungen vor allem im Bereich der sachfachspezifischen Förderung bildungssprachlicher Fähigkeiten (Wieler, 2011). Eine einseitige Fokussierung auf die explizite Vermittlung sprachlicher Teilkompetenzen dürfte jedoch nicht ausreichen um den Anforderungen gerade im schulischen Bereich gerecht werden. Diesem Aspekt soll in der vorliegenden Untersuchung Rechnung getragen werden, indem die Wirkung eines gezielten Trainings mit literarischen Vorgaben und impliziten Vermittlungsstrategien auf bildungssprachliche Fähigkeiten überprüft wurde. Dabei sollten sich bildungssprachliche Leistungen stärker durch die interaktive Arbeit mit Bilderbüchern verbessern lassen als durch die Kontrollgruppenbehandlung. Zudem wurden differentielle Fördereffekte für Kinder mit Deutsch als Zweitsprache erwartet. Anzunehmen war, dass diese mehr von der Arbeit mit Reim und Rhythmus profitieren als die Schülerinnen und Schüler mit Deutsch als Muttersprache.

In der vorliegenden quasi-experimentellen Untersuchung mit Kontrollgruppendesign wurde ein 7-wöchiges Programm zur Förderung schriftsprachlicher Fähigkeiten in der 3. Klassenstufe durchgeführt und evaluiert. An dem Training nahmen 249 Schülerinnen und Schüler aus insgesamt 10 rheinland-pfälzischen Grundschulen teil. Im Rahmen des Unterrichts fanden wöchentlich Vorlesegespräche nach dem Prinzip des dialogischen Lesens (Whitehurst et al., 1988) statt. Im Anschluss wurden die Kinder abwechselnd zu schriftlichen Textproduktionen oder zu mündlichen Textproduktionen in Form des Pretend Reading (Pappas & Brown, 1988) angeregt. Die zwei Interventionsgruppen erhielten zur Hälfte ein Training mit gereimten Bilderbüchern (n=84), die andere Hälfte arbeitete zu denselben Bilderbüchern in Prosaform (n=90). Die Kontrollgruppe (n=75) erhielt ein

soziales Kompetenztraining im gleichen Umfang (Gleichstein, 2005). Zur Überprüfung der Fördereffekte wurden bildungssprachlicher Wortschatz, Hörverstehensleistung und das Verständnis komplexer Satzverbindungen mit Konnektoren vor und nach dem Training standardisiert erhoben (Heppt et al., 2014; Köhne, Kronenwerth, Redder, Schuth & Weinert, 2015; Schuth, Heppt, Köhne, Weinert & Stanat, 2015). Zusätzlich erfolgte die Erfassung des Bildungshintergrundes und der Erstsprache der Kinder über Elternfragebögen, außerdem wurde das Leseverständnis der Kinder mittels eines standardisierten Tests erfasst (ELFE 1-6, Lenhard & Schneider, 2006).

Die Überprüfung der Fördereffekte erfolgte durch Kovarianzanalysen. In Bezug auf den bildungssprachlichen Wortschatz wurde der Effekt der Versuchsbedingung nach Kontrolle des Leseverständnisses signifikant ( $F(2,195) = 3.07$ ;  $p < .05$ ,  $d = 0.36$ ). Jedoch war die Nachtestleistung in der Kontrollgruppe und der Prosa-Interventionsgruppe höher als in der Reim-Interventionsgruppe. Für die Hörverstehensleistung zeigte sich, dass Kinder mit Deutsch als Erstsprache eher von der Arbeit mit gereimten Bilderbüchern profitierten, während Zweitsprachler höhere Testwerte in der Prosa- und der Kontrollgruppe als in der Reimgruppe aufwiesen ( $F(2,195) = 3.20$ ;  $p < .05$ ,  $d = 0.36$ ). Für das Verständnis komplexer Satzverbindungen durch Konnektoren konnten keine signifikanten Fördereffekte nachgewiesen werden.

Kinder scheinen von der Arbeit mit gereimten Bilderbüchern nur dann zu profitieren, wenn sie Deutsch als Erstsprache erworben haben. Ansonsten stellt dieses Angebot für den Erwerb bildungssprachlicher Strukturen möglicherweise eine zu große Hürde dar. Zudem verfügt auch das soziale Kompetenztraining über einen gewissen sprachlichen Anregungsgehalt. Eine mögliche Erklärung für den Lernzuwachs in der Kontrollgruppe könnte der sprachliche Austausch in den Feedbackrunden des Trainings sein, die ebenfalls durch Sprachhandlungen wie Analysieren und Erklären gekennzeichnet sind.

## **"Das Leid mit der verkürzten Schulzeit" - Die diskursive Produktion der Schulzeitdebatte.**

**Stefan Gebhard, Svenja Mareike Kühn**

Universität Koblenz-Landau, Deutschland; [sgebhard@uni-koblenz.de](mailto:sgebhard@uni-koblenz.de)

Theoretischer Hintergrund

Kaum ein anderes bildungspolitisches Thema wurde und wird in der Öffentlichkeit so anhaltend und intensiv diskutiert wie die Frage nach der Schulzeitdauer bis zum Abitur (vgl. zsf. Köller, 2017). Die fast flächendeckende Einführung des achtjährigen gymnasialen Bildungsgangs („G8“) hat eine breite – überwiegend negative – Resonanz in der massenmedialen Berichterstattung hervorgerufen. Nicht zuletzt auf Grund der erheblichen öffentlichen Kritik folgte kürzlich (in einigen Bundesländern) die Option zur Rückkehr zum neunjährigen Bildungsgang (vgl. Kühn, 2017), was ebenfalls Gegenstand der Berichterstattung war bzw. ist.

Der Posterbeitrag stellt ein Promotionsvorhaben zur Analyse der printmedialen Aufbereitung der Schulzeitdebatte vor. Der Agenda-Setting-Ansatz (z.B. Jäckel, 2011; McCombs, 2014), der hier als theoretischer Rahmen herangezogen wird, weist der Presse und den Printmedien generell eine entscheidende Rolle in Bezug auf den Umfang und den Gehalt der für die Öffentlichkeit verfügbaren Informationen zu gesellschaftlich-politischen Themen zu, weshalb ihnen ein wesentliches Einflusspotenzial auf die öffentliche Meinungsbildung zugesprochen wird. Es ist anzunehmen, dass die mediale Berichterstattung die öffentliche Schulzeitdebatte nicht nur widerspiegelt, sondern diese auch mitgestaltet und so Einfluss auf die öffentliche Meinung und auch politische Entscheidungen nimmt. Vor diesem Hintergrund sind die Strukturen und Entwicklungen der Medienberichterstattung zur Schulzeitfrage Gegenstand der Analysen.

Fragestellung

Die zentrale Forschungsfrage des Promotionsprojekts lautet: Wie wird die Schulzeitdebatte in verschiedenen Printmedien aufbereitet? Der üblichen Forschungspraxis folgend werden die Berichterstattungsstruktur, die Argumentationsmuster und story lines (Hajer 2009) sowie darüber hinaus länderspezifische Charakteristika regionaler Debatten und Veränderungen in der Berichterstattung im Zeitverlauf analysiert.

Methode

Das zu erstellende Datenkorpus besteht aus Artikeln verschiedener deutscher Printmedien: Sowohl ausgewählte Tageszeitungen als auch überregionale Nachrichtenmagazine werden hinsichtlich einschlägiger Artikel mithilfe einer Schlagwortsuche gesichtet, um ein möglichst breit gefächertes Textkorpus erstellen zu können, das das Medienphänomen ‚Schulzeit‘ adäquat abbildet. Als Auswahlkriterien werden Auflagenstärke, Überregionalität und ein einschlägiges Themenspektrum herangezogen. Betrachtet werden Artikel von 1999 (Beschluss der flächendeckenden Einführung des achtjährigen Bildungsgangs im ersten westdeutschen Bundesland) bis heute.

Zunächst werden überregionale Nachrichtenmagazine und Tageszeitungen fokussiert, denen eine hohe Reichweite im agenda setting-Prozess zugeschrieben wird. Es wird eine inhaltlich-strukturierende qualitative Inhaltsanalyse durchgeführt, die in Verbindung induktiver und deduktiver Verfahren sowohl zur Aussagenstrukturierung als auch -zusammenfassung genutzt werden kann (vgl. Kuckartz, 2012).

Anhand des Datenmaterials können sowohl die Forschungsfrage ausdifferenziert als auch das weitere Vorgehen präzisiert werden, bevor anschließend die regionale Berichterstattung in den Fokus rückt. Diese wird unter der Annahme betrachtet, dass regionale Berichterstattung einen großen Einfluss auf die regionalen Akteure haben kann, da hier eine besondere Nähe zu den betroffenen Akteursgruppen vorliegt. Die Auswahl der regionalen Tageszeitungen erfolgt auf der Basis der Befunde der überregionalen Analyse, in der die regionalen Schwerpunkte der Berichterstattung identifiziert werden sollen.

Um der Multimodalität der medialen Berichterstattung gerecht zu werden, umfasst das Datenkorpus nicht ausschließlich Textmaterialien, sondern ebenso zusammen mit Texten veröffentlichte Bildmedien, da diese „wesentlich an Produktion und Vermittlung gesellschaftlicher Wissensbestände teilhaben“ (Astheimer, 2016). Diese werden anhand der seriell-ikonografischen Fotoanalyse (Pilarczyk, 2005) untersucht, da diese es ermöglicht den oft implizit bleibenden politischen und soziokulturellen Informationsgehalt von Bildmedien aufzudecken (ebd., S.85).

## **Epistemologische Überzeugungen von Schülerinnen und Schülern über chemische Reaktionen - Möglichkeiten zum Umgang mit Schülervorstellungen im Chemieunterricht**

**Yannic Weber, Janina Scholz, Udo Käser**

Rheinische Friedrich-Wilhelms-Universität Bonn, Deutschland; [YannicWeber@gmx.de](mailto:YannicWeber@gmx.de), [janina.scholz92@gmx.de](mailto:janina.scholz92@gmx.de)

Epistemologische Überzeugungen von Schülerinnen und Schülern (SuS) sind für Schule und Unterricht besonders wichtig, da sie die kognitive Grundlage für die Wahrnehmung und Interpretation von Lernangeboten bilden (z. B. Barke, 2006). Dabei kommt solchen Überzeugungen hohe Persistenz zu: Erfahrungen, die zu bestehenden, inadäquaten Vorstellungen im Widerspruch stehen, führen nicht zwangsläufig dazu, dass es zu einer Anpassung der kognitiven Struktur kommt. Vielmehr bedarf es, wenn adäquate Vorstellungen nachhaltig gelernt werden sollen, der eigenaktiven Überwindung des kognitiven Konflikts (z. B. Seiler, 1980). Dies hat notwendig zur Folge, dass SuS im Unterricht Gelegenheit zum selbstständigen Lösen kognitiver Konflikte gegeben werden muss, was spezifisch konzipierte und sorgsam gestaltete Lehr-Lern-Arrangements erforderlich macht.

Für die Vorstellungen, die SuS mit chemischen Reaktionen verbinden, sind verschiedene Präkonzepte aus der fachdidaktischen und lernpsychologischen Forschung bekannt. Prototypisch sind neben der adäquaten Vorstellung von Massenerhaltung noch das Vernichtungs-, Zerlegungs- und Eigenschaftskonzept (Barke, 2006). Für den Chemieunterricht stellt sich Lehrkräften damit die Aufgabe, diese epistemologischen Überzeugungen (bzw. etwaige Mischformen oder diffuse Vorstellungen) ihrer SuS zu erkennen und durch ihren Unterricht ein fundiertes Verständnis von Massenerhaltung zu vermitteln.

Vor diesem Hintergrund wird in der vorliegenden Studie untersucht, ob eine spezifisch konzipierte Unterrichtseinheit, „Lavoisiers Problem mit der Masse“, welche die nachhaltige Vermittlung adäquater Vorstellungen über chemische Prozesse zum Ziel hat, tatsächlich zu einem langfristigen Lernerfolg führt.

Die Unterrichtseinheit wurde auf Grundlage des Denkstufenmodells nach Schmidkunz und Lindemann (1999) und des Lehr-Lern-Modells nach Leisen (2010) entwickelt. Sie umfasst acht Unterrichtsstunden à 45 Minuten. Die erste Stunde diente zur Einführung, die letzte Stunde der Reflektion und Ergebnissicherung. In Unterrichtsstunde 2 bis 7 bearbeiten die SuS bis zu sechs Stationen (vier Pflichtstationen), die jeweils Angebote zum „guided inquiry learning“ (vgl. Johnstone & Al-Shuali, 2001, Wichmann & Leutner, 2009) enthalten. Dabei wurden bei inhaltlichen Schwierigkeiten Tippkarten eingesetzt und für Kinder mit Migrationshintergrund sprachliche Hilfen angeboten. Bei der Gestaltung der Stationen wurde darauf geachtet, den SuS

- das Anknüpfen an alltägliche Vorerfahrungen durch realitätsnahe Sachverhalte,
- selbstständiges, experimentelles Arbeiten,
- das Erfahren der Grenzen eigener Überzeugungen durch erwartungswidrige Ergebnisse sowie
- die Reflexion der Lernfortschritte durch den Vergleich vorheriger Vorstellungen mit neu erworbenem Wissen zu ermöglichen.

Die Wirksamkeit der Unterrichtseinheit wurde an einer Stichprobe von 105 SuS der Mittelstufe einer Gesamtschule quasiexperimentell im Vergleich von einer Experimentalgruppe (n = 75) mit einer Kontrollgruppe (n = 30) über drei Messzeitpunkte (prä, post, follow-up) untersucht. Dabei wurden die Schülervorstellungen durch einen multiple-choice-Test mit acht Items erfasst. Jedes Item schildert eine chemische Problemsituation, die nicht in den Stationen behandelt worden war und für die Antworten im Sinne der Präkonzepte des Massenerhaltungskonzepts, Vernichtungs-, Zerlegungs- und Eigenschaftskonzept in Concept Cartoons (vgl. Steininger & Lembens, 2011) formuliert worden waren. Auf das Vorliegen eines Präkonzepts wurde bei den Probanden gemäß einer probabilistischen Modellierung unter Beachtung der Ratewahrscheinlichkeit geschlossen, wobei auch Mischkonzepte und diffuse Vorstellungen berücksichtigt wurden. Des Weiteren wurde die Häufigkeit, mit der das Massenerhaltungskonzept im Test angekreuzt wurde, ausgewertet. Als mögliche Kovariaten wurden neben Geschlecht, Alter und der letzten Note im Fach Chemie noch Motivation (nach Sumfleth & Wild, 2001), Interesse (nach Klos, 2008, Fechner, 2009) und Selbstwirksamkeit (nach Schwarzer & Jerusalem, 1995) erfasst.

Während die Präkonzepte innerhalb der Kontrollgruppe stabil waren, lagen in der Experimentalgruppe spürbare Verbesserungen vor: 44% der Jugendlichen zeigten eine kurzfristige Verbesserung hinsichtlich ihrer epistemologischen Überzeugungen, für 29% war die Verbesserung zum follow-up hin stabil ( $p < .001$ ). Für die Häufigkeit korrekter Antworten im Test zeigte sich ebenfalls ein signifikanter, im Längsschnitt stabiler Unterschied zugunsten der Experimentalgruppe ( $p < .001$ ). Während das Geschlecht keinen Einfluss auf die Verbesserung hatte, profitierten vor allem Schülerinnen und Schüler mit hoher Motivation ( $p < .05$ ), hoher Selbstwirksamkeit ( $p < .05$ ) und hohem Vorwissen ( $p < .01$ ) von dem Unterricht.

### **VISIT-Math – Validierung einer Simulation zur Erfassung von Diagnosekompetenzen beim mathematischen Argumentieren von Schülerinnen und Schülern**

**Elias Codreanu<sup>1</sup>, Sarah Reinhold<sup>1</sup>, Sina Huber<sup>1</sup>, Daniel Sommerhoff<sup>2</sup>, Stefan Ufer<sup>2</sup>, Tina Seidel<sup>1</sup>**

<sup>1</sup>Technische Universität München, Deutschland; <sup>2</sup>Ludwig-Maximilians-Universität München, Deutschland; [elias.codreanu@tum.de](mailto:elias.codreanu@tum.de)

#### Theoretischer Hintergrund

Diagnosekompetenzen werden als zentrales Element des professionellen Handelns von Lehrkräften gesehen (Baumert & Kunter, 2006; Südkamp & Praetorius, 2017). Sie sind beispielsweise entscheidend, um Unterricht bestmöglich auf die Lernvoraussetzungen der Klasse als Ganzes auszurichten (Beck et al., 2008; Rogalla & Vogt, 2008), sowie um binnendifferenzierende Elemente und individuelle Fördermaßnahmen zu konzipieren (Artelt & Gräsel, 2009). Hierfür ist es insbesondere erforderlich, dass Lehrkräfte den Leistungsstand der Lernenden sowie die Anforderungen von Lernsituationen einschätzen können (Schrader, 2013). Im Unterrichtsfach Mathematik betreffen solche diagnostischen Entscheidungen häufig prozessbezogene Kompetenzen der Schülerinnen und Schüler (Blum, Roppelt, & Müller, 2013), in der Sekundarstufe insbesondere die mathematische Argumentationskompetenz (KMK, 2003).

Typischerweise werden Lerninhalte in Mathematik im Unterrichtsgespräch zwischen Lehrperson und Lernenden vermittelt (Seidel, 2014), wobei die hohe Interaktionsdichte und die geringe Möglichkeit zur Handlungsunterbrechung herausfordernd für die Lehrkraft ist (Doyle, 1986). Zur gezielten Förderung von Lehrkräften, beispielsweise im Bereich Diagnosekompetenz, werden in der Bildungsforschung sowie auch in der Lehrerbildung Unterrichtsvideos eingesetzt, da diese einerseits eine Unterrichtssituation authentisch wiedergeben, andererseits die Komplexität einer realen Unterrichtssituation reduzieren (Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2011). Auch im medizinischen Bereich, in dem das Diagnostizieren ebenfalls einen wichtigen Bestandteil professionellen Handelns darstellt, werden Simulationen in der universitären Ausbildung genutzt (Peeraer et al., 2007). Ausgehend von dieser Herangehensweise wurde in der vorliegenden Studie VISIT-Math im Rahmen der DFG-Forschergruppe COSIMA eine videobasierte Simulationsumgebung mit Fokus auf das Diagnostizieren von mathematischen Argumentationskompetenzen von Schülern im Bereich Geometrie entwickelt. Als ein Kriterium für die Qualität der Simulation wird in einem ersten Validierungsschritt das Gefühl der Lernenden in die Lernumgebung einzutauchen (Presence) geprüft (Frank, 2015; Scoresby & Shelton, 2011). Es werden die nach Schubert, Friedmann, und Regenbrecht (2001) definierten



Teilaspekte wahrgenommene räumliche Präsenz, Involviertheit und Authentizität erhoben. In einem zweiten Schritt werden in der Hauptstudie zusätzlich Unterschiede von Experten und Novizen in der Diagnosequalität geprüft.

Fragestellungen

Zur Validierung der videobasierten Simulationsumgebung werden folgende Fragestellungen untersucht:

(1) Ermöglicht die Simulation mit den entwickelten Videosequenzen zur Lehrer-Schüler-Interaktion den Probanden ein Empfinden von Presence?

(2) Inwiefern zeigen Experten eine höhere Akkuratheit in der Diagnose der Argumentationskompetenz der in der Simulationsumgebung beobachtbaren Lernenden?

Methode

Die Stichprobe der Pilotierung umfasst N = 5 Experten (erfahrene Mathematiklehrkräfte und Fachkräfte der Mathematikdidaktik) sowie N = 5 Novizen (Lehramtsstudentinnen und -studenten). Die angestrebte Stichprobe der Hauptstudie umfasst N = 30 Teilnehmer, zu gleichen Teilen aus Experten und Novizen. Die Simulationsumgebung umfasst inszenierte Videos, die eine Unterrichtssituation aus dem Mathematikunterricht im Bereich der Geometrie in der achten Klasse mit vier Schülerinnen und Schülern sowie einer Lehrkraft zeigen. Die Schülerinnen und Schüler repräsentieren verschiedene Stufen der mathematischen Argumentationskompetenz, die nach dem Abspielen der Videosequenz durch die Teilnehmer diagnostiziert werden sollen. Zur Generierung zusätzlicher Evidenz können die Teilnehmerinnen und Teilnehmer sich dafür entscheiden, weitere Videosequenzen aus der gleichen Unterrichtssituation abzuspielen. Im direkten Anschluss der Fallbearbeitung werden die Presenceskalen mit insgesamt zwölf Items erhoben. Die Akkuratheit der Diagnose wird unter anderem mittels einer Fragebatterie, welche einzelne Aspekte der Argumentationskompetenz der Schüler beleuchtet, geprüft.

Ergebnisse und Ausblick

Die Daten werden bis Ende des Jahres erhoben und liegen bis zur GEBF 2018 vor. Bezüglich Fragestellung 1 wird erwartet, dass die Probanden ein ausgeprägtes Gefühl von Presence empfinden. Des Weiteren erwarten wir einen Unterschied der Experten und Novizen hinsichtlich ihrer Urteilsgenauigkeit bei den vier Schülerinnen und Schülern (Fragestellung 2). Angenommen wird eine höhere Diagnoseakkuratheit der Experten im Vergleich zu den Novizen. Die Ergebnisse werden mit Blick auf weiterführende Analysen der gewonnenen Daten diskutiert.

## **„Paul, lies doch mal!“ – Der Einfluss der Selbstbestimmung auf die Lesemotivation im Grundschulalter**

**Madeleine Stein, Martina Endepohls-Ulpe**

Universität Koblenz-Landau, Deutschland; [madeleinstein@uni-koblenz.de](mailto:madeleinstein@uni-koblenz.de)

Als notwendige Bedingung für den Wissenserwerb und für die Teilhabe am kulturellen Leben insgesamt kommt der Lesekompetenz eine Schlüsselfunktion für alle Bereiche des Lebens zu (Naumann, Artelt, Schneider & Stanat, 2010). Allerdings belegen Schulleistungsvergleichsstudien seit vielen Jahren, dass Jungen im Vergleich zu ihren gleichaltrigen Mitschülerinnen bereits im Grundschulalter geringere Leseleistungen zeigen. Die signifikanten Leistungsunterschiede in der Lesekompetenz, die im Laufe der Entwicklung noch weiter zunehmen, werden häufig auf eine höhere intrinsische Lesemotivation der Mädchen zurückgeführt (Bos et al., 2007). Die Gründe dafür sind jedoch noch nicht abschließend erforscht. Laut dem motivationstheoretischen Konzept von Ryan und Deci (2000) wird intrinsische Motivation durch die Faktoren selbsterlebte Kompetenz („competence“), soziale Eingebundenheit („relatedness“) und Selbstbestimmung („autonomy“) bedingt. Vor allem die autonome Handlungsregulation geht dabei häufig mit einem Interesse am Lern- bzw. Lesegegenstand einher und beeinflusst die Qualität des Lernens/Lesens sowie folglich die Leistungsergebnisse (Deci&Ryan, 1993).

Die vorliegende Studie untersucht, welchen Einfluss die Selbstbestimmung auf den Spaß am Lesen und die Leseleistung hat. Die Fragestellungen lauten a) Haben SchülerInnen, die selbstbestimmter beim Lesen sind auch mehr Spaß beim Lesen? b) Zeigen selbstbestimmtere SchülerInnen höhere Leistungen im Leseverständnis? c) Zeigen sich Geschlechter- und Altersunterschiede in Bezug auf Selbstbestimmung, Spaß am Lesen und Leseleistung?

Hierzu bearbeiteten Dritt- (n=248) und ViertklässlerInnen (n=210) aus zwölf rheinland-pfälzischen Grundschulen im Klassenverband den Leseverständnistest ELFE 1-6 (Lenhard & Schneider 2006), welcher Wort-, Satz und Textverständnis abbildet. Zusätzlich wurden das Leseverhalten (Beispielitem: „Wie oft liest du in deiner Freizeit?“), die Lesemotivation (Beispielitem: „Ich lese, weil ich es spannend finde.“) und die Selbstbestimmung der Kinder in Bezug auf das Lesen (Beispielitem: „Darfst du in deiner Freizeit bestimmen, was du liest?“) erfasst.

Die Befunde aus Varianzanalysen mit Geschlecht und Klassenstufe als festen Faktoren und den Skalenwerten „Selbstbestimmung“ und „Spaß am Lesen“ als abhängigen Variablen zeigen, dass es signifikante Geschlechts- und Altersunterschiede in Bezug auf die Selbstbestimmung und den Spaß am Lesen gibt. Obwohl die Selbstbestimmung von Klassenstufe 3 zu Klassenstufe 4 bei allen SchülerInnen zunimmt, zeigt sich in beiden Klassenstufen ein höherer Wert bei den Mädchen. Der Spaß beim Lesen nimmt vor allem bei den Mädchen mit steigendem Alter zu, was in Klassenstufe 4 durch einen signifikanten Geschlechtsunterschied deutlich wird. Auch in Bezug auf die Leistung im Leseverständnis zeichnen sich mit zunehmenden Alter Geschlechtsunterschiede zu Gunsten der Mädchen ab. Die Mediationsanalyse ergab zudem, dass Spaß am Lesen den Zusammenhang zwischen Selbstbestimmung und der Leistung im Leseverständnis partiell mediert.

Der Selbstbestimmung sollte somit aufgrund ihres Einflusses auf den Spaß am Lesen und damit auf das intrinsisch motivierte Lesen mehr Beachtung zukommen. Im Kontext Schule könnte dem bereits zu Beginn der Primarstufe Rechnung getragen werden – so zum Beispiel durch Zeit für eigene Lektüren im Unterricht oder auch durch stärkere Mitbestimmung der Kinder bei für den Unterricht ausgewählten Lektüren. Im Freizeitbereich könnte dies durch mehr elterliche Freiheit für die Kinder in der Lektürewahl und bei den Rahmenbedingungen des Leseverhaltens unterstützt werden.

## **Entwicklung und Evaluation eines Tests zur Messung von Zeitmanagement bei Studierenden**

**Daniel O. Thies<sup>1</sup>, Detlev Leutner<sup>2</sup>, Joachim Wirth<sup>1</sup>**

<sup>1</sup>RUHR-Universität Bochum, Deutschland; <sup>2</sup>Universität Duisburg-Essen; [daniel.thies@rub.de](mailto:daniel.thies@rub.de)

Für Studierende stellt im Kontext des selbstregulierten Lernens das Zeitmanagement eine besondere Herausforderung dar, was auch mit akademischer Leistung in Verbindung steht (z. B: Zimmerman, 2013). Überblicksarbeiten postulieren, dass gutes Zeitmanagement mit verschiedenen Vorteilen für Studierende einhergeht, schlechtes Zeitmanagement hingegen mit Nachteilen

wie Anspannung, Stress sowie geringerem Well-Being (Claessens, van Eerde, Rutte & Roe, 2007; Steel, 2007). Entsprechend stellen Fähigkeiten im Zeitmanagement eine grundlegende Bedingung für Studienerfolg dar.

Befunde zum Einfluss von Zeitmanagement im akademischen Kontext beruhen größtenteils auf Selbstberichtsverfahren (z. B. Häfner, Stock & Oberst, 2015). Die Validität dieser Selbstberichtsverfahren ist für die Erfassung von Leistungen jedoch fraglich (Spörer & Brunstein, 2006; Wirth & Leutner, 2008). In dieser Arbeit wird daher ein erster Schritt zur Entwicklung eines Leistungstests für das Zeitmanagement von Studierenden unternommen. Dafür nehmen wir drei Teilfähigkeiten in den Blick: (1) die Fähigkeit einen guten Zeitplan zu erkennen, (2) einen guten Zeitplan aufzustellen und (3) sich an einen aufgestellten Zeitplan zu halten.

Im ersten Schritt entwickeln und evaluieren wir einen Test für die erste Teilfähigkeit. Dafür konstruierten wir zunächst Vignetten für fiktive, aber realistische Studiensituationen (Schreiben einer Hausarbeit, Vorbereitung auf eine mündliche oder schriftliche Prüfung). In einem zweiten Schritt wurden für jede Vignette, unterschiedlich gute Zeitpläne erstellt, indem wir systematisch relevante Zeitmanagementvariablen wie Zielsetzung, Zielformulierung, Teilzielgenerierung, Ableitung verschiedener Aufgaben und Aufgabenhierarchien, Erstellung von „to-do“ Listen sowie Überlegungen zu deren Umsetzung in Zeitpläne variierten (Macan & Sahani, 1990; Britton & Tesser, 1991). In einem forced-choice-Format werden die verschiedenen Zeitpläne für eine Vignette in allen möglichen Kombinationen nacheinander paarweise dargeboten, und die Versuchspersonen haben jeweils die Aufgabe den Zeitplan mit höherer Qualität zu identifizieren. Der Zeitmanagementtest wird online implementiert.

Vorgestellt werden die Ergebnisse aus zwei Validierungsstudien. In der ersten Studie wurden die Items Experten zur Beurteilung der Inhaltsvalidität vorgelegt. In der zweiten Studie mit Erstsemesterstudierenden werden neben der Fähigkeit einen guten Zeitplan zu erkennen weitere Zeitmanagementvariablen, die akademische Leistung, das Stressempfinden sowie Persönlichkeitsfaktoren (z.B. Gewissenhaftigkeit) erhoben.

Neben den Ergebnissen zum ersten Test, wird das Poster einen Ausblick auf die weiteren zu entwickelnden und zu evaluierenden Leistungstests geben, und es wird diskutiert, welche Möglichkeiten sich aufgrund der Testergebnisse für eine professionelle Unterstützung von Studierenden ergeben.

## **Entwicklung techn didaktischer Kompetenz im Verlauf schulpraktischer Studien**

**Friederike Straub<sup>1</sup>, Bernd Geißel<sup>1</sup>, Rehm Markus<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Ludwigsburg, Deutschland; <sup>2</sup>Pädagogische Hochschule Heidelberg, Deutschland;  
[friederike.straub@ph-ludwigsburg.de](mailto:friederike.straub@ph-ludwigsburg.de)

### Theoretischer Hintergrund

Die Bologna-Reform führte zu zahlreichen Veränderungen im Bildungswesen, die auch Lehrpersonen maßgeblich betreffen (vgl. u.a. Bauer & Prenzel, 2012; Winter, 2011). Hierbei trägt unter anderem die Kompetenzforschung einen erheblichen Anteil. Insbesondere das Modell professioneller Handlungskompetenz von Baumert und Kunter (2013) steht aktuell im Zentrum zahlreicher empirischer Arbeiten der Bildungsforschung. Die Autoren deklarieren das professionelle Wissen von Lehrpersonen als einen grundlegenden Bestandteil für die Profession des Lehrberufs (vgl. ebd.). Das Modell orientiert sich an den Arbeiten Shulmans (1987), der verschiedene Facetten des Professionswissens postuliert. Dabei haben drei Facetten eine übergeordnete Bedeutung gewonnen: content knowledge (CK), general pedagogical knowledge (PK) sowie pedagogical content knowledge (PCK) (vgl. Shulman, 1986, 1987). Neben dem Professionswissen gilt die professionelle Unterrichtswahrnehmung als wesentlicher Bestandteil der Lehrerexpertise (vgl. Jahn, Stürmer, Seidel & Prenzel, 2014). Das Konstrukt impliziert die Art und Weise, wie Lehrpersonen Unterrichtsereignisse wahrnehmen und interpretieren, um daraufhin angemessen handeln zu können (vgl. u.a. Seidel, Blomberg & Stürmer, 2010; Stürmer & Seidel, 2015). Daher dient das Konstrukt als Indikator für die Anwendung des Professionswissens und folglich als Indiz für die Kompetenzen von Lehrpersonen.

Im Zuge der Bologna-Reform kam weiterhin die Forderung nach der Erhöhung des Anteils praxisbezogener Ausbildungsanteile auf (vgl. u.a. Arnold et al., 2011; Bach, 2013; Schubarth, Speck & Seidel, 2011; Schubarth et al., 2012; Winter, 2011). Der Anspruch einer stärkeren berufspraktischen Orientierung und Aufwertung von Praktika resultierte im Ausbau des zeitlichen Umfangs sowie in neuen Ansätzen der investierten Lernbegleitung seitens der Hochschulen (vgl. Gröschner & Seidel, 2012). Bisherige Arbeiten zeigen, dass die Ausgestaltung der Praxisphasen im Lehramtsstudium zwischen, sowie innerhalb der Bundesländer erheblich variiert und diesbezüglich ein „bunter Flickenteppich“ (Gröschner et al., 2015, S. 642) vorherrscht. Darüber hinaus fehlt es bislang an empirischer Evidenz ob Praxisphasen eine (erhöhte) Förderwirkung bezüglich der Kompetenzentwicklung angehender Lehrpersonen hervorrufen.

Erste empirische Arbeiten zu Lehrkompetenzen im Fachbereich Technik stammen von Goreth (2017), der mit dem entwickelten vignettengestützten Testinstrument PCK-T (Pedagogical Content Knowledge – Technical Education) fachdidaktische Kompetenzen angehender Techniklehrkräfte erfasst. Es besteht empirische Evidenz, dass das entwickelte Testinstrument in der Lage ist, zwischen verschiedenen Kompetenzniveaus respektive Entwicklungsständen der Studierenden zu differenzieren (vgl. Goreth, Rehm & Geißel, 2016; Goreth, 2017). Im Rahmen des Forschungsvorhabens, soll nun geprüft werden, ob das Testinstrument sensitiv ist, auch Entwicklungen im Verlauf schulpraktischer Studien abzubilden.

### Fragestellung

Unter der Prämisse, dass Lehramtsstudierenden in zentralen Studienphasen neue Kompetenzen erwerben bzw. bestehende erweitern, ist zu prüfen, ob das Vignettentestinstrument PCK-T von Goreth (2017) sensitiv ist, diese unterstellte Entwicklung fachdidaktischer Kompetenzen abzubilden. Die vorliegende Arbeit fokussiert die Entwicklung fachdidaktischer Kompetenzen im Verlauf schulpraktischer Studien. Aufgrund der Diversität, was die Gestaltung von Schulpraktika betrifft, beschränkt sich die vorliegende Arbeit auf das sogenannte integrierte Semesterpraktikum (ISP) an den Pädagogischen Hochschulen in Baden-Württemberg, um eine Vergleichbarkeit der Praktikastrukturen gewährleisten zu können.

### Methode

Der eben skizzierten Forschungsfrage soll mittels Längsschnittstudie im Mehrkohorten-Panel-Design nachgegangen werden. Im Zuge einer Pre-Post-Modellierung werden die Kompetenzstände der Studierenden zu Beginn und zum Ende des integrierten Semesterpraktikums erfasst. Als Kontrollgruppe dienen Lehramtsstudierenden des Faches Technik im höheren Fachsemester, welche das integrierte Semesterpraktikum jedoch noch nicht absolviert haben.

### Ergebnisse

Es sollen erste Ergebnisse der aktuell laufenden Längsschnittstudie im Verlauf des integrierten Semesterpraktikums (ISP) präsentiert und diskutiert werden. Es liegen bereits Daten von N = 83 Probanden vor, die zu beiden Messzeitpunkten an der Erhebung teilnehmen konnten. Der Gesamtdropout liegt bei 28,5%. Der Anteil weiblicher Studierender beläuft sich auf 15,7%, der Anteil männlicher Studierender auf 84,3%. Die Probanden befanden sich im Mittel im M = 5,42 (SD = 1,49) Fachsemester. Weiterhin soll die Stichprobengröße durch Befragung weiterer Kohorten erhöht werden, um die Aussagekraft der Daten zu steigern.

## **Wissenskohärenz durch Lerntagebücher und Lösungsbeispiele bei angehenden Geographielehrkräften**

**Martina Graichen, Matthias Nückles**

Universität Freiburg, Deutschland; [martina.graichen@ezw.uni-freiburg.de](mailto:martina.graichen@ezw.uni-freiburg.de)

Angehende Lehrkräfte erwerben während ihres Universitätsstudiums Wissen zu Fachwissenschaft, Fachdidaktik und Allgemeiner Pädagogik/Psychologie. Jedoch werden diese Wissensbereiche selten miteinander vernetzt oder gemeinsam gelehrt. Deshalb muss die Vernetzung der drei Wissensbereiche von den angehenden Lehrkräften selbstständig erfolgen. Allerdings ist die Integration unterschiedlicher Wissensbereiche kognitiv anspruchsvoll (Ball, 2000) und das nicht-Vernetzen kann zu fragmentiertem und tragem Wissen führen, welches nicht anwendbar ist (Ball, 2000; Harr, Eichler, & Renkl, 2015).

In der Vergangenheit wurde empirisch gezeigt, dass das angeleitete Schreiben eines Lerntagebuchs durch die Anregung kognitiver und metakognitiver Lernstrategien vertieftes Lernen anregt (Nückles, Hübner, & Renkl, 2012). Angeleitet wird das Lerntagebuchs Schreiben durch Prompts, also Hinweisen oder Leitfragen, welche die Organisation, Elaboration und Metakognition der Lerninhalte fördern. Wäschle, Lehmann, Brauch und Nückles (2015) sowie Graichen (2016) zeigten, dass intertextuelle Prompts beim Lerntagebuchs Schreiben, also Prompts, die explizit auf die verknüpfende Organisation und Elaboration von verschiedenen Textinhalten hinweisen, die Wissensintegration fördern, jedoch auf niedrigem Niveau. Ein Ansatz, wie die Wissensintegration während des Lehramtsstudiums besser angeleitet und intensiviert werden kann, könnten beispielhafte Lerntagebucheinträge sein, die konkret die Verknüpfung der Wissensbereiche untereinander zeigen. Lösungsbeispiele sind eine effektive Lernform für Lernende besonders zu Anfang des Fertigkeitserwerbs. Ein Lerntagebucheintrag kann als Lösungsbeispiel betrachtet werden, insofern beispielhaft essenzielle Lernstrategien zur Organisation und Integration von Informationen darin gezeigt werden können (Hübner, Nückles, & Renkl, 2010).

Fachlich wird dieser Ansatz im Unterrichtsfach Geographie verankert. Entsprechend geht es in den fachwissenschaftlichen und fachdidaktischen Texten um Vulkanismus, im allgemein-pädagogischen Text geht es um Visualisierungen, die im Zusammenhang mit Vulkanismus sinnvoll genutzt werden können.

Wir gehen davon aus, dass eine Kombination aus intertextuellen Prompts und einem Beispielerntagebucheintrag (Lösungsbeispiel) anschließend beim eigenständigen Schreiben zu mehr Organisations- und Elaborationsstrategien führen, die auf eine Verknüpfung der Wissensbereiche abzielen, als die Anleitung nur durch intertextuelle Prompts ohne Lösungsbeispiel. Des Weiteren sollte die intertextuelle Gruppe mit Lösungsbeispiel auch in der Anwendung der gelernten Inhalte mehr Textverbindungen wiederaufnehmen als eine Gruppe nur mit Prompts, und eine Kontrollgruppe ohne Prompts und ohne Lösungsbeispiel.

### **Methode**

An der Laborstudie nehmen zu Beginn des Wintersemesters 2017/2018 circa fünfundsiebzig Gymnasiallehramtsstudierende des Faches Geographie teil. Die Teilnehmer werden auf drei Gruppen randomisiert: Eine Kontrollgruppe, die weder Prompts noch ein Lösungsbeispiel erhält, eine Experimentalgruppe, die intertextuelle Prompts erhält und eine zweite Experimentalgruppe, die intertextuelle Prompts, und in der Übungsphase einen Lerntagebucheintrag als Lösungsbeispiel erhält.

In der Übungsphase lesen zunächst alle Teilnehmer pro Wissensbereich einen Text (Gesteinsbildung, Geologie im Unterricht, Lerntexte für Schüler verfassen). Im Anschluss schreiben die beiden Gruppen ohne Beispiel einen eigenen Lerntagebucheintrag, während die Gruppe mit Lösungsbeispiel dieses sorgfältig studiert und selbsterklärt. Eine Woche später, in der Anwendungsphase, lesen die Teilnehmer erneut pro Wissensbereich einen Text (naturwissenschaftliche Mechanismen bei der Vulkanentstehung, Prä-Konzepte bezogen auf den Vulkanismus, Visualisierungen im Unterricht). Im Anschluss an die Lese-Phase (30 Minuten) schreiben alle Teilnehmer einen Lerntagebucheintrag (40 Minuten). Um die Anwendung des Wissens zu überprüfen, bewerten die Teilnehmer am Ende mehrere Visualisierungen zum Thema Vulkanismus hinsichtlich ihrer didaktischen Qualität in Bezug auf Schülerinnen und Schüler einer neunten Klasse (Anwendungsaufgaben).

### **Erwartete Ergebnisse und Diskussion**

Wir gehen davon aus, dass die Gruppe mit intertextuellen Hinweisen und Lösungsbeispiel während der Übungsphase später mehr Verknüpfungen zwischen den drei Wissensbereichen im eigenen Lerntagebucheintrag vornimmt und diese Verbindungen öfters in der Anwendungsaufgabe zur Evaluation heranzieht als die beiden anderen Gruppen.

Wenn sich unsere Hypothesen durch diese Interventionsstudie bestätigen, dann zeigt sich, dass sich das kognitiv anspruchsvolle Verknüpfen von Textinhalten durch eine umsichtig gestaltete instruktionale Lernumgebung fördern lässt. Somit wäre es erstrebenswert, diese Intervention längerfristig durch eine Implementation der Lerntagebücher mit Lösungsbeispiel und intertextuellen Prompts im neugestalteten Master of Education an der Universität Freiburg auf seine Anwendbarkeit und Ergiebigkeit in der Praxis zu evaluieren.

## **Nachhaltiges Lernen: Was wissen Neuntklässler im Fach Chemie?**

**Anja Kramer, Eva Pischel, Udo Käser**

Rheinische Friedrich-Wilhelms-Universität Bonn, Deutschland; [anja8892@gmx.de](mailto:anja8892@gmx.de)

Wenn Schülerinnen und Schüler (SuS) lernen sollen, sich als Individuen mit ihrer inneren und ihrer äußeren Welt reflektiert auseinander zu setzen, müssen schulische Lernprozesse den nachhaltigen Erwerb von Wissen und Fähigkeiten zum Ziel haben (vgl. Schüßler, 2006). Für die Schulorganisation bedeutet dies mit Blick auf den Erwerb der Allgemeinen Hochschulreife in der Oberstufe, dass Kenntnisse, welche in der Sekundarstufe I vermittelt werden, so weit gefestigt sein sollten, dass im Unterricht der Sekundarstufe II SuS auf sie automatisiert zurückgreifen und Lehrpersonen hierauf ohne Schwierigkeiten Bezug nehmen können.

Durch zentrale Prüfungen, die zum Ende der Sekundarstufe I durchgeführt werden, erhalten – bei aller möglicher Kritik (vgl. Hagemeyer, 2007) – SuS, Eltern und Lehrkräfte zumindest für Hauptfächer eine gewisse Rückmeldung darüber, inwieweit

Grundlagen der Sekundarstufe I vorliegen. Für andere Fächer fehlen diagnostische Instrumentarien, die die Nachhaltigkeit schulischen Lernens überprüfen, oder werden nur in Einzelfällen eingesetzt. Es bleibt dann die alleinige pädagogische Aufgabe der Lehrkraft, die zu Beginn der Oberstufe unterrichtet, das Beherrschen der notwendigen Grundlagen zu überprüfen und sicherzustellen.

Für das Fach Chemie liegen nur wenige Befunde zur Nachhaltigkeit des Wissenserwerbs in der Sekundarstufe I vor. Beispielsweise wurde in PISA 2006 / 2015 zwar ein Schwerpunkt auf naturwissenschaftliche Grundbildung gelegt, ohne dass aber ein curricularer Bezug vorliegt. Im Kontext der NAW-Studien untersucht Gehlen (2016) Kompetenzen von SuS für das Fach Chemie, nimmt aber im Wesentlichen Schülervorstellungen über den Prozess der Erkenntnisgewinnung in den Blick. Die Entwicklung des Wissens im Fach Chemie von der siebten zur neunten Klasse wird von Ferber (2014) untersucht. Hier zeigt sich eine hohe Schwierigkeit bei Aufgaben auf Teilchenniveau, welche auf eine geringe Nachhaltigkeit bei der Vermittlung komplexer Lerninhalte in der Sekundarstufe I hinweist.

Vor diesem Hintergrund wurde eine Studie an einer Stichprobe von 260 Neuntklässlern aus elf Klassen (94 Jungen, 166 Mädchen) realisiert, um folgende Fragen zu untersuchen:

F1: Wie viel wissen die SuS am Ende der Sekundarstufe I über Chemie?

F2: Welche Inhalte des Fachs erweisen sich als besonders schwierig oder einfach?

F3: Was sind Prädiktoren, die das Ausmaß des Wissens über Chemie bedingen?

Hierfür wurde in Orientierung am Curriculum des Landes NRW ein diagnostischer Wissenstest konzipiert. Der Test besteht aus 21 Items und deckt die thematischen Schwerpunktthemen des Chemieunterrichts der Sekundarstufe I valide ab (chemische Reaktionen, Struktur der Materie, Energie). Dabei wurden Single-Choice-Aufgaben, Zuordnungsaufgaben, Lückentexte sowie halboffene und offene Aufgabestellungen eingesetzt und die Schülerleistungen auf einer Skala von 0 bis 100 Punkten gemessen. Daneben wurden Fachinteresse (10 Items) und Selbsteinschätzung der Fachkompetenz (5 Items) erfasst. Für die drei Skalen lagen ausreichend hohe Reliabilitäten (Cronbachs  $\alpha > .75$ ) vor. Außerdem wurden personenbezogene Daten (Geschlecht, Alter, Chemienote, Kurswahl für die Oberstufe) erfasst.

Die Leistungen der SuS fielen im Mittel schwach aus und variierten sowohl im Vergleich der Klassen als auch in der Stichprobe in hohem Maße ( $M = 19.2$ ,  $SD = 10.6$ ,  $Min = 0.0$ ,  $Max = 57.5$ ). Jungen schnitten signifikant besser als Mädchen ab ( $p < .001$ ). Auch gab es signifikante Unterschiede zugunsten von SuS, die das Fach Chemie für die Oberstufe weiter gewählt hatten ( $p < .001$ ). Die Selbsteinschätzung des eigenen Wissens und die Bewertung durch die Lehrkraft korrelierten positiv mit der Leistung im Wissenstest (je  $p < .001$ ). Die Itemschwierigkeit hing stärker von den Aufgabeninhalten als von den Aufgabenformaten ab. Wie bei Ferber (2014) bestätigte sich die hohe Schwierigkeit von Aufgaben auf Teilchenniveau.

Die Befunde regen eine Diskussion über die Nachhaltigkeit des Wissenserwerbs im Chemieunterricht an. So zeigt sich, dass ein Großteil der Fachkompetenzen, die am Ende der Sekundarstufe I vorliegen sollten, nur unzureichend vorhanden sind und ein nachhaltiges Lernen nicht stattgefunden hat. Dies wirft die Frage auf, wie ein moderner Chemieunterricht hiermit umgehen will.

## **Adaptive Unterstützung beim Experimentieren in leistungsheterogenen Lerngruppen im Fach Chemie**

**Dennis Kirstein, Sebastian Habig, Maik Walpuski**

Universität Duisburg-Essen, Deutschland; [dennis.kirstein@uni-due.de](mailto:dennis.kirstein@uni-due.de)

Die immer vielfältigeren Lernvoraussetzungen von Schülerinnen und Schülern machen das Lehren und Lernen zu einem komplexen Tätigkeitsbereich im Handlungsfeld Schule. Nicht zuletzt erfährt ein produktiver Umgang mit Heterogenität aktuell auch im Kontext der inklusiven Schulentwicklung großen Zuspruch, sodass hier einmal mehr die Notwendigkeit adaptiver Lerngelegenheiten für individualisiertes Lernen im Fachunterricht hervorgehoben werden muss.

Jedoch steht der Fachunterricht in einem Spannungsfeld zwischen verbindlichen Bildungszielen (vgl. KMK, 2004, AAAS, 1990) und den Ansprüchen an die Berücksichtigung unterschiedlicher Voraussetzungen und Bedingungen seitens der Schülerinnen und Schüler (vgl. PISA 2015; IQB Ländervergleich 2012). Für den Chemieunterricht liegt ein bedeutungsvoller Zugang zur Auseinandersetzung mit diesem Spannungsfeld in der leistungsbezogenen Dimension von Heterogenität (vgl. Trautmann & Wischer, 2011), da zahlreiche konzept- und prozessbezogene Kompetenzen erworben werden sollen. Von zentraler Bedeutung dafür sind im Chemieunterricht unter anderem die Experimentierphasen. Eine geeignete Lernumgebung, die die Kompetenzorientierung konzeptionell umsetzt, stellt das problemorientierte Experimentieren in Kleingruppen mit sogenannten Interaktionsboxen dar (Rumann, 2005; Walpuski, 2006; Habig, 2017). Zusätzlich tragen zentrale Aspekte wie die hohe Schülerorientierung, die Offenheit der Lernumgebung sowie Möglichkeiten für unterschiedliche Unterstützungsformate zu einer adaptiven Lernprozessgestaltung bei (vgl. Altrichter et al., 2009). Bezüglich möglicher Unterstützungsformate liegen zwar bereits Erkenntnisse zum Einfluss der Unterstützungsmaßnahmen „Strukturierungshilfen“, „Feedback“ und „Kommunikationsförderung“ auf den Lernerfolg vor (Walpuski, 2006; Wahser, 2007; Knobloch, 2011), allerdings sind diese bisher noch nicht auf die Unterschiede in der leistungsbezogenen Kleingruppenzusammensetzung bezogen worden. Aus unterschiedlichen Studien können in diesem Zusammenhang besonders homogene Kleingruppen mit geringem Vorwissen (H1), homogene Kleingruppen mit mittlerem Vorwissen und hohen kognitiven Fähigkeiten (H2) sowie heterogene Kleingruppen (H3) für eine systematische Untersuchung herangezogen werden (vgl. u. a. Lou et al., 2000; Gröhlich et., 2009).

Vor diesem Hintergrund wird in diesem Projekt der Frage nachgegangen, wie ein produktiver Umgang mit Leistungsheterogenität im Chemieunterricht insbesondere in Experimentierphasen gestaltet sein muss, um die Schülerinnen und Schüler gezielt fördern zu können. In einem ersten Schritt ist in einer Reanalyse mit Hilfe quantitativer und qualitativer Methoden untersucht worden, inwiefern mit bereits existierenden Unterstützungsmaßnahmen beim Experimentieren in Kleingruppen eine gezielte Förderung einzelner Kleingruppen in Abhängigkeit von ihrer Zusammensetzung möglich ist.

Die Reanalyse wurde mit bereits vorliegenden Leistungs- und Prozessdaten vorangegangener Studien (Walpuski, 2006; Wahser, 2007; Knobloch, 2011) durchgeführt. Auf der Grundlage der Individuelleistungen in den Tests zu Fachwissen und kognitiven Fähigkeiten wurden die Kleingruppen hinsichtlich ihrer Zusammensetzung charakterisiert. Diese leistungsbezogene Kleingruppencharakterisierung wurde anschließend dazu genutzt, um Effekte unterschiedlicher Unterstützungsmaßnahmen auf den fachlichen Lernzuwachs statistisch zu untersuchen. Die Prozessdaten wurden anschließend herangezogen, um mögliche Erklärungsansätze für den besonderen Erfolg bestimmter Unterstützungsmaßnahmen in einer komparativen offenen Analyse (Grounded Theory, vgl. Glaser & Strauss, 2005) herauszuarbeiten.

Die Ergebnisse zeigen einen signifikanten Einfluss der eingesetzten Unterstützungsmaßnahmen auf den Lernzuwachs im Fachwissen für die Kleingruppen H2 ( $F(3, 39) = 4.128$ ;  $p = .012$ ;  $\eta^2 = .241$ ) und H3 ( $F(3, 39) = 4.254$ ;  $p = .011$ ;  $\eta^2 = .247$ ), der

stets zugunsten des Einsatzes von „Feedback“ ausfällt. Für die Kleingruppen H1 lässt sich ein nicht signifikantes, aber ähnliches Bild feststellen. Die Analysen der Prozessdaten legen den Schluss nahe, dass in Kleingruppen mit einer Unterstützung durch „Feedback“ besonders häufig chemische Konzepte in den Lernprozess eingebunden werden und dieser auch viel stärker strukturiert ist.

Da eine gezielte Förderung einzelner Kleingruppen durch alleinige Verwendung des Feedbacks wenig spezifisch ist, wird in einer aktuell laufenden Folgestudie untersucht, welche Unterstützungsmaßnahmen notwendig sind, um gezielt auf die Bedürfnisse einzelner Kleingruppen eingehen zu können. Dabei sollen die Limitationen der oben beschriebenen Studie überwunden werden, indem die Lernprozesse beim Experimentieren mit Interaktionsboxen auf einen Zusammenhang zwischen der leistungsbezogenen Kleingruppenzusammensetzung und typischen Schwierigkeitsbereichen näher untersucht wird. Erste Ergebnisse können ebenfalls auf der Jahrestagung präsentiert werden.

## **Die Kooperation von LehrerInnen und IntegrationshelferInnen im Kontext von schulischer Inklusion**

**Stefanie Czempiel, Bärbel Kracke**

Friedrich-Schiller-Universität Jena, Deutschland; [Stefanie.czempiel@uni-jena.de](mailto:Stefanie.czempiel@uni-jena.de)

„Inklusion ist [...] ein Transformationsprozess, in dem durch Teilhabe statt durch Fürsorge die Verankerung in der eigenen Generation ermöglicht wird“ (Sasse/Lada, 2014). Von dieser Definition ausgehend stellt sich die Frage, ob und wie der Begriff der Inklusion durch das Handeln wesentlicher professioneller Akteure im schulischen System realisiert und wie er reflektiert wird. Eine zentrale Anforderung an LehrerInnen im inklusiven Unterricht ist die Kooperation in multiprofessionellen Teams (Wittek, 2016). Abhängig vom Schulkonzept und der Anzahl der SchülerInnen mit sonderpädagogischem Förderbedarf müssen LehrerInnen konkrete unterrichtsbezogene Zusammenarbeit mit SonderpädagogInnen, SozialpädagogInnen und/oder IntegrationshelferInnen realisieren. IntegrationshelferInnen sind dabei eine personelle Ressource, die laut Gesetzeslage dem einzelnen Kind oder Jugendlichen mit Förderbedarf zugesprochen wird. Im Rahmen der Eingliederungshilfe übernehmen auch pädagogische Laien den beruflichen Auftrag, als IntegrationshelferInnen zu arbeiten (Dworschak, 2010). Dabei ist das berufliche Tätigkeitsfeld noch kaum definiert und es existieren keine professionellen Standards. Bisher wird die Integrationshilfe wortwörtlich als Helferberuf verstanden, obwohl das Ziel, umfassende Teilhabe im Bildungssystem zu sichern, sehr komplex und herausfordernd ist (Knuf, 2013). Wie IntegrationshelferInnen tatsächlich eingebunden werden und welche Handlungsspielräume sie im Unterricht haben, hängt neben dem eigenen Selbstverständnis auch vom Entwicklungsstand der jeweiligen Schule in Bezug auf Inklusion ab. Sommer, Czempiel, Kracke und Sasse (2017) beschreiben vor dem Hintergrund der Diskussion um Qualitätsmerkmale inklusiver Schulen (Landwehr, 2012; Wilbert, 2016; Bertelsmann Stiftung, 2016) drei Entwicklungsstände von Schulen in Bezug auf Inklusion. Sie unterscheiden „Anfänger“, bei denen keine Bedingung inklusiver Schule (Anteil der SuS mit sonderpädagogischem Förderbedarf, Inklusion ist Teil des Schulkonzepts, Kooperation im Kollegium, vorbereitete Lernumgebung, Individualisierung, Leistungskonzept) erfüllt ist von „Fortgeschrittenen“, die alle Merkmale aufweisen. Eine „mittlere Gruppe“ weist 1 bis 6 Merkmale auf.

Das Ziel des vorliegenden Beitrags ist auf der Basis von Interviewaussagen genauer zu beschreiben, wie LehrerInnen und IntegrationshelferInnen miteinander oder nebeneinander arbeiten und ihre Rollen im inklusiven Unterricht reflektieren. Es wird erwartet, dass sich die Beschreibung der Kooperation in den Schulen je nach Stand der inklusiven Schulentwicklung unterscheidet.

Im Rahmen des Projekts „Expertise zum Gemeinsamen Unterricht in der Stadt Jena“ (Sasse, Kracke, Sommer & Czempiel, in Vorbereitung), das die Perspektiven der Pädagogischen Psychologie und der Inklusionspädagogik vor dem Hintergrund der Schulversuche zur Integration (z.B. Eberwein & Knauer, 2009; Preuss-Lausitz, 2009), verbindet, wurden leitfadengestützte Interviews mit 41 LehrerInnen und 16 IntegrationshelferInnen geführt. Diese wurden mittels Qualitativer Inhaltsanalyse nach Mayring (2010) mit einem weitgehend deduktiven Kategoriensystem ausgewertet. Folgende Fragen sind forschungsleitend in der Analyse des Materials: Wie gestalten LehrerInnen und IntegrationshelferInnen unterrichtsbezogene Kooperation im Kontext schulischer Inklusion? Welche Erwartungen an das Gegenüber haben die LehrerInnen und die IntegrationshelferInnen? Welche Rolle und Aufgaben schreiben sie sich selbst und dem Anderen zu? In Bezug auf die Qualität von Kooperation wurden die Ausprägungen der Kooperation von Gräsel, Fussangel und Pröbstel (2006) zugrunde gelegt: Arbeitsteilung – Austausch – Ko-Konstruktion.

Die Ergebnisse zeigen, dass vor allem LehrerInnen der „Anfänger“-Schulen kaum Wissen darüber haben, welche Aufgaben IntegrationshelferInnen übernehmen können und große Unsicherheit darüber besteht, wie eine gelungene unterrichtsbezogene Kooperation gestaltet werden kann. Die Zusammenarbeit lässt sich vor allem als paralleles Arbeiten bzw. als arbeitsteilig beschreiben. Diese Wahrnehmung wird von den IntegrationshelferInnen geteilt. Dagegen bestehen an den „Fortgeschrittenen“-Schulen Routinen und Strukturen, wie IntegrationshelferInnen klassenbezogen eingesetzt werden. Dadurch entwickelt sich eine im Sinne ko-konstruktiven Zusammenarbeitens enge Kooperationsbeziehung, die beiden Akteuren Sicherheit im Handeln gibt, aber auch offen im pädagogischen Prozess beiderseits gestaltet wird.

Die Ergebnisse lassen sich nutzen, um die Qualität des professionellen Handelns von LehrerInnen im inklusiven Unterricht konkreter zu beschreiben. Sie kann darüber hinaus Hinweise zur Gestaltung der LehrerInnenaus- und Fortbildung geben.

## **Gleich und gleich gesellt sich gern? Eine experimentelle Studie über die Effekte von homogener oder heterogener Gruppenformation auf Zufriedenheit und Leistungserfolg**

**Adrienne Müller<sup>1</sup>, Henrik Bellhäuser<sup>2</sup>, Johannes Konert<sup>3</sup>, René Röpke<sup>4</sup>**

<sup>1</sup>Johannes Gutenberg Universität Mainz, Pädagogische Psychologie, Wallstrasse 4-6, 55113 Mainz; <sup>2</sup>Goethe Universität Frankfurt, Pädagogische Psychologie, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; <sup>3</sup>Beuth Hochschule für Technik, FB VI Informatik und Medien, Luxemburger Str. 10, 13353 Berlin; <sup>4</sup>RWTH Aachen University, Informatik 9 (Learning Technologies), Ahornstraße 55, 52074 Aachen; [adriennemueller@yahoo.de](mailto:adriennemueller@yahoo.de)

Aktives und kooperatives Lernen sowie die Bildung von zwischenmenschlichen Beziehungen ist insbesondere im Studium von hoher Relevanz. Lernen in einem solchen Gruppensetting erweist sich als wirkungsvolle didaktische Strategie, um die Leistung von Studierenden zu steigern, da sie Lern- als auch Sozialisationsprozesse anstößt (Magnisalis, Demetriadis, & Karakostas, 2011). Daraus resultieren deutliche Vorteile Wissen zu etablieren (Fernández-Breis, Castellanos-Nieves & Valencia-García, 2009) und das Potential für Forschung und Förderung zeigt sich. Während die Vorteile aus Gruppenarbeit gut belegt sind, ist die Frage nach der Bedeutung der Gruppenkomposition noch weitgehend ungeklärt. Der vorliegende Beitrag beschäftigt sich mit den Kriterien, nach denen eine Gruppe erfolgreich zusammengestellt werden kann, um maximal produktiv und zufrieden mit dem jeweiligen Ergebnis zu sein.

Bisherige Forschung zu Effekten von homogener oder heterogener Gruppenkomposition fand größere Effekte für Persönlichkeitseigenschaften wie Extraversion und Gewissenhaftigkeit als für demografische Eigenschaften wie Geschlecht oder Bildungsniveau (Bell, 2007). Dabei wurden Vorteile für heterogene Verteilung von Extraversion (Kramer, Bhawe, & Johnson, 2014) gefunden, sowie für homogene Verteilung von Gewissenhaftigkeit (Prewett, Walvoord, Stilson, Rossi, & Brannick, 2009). Diese Befunde stammen allerdings aus korrelativen Studien, sodass Rückschlüsse auf kausale Zusammenhänge nicht zulässig erscheinen.

In einem interdisziplinären Projekt zwischen Informatik und Psychologie (Konert, Bellhäuser, Röpke, Gallwas & Zucik, 2016) wurde dazu die Software MoodlePeers entwickelt, welche eine Gruppenformation auf Basis vordefinierter Kriterien erlaubt. Nachdem die gewählten Kriterien in einem Fragebogen für jede Person diagnostiziert wurden, bildet der Algorithmus Gruppen, die entweder homogen oder heterogen in Bezug auf das jeweilige Kriterium sind (Konert et al., 2014).

Eine erste experimentelle Studie konnte zeigen, dass eine Kombination mehrerer Kriterien (Extraversion, Gewissenhaftigkeit, Teamorientierung, Motivation und Vorwissen) zu größerer Zufriedenheit mit der Gruppenzusammenstellung und zu besseren Leistungen in der Gruppenarbeit führte im Vergleich zu zufällig zusammengestellten Lerngruppen (Konert et al., 2016). Offen blieb bei dieser Untersuchung allerdings die Frage, welche der verwendeten Kriterien welchen Anteil an den berichteten Effekten hatten.

Im Vortrag werden die Ergebnisse von zwei neuen Evaluationsstudien berichtet und miteinander verglichen. Beide Studien wurden im Kontext eines Online-Mathematikvorkurses mit angehenden Studierenden durchgeführt, der das Ziel verfolgt, innerhalb von vier Wochen das mathematische Schulwissen zu wiederholen.

In Studie 1 wurden N=424 Studierende (MW Alter = 19,39; SD = 2,20; 92 Frauen, insgesamt 106 Gruppen zu je 4 Personen) in einem vollständig gekreuzten 3x3-Design randomisiert einer von neun Bedingungen zugewiesen; dabei wurden Extraversion und Gewissenhaftigkeit jeweils entweder homogen oder heterogen verteilt oder bei der Gruppenformation ignoriert. Die gebildeten Lerngruppen bearbeiteten gemeinsam wöchentliche Hausaufgaben.

Es zeigte sich, dass heterogene Extraversion zu größerer Zufriedenheit mit der

Gruppenzusammenstellung, mehr Kommunikation zwischen den Gruppenmitgliedern und besseren Ergebnissen in den Hausaufgaben beiträgt. Für Gewissenhaftigkeit zeigte sich, dass Heterogenität (und nicht etwa Homogenität) in den genannten Bereichen positive Auswirkungen hat.

Studie 2 stellt eine Replikationsstudie dar, die vor Beginn des Wintersemesters 2017/18 durchgeführt wird. Vorläufige Ergebnisse scheinen die Ergebnisse von Studie 1 teilweise zu bestätigen.

Die vorliegenden experimentellen Studien weisen zum Teil fehlende Übereinstimmung mit korrelativen Befunden aus der Literatur auf. Aufgrund der hohen Relevanz von Gruppenarbeit für die universitäre Lehre ergibt sich als Forschungsdesiderat, die Gruppenkomposition als wichtige Einflussgröße weiter zu untersuchen, um die zugrundeliegenden Wirkmechanismen besser zu verstehen. Die Open-Source-Software MoodlePeers bietet dafür durch ihren offenen Quellcode ein geeignetes Instrument, das auch bereits in Lehrveranstaltungen Einsatz gefunden hat.

## **Insenzierte Videovignetten zur Förderung der Klassenführung: Entwicklung und Wirkung**

**Rijana Nissing<sup>1</sup>, Julia Bönnte<sup>1</sup>, Gerlinde Lenske<sup>2</sup>, Theresa Dicke<sup>3</sup>, Detlev Leutner<sup>1</sup>**

<sup>1</sup>Universität Duisburg-Essen, Deutschland; <sup>2</sup>Universität Koblenz-Landau, Deutschland; <sup>3</sup>Australian Catholic University, Sydney; [rijana.nissing@uni-due.de](mailto:rijana.nissing@uni-due.de), [julia.boente@uni-due.de](mailto:julia.boente@uni-due.de)

Theoretischer Hintergrund

Im Vergleich zu textbasiertem Unterrichtsmaterial ermöglicht videobasiertes Unterrichtsmaterial eine komplexe und authentische Abbildung des Klassenraumgeschehens (Syring et al., 2015). Unterrichtsvideos sind in der Lage, die tatsächlichen Berufenanforderungen von Lehrkräften annäherungsweise aufzuzeigen (Grossmann et al., 2009; Seidel & Thiel, 2017). Innerhalb der videobasierten Fallarbeit kann damit ein reflexiver Austausch über die Möglichkeiten des didaktischen Handelns frei von Handlungsdruck entstehen (Krammer & Reusser, 2005). Das Medium Video kann aufgrund dieser Charakteristika einen empirischen Zugang für Lehramtsnovizen bieten (vgl. Seidel & Thiel, 2017) und somit als Hilfsmittel zur Erlangung von Klassenführungswissen dienen (ebd.; vgl. König & Kramer, 2016). Gerade im Bereich der Klassenführung fühlen sich Novizen jedoch häufig nicht ausreichend vorbereitet (Evertson & Weinstein, 2006). Das Projekt Professionalisierung für Vielfalt (ProViel), welches in der Qualitätssoffensive Lehrerbildung angesiedelt ist, befasst sich durch das Teilprojekt „Videobasiertes Lehr-Lernmaterial“ (CLIPSS) daher explizit mit der Thematik der Klassenführung.

Klassenführung umfasst dabei die Reduktion von Unterrichtsstörungen bei simultaner Aufrechterhaltung des Unterrichtsflusses (Thiel, 2016; Seidel & Thiel, 2017). Die Qualität der Klassenführung gilt als Prädiktor für Unterrichtsqualität, der sich spezifischer und konkreter auf die Unterrichtspraxis bezieht als generelles pädagogisches Wissen (König & Kramer, 2016; vgl. auch Hattie, 2009; Voss, Kunter, Seiz, Hoehne & Baumert, 2014; Lenske et al., 2016; Kramer, König, Kaiser, Ligtvoet & Blömeke, 2017). Da Klassenführung zudem stark vom Erfahrungsschatz der Lehrkraft abhängt (König & Kramer, 2016), gestaltet sich die Handhabung von Unterrichtsstörungen für Novizen als herausfordernd (Dicke, Elling, Schmeck & Leutner, 2015). Im CLIPSS-Projekt werden daher Unterrichtsvideovignetten entwickelt, welche den Erwerb von Klassenführungswissen insbesondere bei Lehramtsnovizen fördern.

Fragestellung

Bei der Videoentwicklung steht die Frage im Mittelpunkt, wie Unterrichtsvideos trotz Inszenierung einen hohen Grad an Authentizität aufweisen können und zugleich die gezielte Darstellung von Klassenführungsstrategien abbilden. Die Wirkungsweise der Videovignetten wird anschließend in zwei anknüpfenden Forschungsvorhaben untersucht: Ein Forschungsschwerpunkt liegt auf der Förderung von nonverbalen Klassenführungsstrategien durch Videovignetteninsatz, ein weiterer auf der Komplexitätsreduktion bei Videovignetten zur Verbesserung der Lernwirksamkeit.

Methode

Damit Unterrichtsvideos von authentischen Klassenraumsituationen entstehen können, wurden zunächst Drehbücher auf Basis von Schulhospitationen und realen Unterrichtsmitschnitten erstellt. Aktuell werden die Drehbücher innerhalb von Theater-AGs in der Primar- und Sekundarstufe inszeniert und aus mehreren Perspektiven (Lehrer-, Schüler- und Total-Kamera) videografiert. Diese sogenannten Staged Videos besitzen den Vorteil, dass das Videomaterial spezifisch veränderbar und beeinflussbar ist (vgl. Seidel & Thiel, 2017). Zudem sind datenschutzrechtliche Hürden für die Nutzung des Videomaterials irrelevant. Das Vorgehen ermöglicht eine Fokussierung auf professionsrelevante Anforderungen im Lehrerberuf sowie einen

Querschnitt der einzelnen Klassenführungsdimensionen. Als theoretische Grundlage dienen die Klassenführungsdimensionen nach Lenske und Mayr (2015) sowie die Techniken der Klassenführung nach Kounin (1976). Innerhalb von Theater-AGs wird der Unterricht inszeniert, sodass datenschutzrechtlich unbedenkliches Videomaterial entsteht. Die Dimensionenpassung und der Authentizitätsgrad werden prozessbegleitend von ExpertInnen bewertet. Nach der Dreharbeit wird das Videomaterial technisch aufbereitet, um im Anschluss daran die Wirkungsweise in verschiedenen Vergleichsgruppen zu beleuchten.

#### Ergebnisse

Seit Herbst 2017 stehen erste technisch aufbereitete Unterrichtsvideovignetten für Lehrzwecke zur Verfügung. Die anknüpfenden Studien werden im WiSe2017/2018 und SoSe2018 durchgeführt. In verschiedenen Seminargruppen werden die abgebildeten Klassenführungsstrategien der Lehrkräfte innerhalb der Unterrichtsvideovignetten beleuchtet, sodass ein reflexiver Diskurs über die Qualität des didaktischen Handelns entstehen kann. Die Authentizität des Videomaterials wird hier durch die Studienteilnehmenden eingeschätzt. Erwartet wird, dass das Videomaterial eine gezielte Fokussierung auf einzelne Klassenführungsfacetten ermöglicht und von den Studienteilnehmenden authentisch eingeschätzt wird. Erste Aussagen über die Wirkungsweise der Unterrichtsvideovignetten können daher im Frühjahr 2018 getroffen werden.

### **Einstellungen zur Theorie- und Praxisorientierung des Lehramtsstudiums im Praxissemester**

**Thomas Fischer<sup>1</sup>, Bach Andreas<sup>2</sup>, Kathrin Rheinländer<sup>1</sup>**

<sup>1</sup>Europa-Universität Flensburg, Deutschland; <sup>2</sup>Paris-Lodron-Universität Salzburg; [andreas.bach@sbg.ac.at](mailto:andreas.bach@sbg.ac.at)

Einstellungen bzw. Überzeugungen bilden eine bedeutsame Komponente der professionellen Kompetenz von Lehrkräften (Baumer & Kunter, 2006). In vorliegenden Untersuchungen wurden häufig domänenspezifische Überzeugungen untersucht (Reusser & Pauli, 2014). Weitaus seltener sind generelle berufsbezogene Einstellungen in Blick genommen worden. Mit der Einführung von Praxissemestern wurde die Verbesserung der Integration von Theorie und Praxis intendiert. Es liegen bislang wenig Befunde dazu vor, welche Effekte die Teilnahme von Lehramtsstudierenden an Praxissemestern auf die Veränderungen von Einstellungen hat. Der vorliegende Beitrag greift dieses Desiderat auf und fokussiert die Einstellungen zur Theorie- und Praxisorientierung des Lehramtsstudiums im Praxissemester vor dem Hintergrund folgender Forschungsfragen: 1. Über welche Einstellungen zur Theorie- und Praxisorientierung des Lehramtsstudiums verfügen die Studierenden? 2. Wie verändern sich diese im Verlauf des Praxissemesters?

Die Datenbasis stammt aus dem Projekt „Intensität und Stabilität von berufsbezogenen Einstellungen im Lehramtsstudium“, das als quantitative Befragung von Studierenden im Längsschnittdesign durchgeführt wird. Die Stichprobe der vorliegenden Prä-Post-Messungen umfasst die Kohorte der Masterstudierenden im Praxissemester, das an der Europa-Universität Flensburg im 3. Fachsemester durchgeführt wird. Insgesamt nahmen 239 Personen von 304 Studierenden am Prä-Messzeitpunkt teil. 225 konnten für die Post-Erhebung erneut befragt werden. Die Theorie- und Praxisorientierung des Lehramtsstudiums wurde entlang der Dimensionen (1) theoretisch-reflexive Professionalisierungsfunktion des Lehramtsstudiums“ (3 Items,  $\alpha = .76$ ) sowie (2) berufseinführende Professionalisierungsfunktion des Lehramtsstudiums“ (4 Items,  $\alpha = .94$ ) operationalisiert und gemessen wurden. Die Faktorenstruktur der Skala wurde mittels konfirmatorischer Faktorenanalyse (CFA) überprüft.

Die Ergebnisse zeigen, dass sich die Theorie- und Praxisorientierung des Lehramtsstudiums durch Likertskalen auf Einstellungsebene operationalisieren lassen. Für das spezifizierte Modell konnten akzeptable Modell-Fit-Werte ermittelt werden. Ingesamt befürworten die Studierenden die berufseinführende Professionalisierungsfunktion des Lehramtsstudiums vor der theoretisch-reflexiven Professionalisierungsfunktion zum Prä-Messzeitpunkt. Die Ergebnisse der Datenanalysen zeigen, dass es keine signifikanten Veränderungen der Einstellungen zum Theorie- und Praxisorientierung zum Lehramtsstudium innerhalb des zehnwöchigen Praxissemesters gibt. Die Implikationen der Ergebnisse für die Wirksamkeit von verlängerten Praxisphasen sollen vor dem Hintergrund der Limitationen der Untersuchung diskutiert werden.

### **Polyvalenzwahrnehmung und Polyvalenzattraktivität im Studium der Wirtschaftspädagogik – eine quantitative Analyse**

**Clemens Frötschl, Benjamin Jäcklin, Karin Heinrichs**

Otto-Friedrich-Universität Bamberg, Deutschland; [clemens.froetschl@uni-bamberg.de](mailto:clemens.froetschl@uni-bamberg.de), [karin.heinrichs@uni-bamberg.de](mailto:karin.heinrichs@uni-bamberg.de)

#### Hintergründe

Die Studienwahl legt den Grundstein für die Wahl späterer beruflicher Tätigkeit. Allerdings soll-te ein Studium auf den flexiblen Einsatz in (angrenzenden) beruflichen Aufgaben vorbereiten. Bei reinen Lehramtsstudiengängen fehlt diese so genannte Polyvalenz. Hier ist die Studienwahl vielmehr häufig gleichzeitig Berufswahl. So untersucht Rauin (2007, 61f.) in einer Längsschnitt-studie über 15 Jahre Studierende in Lehramtsstudiengängen an pädagogischen Hochschulen in Baden-Württemberg. Die Ergebnisse zeigen, dass viele Lehramtsstudierende das Studium trotz selbst eingeschätzter Nichteignung für die spätere Lehrtätigkeit weiter fortführen, um eine fi-nanzielle Absicherung bei Verbeamtung zu erreichen. Umgekehrt werden Programme aufgelegt, die in Phasen geringer Einstellungschancen Lehramtsstudierenden Weiterqualifikationen anbieten und die Vorbereitung auf außerschulische Handlungsfelder nachgeholt wird, z.B. Initiativen an bayerischen Universitäten vergleichbar zum Bamberger Projekt „Move On“. Das Studium der Wirtschaftspädagogik ist dagegen bereits traditionell durch Polyvalenz gekennzeichnet (DGfE 2014). Absolventen wirtschaftspädagogischer Studiengänge werden systematisch auf diverse Betätigungsfelder vorbereitet. Neben der Tätigkeit als Lehrpersonen an berufsbildenden Schulen stehen Wirtschaftspädagog/innen auch Berufe in Unternehmen, an Hochschulen oder in der (Bildungs-)Verwaltung offen.

#### Fragestellung

Empirisch unklar ist allerdings, ob die polyvalente Struktur wirtschaftspädagogischer Studiengänge auch aus Sicht der Studierenden relevant ist, inwiefern das polyvalente Angebot die Studienwahl beeinflusst, inwiefern sich Tätigkeitspräferenzen im Laufe des Studiums ändern oder welchen Einflüssen diese unterliegen. Es mangelt an empirischer Evidenz, inwiefern die Polyvalenz bei der Studienwahl, während des Studiums und beim Übergang in den Beruf Relevanz entfaltet. Erste Einblicke liefert eine Querschnittstudie aus Österreich (Stock, Fernandez & Schelch 2008). Die Übertragbarkeit auf wirtschaftspädagogische Studiengänge, die auf eine zweiphasige Lehrerbildung ausgerichtet sind, ist aufgrund der einphasigen Struktur in Österreich fraglich. So gilt es, vergleichbare Fragen im deutschen System der zweiphasigen Lehrerausbildung zu untersuchen und die Relevanz der Polyvalenz aus Sicht der Studierenden und für deren Studienwahl zu explorieren.

#### Methode

Entsprechend wurden an der Otto-Friedrich-Universität Bamberg Studierende der Wirtschaftspädagogik zur Polyvalenzattraktivität, zum präferierten Tätigkeitsbereich, zu theoretisch relevanten Einflussfaktoren der Berufswahl (Lent, Brown & Hackett 1994) und zu möglichen Einflussfaktoren während des Studiums auf die berufliche Präferenz konsultiert. Es liegen Daten von insgesamt 215 Studierenden vor, von denen sich 163 Studierende im Masterstudium befinden (1.-4. Semester). Erste Analysen sollen die folgenden Fragen adressieren:

- Spielt die Polyvalenz für die Berufswünsche der Studierenden der Wirtschaftspädagogik an der Otto-Friedrich-Universität Bamberg eine Rolle?
- Wird diese Mehrwertigkeit des Abschlusses im Studienverlauf von den Studierenden erkannt?

Zur Beantwortung der Fragen werden unter anderem Skalen zur Laufbahnunsicherheit (Cronbachs  $\alpha = ,909$ ) sowie zur beruflichen Selbstwirksamkeit im Studium (Cronbachs  $\alpha = ,841$ ) eingesetzt. Zur Erfassung der wahrgenommenen Attraktivität der Polyvalenz wurde eine neue Skala entwickelt (Beispielitem: „Die Möglichkeit, eine Tätigkeit außerhalb der Schule aufnehmen zu können, macht den Studiengang für mich attraktiv“; Cronbachs  $\alpha = ,730$ ). Zu allen Skalen liefern die Studierenden Selbsteinschätzungen auf einer jeweils sechsstufigen Antwortskala („trifft gar nicht zu“ bis „trifft voll zu“).

#### Ergebnisse

Erste Analysen zeigen, dass Studierende der Wirtschaftspädagogik die Polyvalenz des Studiums weitgehend als attraktiv bewerten ( $M = 4,51$ ;  $SD = ,82$ ). Es gibt einen nicht zu vernachlässigenden Anteil an Studierenden, die klar äußern, nicht Lehrkraft an einer beruflichen Schule werden zu wollen ( $t_1$ : 17%,  $t_2$ : 28%). Demgegenüber stehen Teilnehmende, die eine Tätigkeit als Lehrperson anstreben. Selbst letztgenannte Personen messen der Polyvalenz Attraktivität bei ( $M = 4,23$ ,  $SD = 1,02$ ).

Im weiteren Analysen wird derzeit untersucht, inwieweit sich Profile des Antwortverhaltens der Teilnehmenden mit Blick auf die angestrebten beruflichen Handlungsfelder identifizieren lassen und geprüft, inwiefern es im Verlauf der Längsschnittstudie zu (signifikanten) Veränderungen hinsichtlich der angestrebten Tätigkeit kommt bzw. welchen (personellen und strukturellen) Einflüssen diese unterliegen. Über die zugehörigen Ergebnisse soll im Rahmen der Ergebnispräsentation ebenfalls berichtet werden.

### **Der Einfluss kognitiver und nicht-kognitiver Merkmale angehender Ärzt\*innen auf die Gesprächsführungskompetenz in simulierten Anamnesegesprächen**

**Kristina Schick<sup>1</sup>, Martin Gartmeier<sup>1</sup>, Sigrid Harendza<sup>2</sup>, Martina Kadmon<sup>3</sup>, Pascal O. Berberat<sup>1</sup>**

<sup>1</sup>Technische Universität München, TUM School of Medicine, TUM Medical Education Center, Lehrstuhl für Medizindidaktik, medizinische Lehrentwicklung und Bildungsforschung, München, Deutschland; <sup>2</sup>III. Medizinische Klinik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Deutschland; <sup>3</sup>Medizinische Fakultät, Universität Augsburg, Deutschland; [kristina.schick@tum.de](mailto:kristina.schick@tum.de)

#### Theoretischer Hintergrund

Gute ärztliche Gesprächsführung ist eine wichtige Voraussetzung einer qualitativ hochwertigen Patient\*innenversorgung (Fong Ha, Surg Anat, & Longnecker, 2010). Bereits während der medizinischen Ausbildung erhält die Vermittlung von Gesprächsführungskompetenzen daher besondere Aufmerksamkeit (NKLM; 2015). Studierenden werden in Seminaren zur ärztlichen Gesprächsführung Kommunikationstechniken u.a. mithilfe von standardisierten Schauspielpatient\*innen vermittelt (Jünger & Köllner, 2003). Dabei stellen sich für die Gestaltung der medizinisch-kommunikativen Ausbildung grundsätzlich bedeutsame Fragen: Ist es sinnvoll, dabei primär Kommunikationstechniken zu fokussieren? Oder wäre es nicht grundsätzlich auch notwendig, die Entwicklung diesbezüglich günstiger Persönlichkeitsstrukturen von Studierenden zu fördern? Und wie beeinflusst medizinisches Fachwissen generell die Entwicklung kommunikativer Kompetenzen?

In der Literatur werden nicht-kognitive Attribute (z.B. Persönlichkeit, Empathie), aber auch kognitive Fähigkeiten (z.B. das Fachwissen) als Prädiktoren der Performanz der Studierenden während der medizinisch-kommunikativen Ausbildung genannt (Hojat, Erdmann, & Gonnella, 2013). Bisher liegt jedoch wenig belastbare Evidenz dazu vor: Haight, Chibnall, Schindler, und Slavin (2012) konnten zeigen, dass die Performanz in akademischen Tests als Prädiktor für die kognitiven Fähigkeiten von Medizinstudierenden geeignet ist. Des Weiteren zeigte die Forschungsgruppe, dass Persönlichkeitsfacetten wie Extraversion, Gewissenhaftigkeit und Empathie als nicht-kognitive Attribute die Performanz bedingen können. Jedoch konnte insgesamt bisher nur unzureichend gezeigt werden, wie die Persönlichkeitsstruktur und das Fachwissen die ärztliche Gesprächsführungskompetenz beeinflussen. Vorangegangene Studien setzten zur Erfassung der ärztlichen Gesprächsführungskompetenz Audioaufzeichnungen oder Fragebögen ein (Chapmann, Duberstein, Epstein, Fiscella, & Kravitz, 2008; Duberstein, Meldrum, Fiscella, Shields, & Epstein, 2007). Die Einschätzung der Gesprächsführungskompetenz erfolgte zum Teil durch ein intuitives Rating zuvor nicht-geschulte Rater\*innen (Duberstein et al., 2007). Die Performanz wurde in diesen Fällen nur mit den Persönlichkeitsmerkmalen in Beziehung gesetzt. Das Fachwissen wurde in keiner der Studien berücksichtigt. Ausgehend von diesen vorangegangenen Forschungsergebnissen erfolgte nun eine strukturierte Beobachtung der simulierten Anamnesegespräche mittels Videoaufzeichnung (Janik, Seidel, & Najvar, 2009). Für die Einschätzung der Gesprächsführungskompetenz wurden die Rater\*innen anhand eines etablierten Konzeptes zur Gesprächsführung geschult, sodass die Subjektivität der Einschätzung reduziert wurde (Langer & Schulz von Thun, 1974, 2007). Mithilfe dieser Methodik soll untersucht werden, wie das Fachwissen und die Persönlichkeit angehender Ärzt\*innen die ärztliche Gesprächsführungskompetenz beeinflusst.

#### Fragestellung

In der Studie werden die folgenden Fragestellungen untersucht:

- (1) Welche Aspekte der Persönlichkeit sind Prädiktoren für Gesprächsführungskompetenz in Anamnesegesprächen?
- (2) Beeinflusst das Maß an Fachwissen die Gesprächsführungskompetenz angehender Ärzt\*innen?

#### Methode

Die Beantwortung der Forschungsfragen erfolgt auf Basis der Daten aus dem Kooperationsprojekt ÄKHOM – Ärztliche Kompetenzen Hamburg Oldenburg München. In diesem Projekt durchliefen  $N = 70$  Medizinstudierende einen simulierten ersten Arbeitstag als Assistenzarzt/-ärztin. Die Teilnehmenden befanden sich zum Teil im letzten klinischen Studienjahr und zum Teil bereits im praktischen Jahr der Medizinausbildung. Sie führten zu Beginn des Assessments fünf Anamnesegespräche mit standardisierten Patient\*innen. Diese Gespräche wurden videographiert und anschließend mit der deutschen Version des Kalamazoo Communication Skills Assessment Form (KCSAF-d) durch geschulte Rater\*innen ausgewertet. Der KCSAF-d



umfasst 9 Kompetenzfacetten „Ermöglichte Beziehungsaufbau“, „Eröffnete das Gespräch“, „Sammelte Informationen“, „Versetzt sich in die Lage der Patient\*in“, „Teilt Informationen“, „Erzielt Einigung“, „Beendet das Gespräch angemessen“, „Zeigt Empathie/Einfühlungsvermögen“ und „Vermittelt korrekte Informationen“. Zusätzlich wurde ein Wissenstest auf dem Niveau der zweiten ärztlichen Prüfung mit 100 Fragen zu sämtlichen Fachbereichen durchgeführt. Die Persönlichkeitsstruktur wurde mittels des Neo-FFI erhoben. Für die Erfassung der Empathiefähigkeit wurde der Interpersonal Reactivity Index Empathy Questionnaire (IRI) eingesetzt.

Mittels eines linearen Regressionsmodells werden die Daten ausgewertet. Die abhängige Variable stellt die gemessene Gesprächsführungskompetenz mittels des KCSAF-d dar. Die unabhängigen Variablen sind die Empathiefähigkeit (IRI), die Persönlichkeitsfacetten des Neo-FFI und das Fachwissen.

Ergebnisse

Aktuell wurde die Datenerfassung abgeschlossen und mit der Datenaufbereitung und Datenauswertung begonnen. Es wird erwartet, dass sowohl das Fachwissen die Persönlichkeitsfacetten „Extraversion“ und „Gewissenhaftigkeit“ als auch die Empathiefähigkeit positiv mit der Gesprächsführungskompetenz korrelieren. Im Februar 2018 werden die Ergebnisse der Untersuchung vorliegen.

### **Praktikumskulturen in Dokumenten und Dialogen. Dokumentenanalyse und Ratingkonferenzen als Instrumente zur Erfassung differenter Praktikumskulturen von Lehrerinnen- und Lehrerbildungseinrichtungen**

**Anja Winkler<sup>1</sup>, Miriam Grüning<sup>2</sup>, Tina Hascher<sup>1</sup>, Sibylle Rahm<sup>2</sup>**

<sup>1</sup>Universität Bern, Schweiz; <sup>2</sup>Universität Bamberg, Deutschland; [anja.winkler@edu.unibe.ch](mailto:anja.winkler@edu.unibe.ch), [miriam.gruening@uni-bamberg.de](mailto:miriam.gruening@uni-bamberg.de)

In der Lehrerbildungsforschung gelten Praktika als zentrales Element für die Entwicklung professionellen Handelns von Lehrpersonen. Sie unterscheiden sich jedoch wesentlich hinsichtlich ihrer Gestaltung (z. B. Arnold et al., 2011). Neben formalen Unterschieden (z.B. Anzahl, Dauer, Platzierung im Studienverlauf) bestehen Differenzen in den Praktikumskulturen. In Praktikumskulturen spiegeln sich Verhaltenserwartungen und Handlungsorientierungen der beteiligten Akteure wider (Helsper, 2010). Sie stellen deshalb eine wesentliche Grundlage für die Lernerfahrungen der Studierenden dar und müssen in der Interpretation von Forschungsergebnissen berücksichtigt werden. Wenig ist jedoch darüber bekannt, was Praxiskulturen sind und wie sie sich erfassen lassen.

Hier setzt unsere Studie an, die sich im Rahmen des binationalen Forschungsprojektes COPRA (Coaching im Praktikum) verortet. Im Projekt COPRA geht es um die Gewinnung von Erkenntnissen zur Wirksamkeit von Coaching in schulpraktischen Studien an mehreren Lehrerbildungsstandorten in der Schweiz und in Deutschland. Die Praktikumskulturen von vier ausgewählten Institutionen wurden anhand zweier qualitativer Methoden untersucht, da sich vor allem qualitative Methoden anbieten, um organisationale Identitäten zu rekonstruieren. (a) Mit Hilfe einer Dokumentenanalyse (N = 34 Dokumente) wurden die spezifischen Rahmenbedingungen untersucht, um die Grundhaltungen und Orientierungen des jeweiligen Standortes zu identifizieren (Mayring, 2002). Im Bereich der schulpraktischen Anteile wurde dabei beachtet, dass diese auf drei Ebenen im Systemzusammenhang gesteuert werden: auf der Makroebene der Ordnungen und Reglemente, der Mesoebene der Organisationseinheit und Praktikumsseinrichtungen und auf der Mikroebene der Handlungsanweisungen (Rolf, 2013).

(b) Anhand von Ratingkonferenzen mit AkteurInnen der LehrerInnenbildung (N=24 Studierende, N=22 Praxislehrpersonen, N=25 Dozierende) wurden die individuellen Sichtweisen und Orientierungen analysiert und somit die sog. organisationalen Identitäten der Institutionen rekonstruiert. Als eine Kombination aus Kurzfragebogen und Gruppeninterview ermöglicht die Ratingkonferenz eine direkte kommunikative Validierung der zuvor gegebenen Antworten (Keller, Heinemann & Kruse, 2012).

Die Vorteile der Kombination von Dokumentenanalysen und Ratingkonferenzen lassen sich aus empirischer Perspektive mit der Triangulation begründen. Das heißt, ein Forschungsgegenstand wird aus unterschiedlichen Perspektiven betrachtet und analysiert. Damit können die Stärken der einzelnen Methoden sich in ihren Stärken ergänzen und die Schwächen können wechselseitig aufgezeigt werden (Flick, 2004). In der Auswertung des Materials zeigt sich, was Fend (2006) unter der „Rekontextualisierung“ versteht: Die in den Dokumenten festgelegten Handlungsanweisungen durch die Institution werden durch ihre jeweiligen AkteurInnen individuell reinterpretiert. Hinweise darauf geben die Aussagen der Dozierenden, Praxislehrpersonen und Studierenden bei den Ratingkonferenzen.

Am Beispiel der Mitbestimmung Studierender bei der Praktikumsgestaltung soll gezeigt werden, wie die Kombination beider Methoden und damit unterschiedliches Datenmaterial als Interpretationsgrundlage zur Erstellung von Institutionsprofilen verwendet werden kann.

### **Antizipierte Gelingensbedingungen der "School-Cloud". Eine Delphi-Studie**

**Jürgen Schneider, Samuel Merk, Thorsten Bohl**

Universität Tübingen, Deutschland; [juegen.schneider@ife.uni-tuebingen.de](mailto:juegen.schneider@ife.uni-tuebingen.de), [samuel.merk@uni-tuebingen.de](mailto:samuel.merk@uni-tuebingen.de)

Schulen stehen im Zuge der Digitalisierung vor großen Herausforderungen. Zum einen ermöglichen neue Medien eine weitere Pluralisierung der Didaktik und Unterrichtsentwicklung, die über das momentan vorherrschende Verfassen von Texten oder die Informationssuche im Internet hinausgehen (Medienpädagogischer Forschungsverbund Südwest, 2017). Wie können neue Medien (z.B. Tablets) jedoch so eingesetzt werden, dass sie nicht Selbstzweck sind, sondern lernförderlich wirken (Bastian & Aufenanger, 2017)?

Zum anderen fordert die Dynamik von Digitalisierung eine bisher wenig beachtete Dimension der Infrastrukturentwicklung und -pflege. Die Auswahl, Beschaffung und Wartung der Hardware sowie insbesondere die Einhaltung sich stets wandelnder Sicherheits- und Datenschutzerfordernungen stellt das (zumeist nicht professionell dafür ausgebildete) zuständige Personal vor große Aufgaben.

Ein Versuch die Barrieren zu senken stellt das Projekt "Schul-Cloud" des Hasso Plattner Instituts dar. Die Schul-Cloud entlastet Lehrpersonen bei der Wartung schuleigener Server, lokaler Rechner und Software indem sie webbasierte Dienste, Rechenressourcen und Speichermöglichkeiten per Cloud Computing anbietet. Dies ermöglicht veraltete und somit unsichere Software zu vermeiden und sensible Daten zu schützen, da die Cloud Infrastruktur zentral und für viele Schulen simultan von professionellen IT-Technikern entwickelt und gepflegt wird. Aktuell wird eine solche Schul-Cloud an 27 Schulen des MINT-EC-Schulnetzwerks bundeslandübergreifend in einem Pilotprojekt eingesetzt. Dieses nationale Exzellenz-Netzwerk technikaffiner

Schulen setzt sich zum Ziel, die Qualität der MINT-Bildung in Deutschland zu fördern, indem durch Pilotprojekte Leuchtturmschulen gebildet werden.

Der Zugang zur Cloud erfolgt anhand einfacher Endgeräte, die lediglich über einen Bildschirm und einen Webbrowser verfügen müssen um auf eine Vielzahl von Diensten zugreifen zu können. Zu diesen zählt die Bereitstellung von Lerninhalten, die Möglichkeit Lehrmaterialien zu erstellen, Dateimanagement, Hausaufgabenverwaltung und Kollaborationsmöglichkeiten. Solche Dienste über ein Cloudsystem für Schulen und Lehrpersonen nutzbar zu machen entspricht einem Novum und wurde bisher nicht systematisch untersucht. Neue Medien sind jedoch nicht Selbstzweck, sondern ein Tool (Mayer, 2014), dessen Anwendungsgebiete und Gelingensbedingungen exploriert werden müssen. Welche Dienste der Schul-Cloud in welcher Weise und unter welchen Bedingungen für Lehrpersonen von Nutzen sein können stellt somit ein Desiderat dar.

Im Forschungsprojekt soll in einer ersten Forschungsfrage exploriert werden, welche Dienste der Schul-Cloud in welchen Bereichen des schulischen Alltags zum Einsatz kommen und welchen Gelingensbedingungen sie unterliegen. Weiterhin soll in einer zweiten Forschungsfrage untersucht werden, welche didaktischen Arrangements durch die inhaltlichen Schul-Cloud Dienste im Unterricht möglich sind und welchen Gelingensbedingungen diese unterliegen.

Die Forschungsfragen zu Anwendungsbereichen und didaktischen Arrangements der Schul-Cloud werden anhand zweier Delphi-Studien untersucht. Delphi-Studien dienen dazu, Lösungsmöglichkeiten für komplexe Fragestellungen (z.B. Zukunftsszenarien) zu liefern (Häder, 2014). Die Lösungsmöglichkeiten entstehen als Produkt einer mehrphasigen Gruppenkommunikation zwischen Experten, deren Arbeitsbereiche sich auf unterschiedliche Weise mit dem fokussierten Feld der Fragestellung überschneiden. In einer ersten Phase formulieren die Experten einzeln Lösungsmöglichkeiten, die durch ein Leitungsteam strukturiert und gesammelt an alle Experten zurückgespielt werden. So können diese die erweiterten Lösungsmöglichkeiten überarbeiten, konkretisieren und der Bedeutung nach ordnen. Das Forscherteam orchestriert den Austausch, die Überarbeitung und die Strukturierung der Lösungsmöglichkeiten und arbeitet abschließend einen umfassenden Katalog zur Beantwortung der Fragestellung aus.

Vorliegend stammen die Expert/innen aus unterschiedlichen Bereichen, die jeweils Berührungspunkte mit der Schul-Cloud aufweisen. Hierzu zählen die Lehrpersonen der 27 MINT-EC Schulen, die Erfahrung mit dem Einsatz der Schul-Cloud gemacht haben. Weiterhin werden Expert/innen aus der Wissenschaft die Fragen auf Basis des aktuellen Forschungsstands und theoretischer Modelle bearbeiten. Vertreter aus der Schulverwaltung hingegen, haben vermehrt strukturelle Merkmale und Rahmenbedingungen im Blick. Letztlich weisen an der Schul-Cloud beteiligte Entwickler\*innen des Hasso-Plattner-Instituts Detailwissen über die Merkmale der Dienste selbst auf.

Bis zur Konferenz kann bezüglich der ersten Forschungsfrage eine nach Bedeutung geordnete Liste der Einsatzbereiche für die Dienste der Schul-Cloud mitsamt deren antizipierter Gelingensbedingungen dargestellt werden. Bezüglich der zweiten Forschungsfrage können erste, anhand der Delphi-Studie antizipierte Anwendungsszenarien im Unterricht dargestellt werden.

## **Erfassung kognitiver Lernstrategien mit Blickbewegungsdaten. Sensitivität von Indikatoren auf Textebene**

**Tristan Haselhuhn, Klaus-Peter Wild**

Universität Regensburg, Deutschland; [tristan.haselhuhn@ur.de](mailto:tristan.haselhuhn@ur.de)

Kognitive Lernstrategien nehmen eine wichtige Rolle bei der Untersuchung selbstgesteuerten Lernens ein und können nach Wiederholungs-, Elaborations- und Organisationsstrategien klassifiziert werden (Weinstein, Acee & Jung, 2011; Wild, 2010). Die Diagnostik kognitiver Lernstrategien stützt sich bislang mehrheitlich auf Selbstberichtsdaten. Mit der Zugänglichkeit der Eye-Tracker Technologie eröffnet sich prinzipiell die Möglichkeit, entsprechende Lernprozessdaten auch unabhängig von Selbstberichten zu operationalisieren. Aus psychometrischer Sicht stellt sich die Frage, ob auf Basis von Blickbewegungsdaten reliable und valide Indikatoren zur Erfassung kognitiver Lernstrategien gewonnen werden können (Catrysse et al., 2016). Da sich kognitive Lernstrategien konzeptionell besonders trennscharf auf das Lernen mit Texten anwenden lassen, wurde diese Fragestellung im Rahmen einer experimentellen Lesestudie mit spezifischen Lernstrategieinstruktionen untersucht. Um von Blickbewegungen auf ablaufende Kognitionen zu schließen, wird eine Interpretation des Blickbewegungsmusters beim Einsatz der Lernstrategie benötigt. Dies kann über Indikatoren erfolgen, die sensitiv auf die Anwendung einer oder mehrerer Lernstrategien reagieren. Aktuelle Studien legen verschiedene, potentiell geeignete Indikatoren zur Lernstrategieerfassung nahe: a) „Lesezeit“ (Summe der Fixationszeiten je Textabschnitt), als Maß für die Dauer der kognitiven Beschäftigung mit dem Textabschnitt; b) „Second Pass“, das wiederholte Lesen einzelner Textbestandteile, nachdem der Text einmal komplett gelesen wurde (Mason, Pluchino, Tornatora & Ariasi, 2013); c) „Looking back“, das wiederholte Lesen von bereits gelesenen Sätzen (Penttinen, Anto & Mikkilä-Erdmann, 2013).

Auf dem Hintergrund der bisherigen Forschung wird davon ausgegangen, dass sich die Verwendung spezifischer kognitiver Lernstrategien (hier „Wiederholungsstrategien“ und „Elaborationsstrategien“) in jeweils lernstrategiespezifischen Blickbewegungsmustern niederschlägt. Es wird vermutet, dass die oben benannten Indikatoren in differentieller Weise zur Erfassung von kognitiven Lernstrategien sensitiv und damit auch indikativ sind.

Das Untersuchungsdesign der Studie folgt der Annahme, dass sich das Lernverhalten in einer Lesesituation durch eine lernstrategiespezifische Instruktion gezielt verändern lässt (de Boer, Donker & van der Werf, 2014). Unterschiede in den Blickbewegungsdaten zwischen Gruppen mit unterschiedlichen Instruktionen sollten sich daher durch unterschiedliches Lernverhalten erklären lassen. Jeweils größere Effektstärken eines Indikators sollten auf eine höhere diagnostische Relevanz verweisen.

An der Studie nahmen 69 Studierende teil, die drei experimentellen Instruktionsgruppen (Wiederholungsstrategien, Elaborationsstrategien und Kontrollgruppe) ausbalanciert hinsichtlich Intelligenz (PSB-R 6-13; Horn, 2003) und Leseverständnis (LGVT 6-12; Schneider, Schlagmüller & Ennemoser, 2007) zugeordnet wurden. Die Kontrollgruppe erhielt keine lernstrategiespezifische Leseinstruktion, sondern wurde aufgefordert, den präsentierten Text wie einen Zeitungsartikel zu lesen. Die Gruppen „Elaborationsstrategien“ und „Wiederholungsstrategien“ erhielten eine ausführliche Instruktion zum gewünschten Lernverhalten. Jede Gruppe hatte zwei Texte zu lesen. Die Reihenfolge war ausbalanciert und Gegenstand der statistischen Analysen. Eine doppelblinde Durchführung sollte Versuchsleitereffekte vermeiden.

Der Indikator „Lesezeit“ zeigt einen starken Effekt und signifikante Unterschiede zwischen den Gruppen „Wiederholungsstrategien“ und „Elaborationsstrategien“, sowie zwischen „Wiederholungsstrategien“ und „Kontrolle“. Die Gruppe „Wiederholungsstrategien“ wies die höchsten Lesezeiten auf. Eine hohe Effektstärke verweist auf die diagnostische Relevanz dieses Indikators. Der Indikator „Second Pass“ zeigt einen starken Effekt und signifikante Unterschiede zwischen den Gruppen „Wiederholungsstrategien“ und „Elaborationsstrategien“ sowie zwischen „Wiederholungsstrategien“ und „Kontrolle“.

Die höchsten „Second Pass“ Zeiten zeigten sich für die Gruppe „Wiederholungsstrategien“. Durch die hohe Effektstärke scheint auch hier eine diagnostische Relevanz des Indikators gegeben. Für den Indikator „Looking-Back“ wurde ein schwacher Effekt gefunden; signifikante Unterschiede zeigen sich zwischen den Gruppen „Wiederholungsstrategien“ und „Elaborationsstrategien“ sowie zwischen den Gruppen „Wiederholungsstrategien“ und „Kontrolle“. Die Gruppe „Elaborationsstrategien“ weist die höchsten „Looking-back“ Zeiten, die Gruppe „Wiederholungsstrategien“ die niedrigsten Zeiten auf. Durch die schwache Effektstärke erscheint die diagnostische Relevanz des Indikators als problematisch. Anscheinend kann bei allen beschriebenen Indikatoren nur die Gruppe Wiederholungsstrategien eindeutig von Kontroll- und Elaborationsgruppe getrennt werden. Hier erscheint eine weitere Untersuchung dieser Indikatoren vielversprechend. Die mangelnde Trennung zwischen Elaborationsstrategien und Kontrollgruppe könnte eine mangelhafte Implementierung der Lernstrategieinstruktion zurückzuführen sein. Hier werden erst weitere Studien mit einer veränderten Implementierung der experimentellen Bedingungen und auch einer größeren Bandbreite der Blickbewegungsindikatoren zu einer abschließenden Bewertung führen können.

## **Förderung des mathematischen Modellierens in der Physikalischen Chemie**

**Ines Komor, Helena van Vorst, Elke Sumfleth**

Universität Duisburg-Essen, Deutschland; [ines.komor@uni-due.de](mailto:ines.komor@uni-due.de)

### Theoretischer Hintergrund

Mathematisches Modellieren spielt beim Bearbeiten von chemischen Aufgaben eine wichtige Rolle. Beschrieben wird der Vorgang des mathematischen Modellierens durch den mathematischen Modellierungskreislauf für das Fach Chemie (Goldhausen, 2015). Es handelt sich dabei um ein aus der Mathematikdidaktik adaptiertes Kreislaufmodell von Blum & Leiß (2005). Beim Durchlaufen des Kreislaufes müssen Lernende zunächst eine reale Situation (1) erfassen und ein mentales Modell dieser Situation (2) entwickeln. Durch Prozesse des Vereinfachens und Strukturierens wird daraus ein chemisches Modell (3) abgeleitet. Aus der Mathematisierung dieses Modells resultiert dann ein mathematisches Modell (4). Mit Hilfe des mathematischen Modells wird durch mathematisches Arbeiten ein mathematisches Ergebnis ermittelt (5). Dieses wird in ein chemisches Ergebnis übersetzt. Abschließend wird nach der Validierung (6) des Ergebnisses ein Zusammenhang zur ursprünglichen Situation hergestellt (7). Nach Goldhausen (2015) haben Lernende aufgrund mangelnden chemischen Grundwissens Schwierigkeiten, ein angemessenes chemisches Modell aufzustellen. Daraus ergeben sich wiederum Probleme beim Schritt der Mathematisierung. Für die Allgemeine Chemie konnte bereits gezeigt werden, dass Studierende Schwierigkeiten beim Mathematisieren (4) haben, das benötigte mathematische Grundwissen jedoch vorhanden ist (Kimpel, eingereicht). Ziel der Vorstudie ist es, zu untersuchen, inwiefern sich die bereits gefundenen Ergebnisse auf die Physikalische Chemie übertragen lassen und welche zusätzlichen Schwierigkeiten sich hier zeigen. Vermutlich bereitet der Schritt der Mathematisierung den Studierenden in der Physikalischen Chemie größere Probleme, da dieser Schritt komplexer ist als in der Allgemeinen Chemie und die Rechenoperationen anspruchsvoller sind.

### Ergebnisse der Vorstudie

Im Rahmen einer Vorstudie wurden daher zunächst die schwierigkeiterzeugenden Schritte im Modellierungskreislauf für Aufgaben aus der Physikalischen Chemie mithilfe von Studierendeninterviews (N = 10) analysiert. Dazu wurden drei Aufgaben mit steigendem Schwierigkeitsgrad entwickelt, die als Interviewleitfaden dienten. Diese Aufgaben wurden dann von Studierenden der Vorlesung zur Physikalischen Chemie mit der Methode „Lautes Denken“ gelöst. Die Aufgabenstellungen waren so konzipiert, dass die einzelnen Schritte des Modellierungskreislaufs detailliert und möglichst isoliert betrachtet werden konnten.

Die Ergebnisse der Vorstudie zeigen, dass schon das Erfassen und Strukturieren des Sachverhalts (Schritte 1 und 2) sowie der spätere Rückbezug des Ergebnisses auf die Aufgabenstellung (Schritt 6 und 7) für Studierende eine Herausforderung darstellen. Dabei ist anzunehmen, dass Studierende lediglich Schwierigkeiten bei der expliziten Verbalisierung der notwendigen Vereinfachungen haben, obwohl sie ihnen wahrscheinlich (unbewusst) bekannt sind. Wie schon in der Allgemeinen Chemie traten auch in der Physikalischen Chemie die größten Schwierigkeiten bei der Entwicklung eines mathematischen Modells (3) auf. Konnten die Studierenden bei der Aufgabenbearbeitung auf ihr Vorwissen aus der Allgemeinen Chemie zurückgreifen, gelang ihnen der Schritt der Mathematisierung gut. Bei der Mathematisierung von Inhalten, welche erstmalig in der Vorlesung zur Physikalischen Chemie behandelt wurden, hatten sie hingegen größere Schwierigkeiten. Zudem fehlten den Studierenden entscheidende mathematische Kompetenzen aus dem Bereich der Integral- und Differentialrechnung. Dies äußerte sich darin, dass die Studierenden die zur Herleitung einer Formel benötigten Umformungen in der Regel nicht durchführen konnten, sobald sie integrieren mussten. Das Einsetzen und Ausrechnen der entsprechenden Werte gelang den Studierenden (4) zwar, sie hatten aber Probleme bei der Umformung von Einheiten, insbesondere bei wenig vertrauten Inhalten aus der Physikalischen Chemie.

### Ziele der geplanten Studie

Ein Ziel ist es nun auf Grundlage dieser Ergebnisse sowohl ein Messinstrument zur Erfassung der mathematischen Modellierungskompetenz als auch ein individualisiertes Förderprogramm für die Physikalische Chemie zu entwickeln. Diese beiden Elemente sollen dann im Rahmen einer Interventionsstudie mit Studierenden der Veranstaltung Physikalische Chemie evaluiert werden. Das Diagnoseinstrument soll die Teilschritte des Modellierungskreislaufs abdecken und dabei ein besonderes Augenmerk auf die gefundenen typischen Schwierigkeiten legen. Mit Hilfe des Diagnoseinstruments sollen individuelle Defizite bei Studierenden identifiziert werden, sodass im Anschluss an die Diagnose jeder Studierende ein individuell angepasstes Förderangebot bearbeiten kann. Bei diesem Förderangebot sollen neben chemischen Fachinhalten auch mathematische Grundlagen vermittelt werden.

## **Stress is in the (h)air - Stressbelastung von Erstsemestristudierenden in Abhängigkeit vom elterlichen Bildungshintergrund**

**Alex Bertrams<sup>1</sup>, Nina Minkley<sup>2</sup>**

<sup>1</sup>Universität Bern, Schweiz; <sup>2</sup>Ruhr-Universität Bochum, Deutschland; [alexander.bertrams@edu.unibe.ch](mailto:alexander.bertrams@edu.unibe.ch), [Nina.Minkley@rub.de](mailto:Nina.Minkley@rub.de)

### Theoretischer Hintergrund

Die Ergebnisse soziologischer Studien verweisen darauf, dass im deutschsprachigen Raum an den verschiedenen Übergängen im Bildungssystem soziale Ungleichheiten bestehen. So ist beispielsweise die Wahrscheinlichkeit, dass eine Person mit Hochschulzugangsberechtigung ein Hochschulstudium aufnimmt, höher, wenn sie aus einem akademischen

(anstatt aus einem nichtakademischen) Elternhaus stammt (Becker, 2012). Dieser Befund kann unter anderem dadurch erklärt werden, dass Personen aus akademischen Elternhäusern eine höhere Motivation aufweisen (und ggf. auch einen höheren Druck verspüren), den sozialen Status der Elterngeneration zu erhalten. Daher tendieren sie dazu, selbst bei relativ niedrigem Leistungsniveau und geringen Erfolgserwartungen ein Studium zu beginnen (Becker, 2012).

Die eben erwähnten Befunde lassen im Mittel eine erhöhte Anpassungserfordernis bei Erstsemesterstudierenden aus akademischen Elternhäusern erwarten. Aktuelle Befunde von Janke, Rudert, Marksteiner und Dickhäuser (2017) implizieren hingegen, dass Studierenden aus nichtakademischen Elternhäusern kurz nach dem Studieneintritt erschwerte Anpassungsbedingungen vorfinden, da diese Studierendengruppe sich selbst als leistungsängstlicher beschrieb und weniger wahrgenommene Zugehörigkeit zur Gruppe der Akademiker angab als Studierende mit akademischem Bildungshintergrund.

Solche erschwerten Anpassungsbedingungen könnten zu erhöhten Stressreaktionen führen, die gemäss dem Transaktionalen Stress-Modell (Lazarus & Folkman, 1984) aus der Wahrnehmung von Bedrohung in Verbindung mit einer gering empfundenen Fähigkeit, diese zu bewältigen, entstehen. Da verschiedene Untersuchungen zeigen konnten, dass die Cortisolkonzentration im Haar ein Marker für die chronische Stressbelastung ist (z.B. Karlén et al., 2011; van Uum et al., 2008, Minkley et al., 2015), kann diese genutzt werden, um die Stressbelastung zu Studiumsbeginn zu erfassen und in Zusammenhang mit dem elterlichen Bildungshintergrund zu bringen.

#### Fragestellung

Vor dem Hintergrund der zuvor beschriebenen Befunde lassen sich gegensätzliche Hypothesen formulieren, die wir in der vorliegenden Pilotstudie getestet haben: Es besteht ein positiver versus ein negativer versus ein V-förmiger Zusammenhang zwischen dem Stress zu Studiumsbeginn und der Anzahl studierter Elternteile.

#### Methode

An der Pilotstudie nahmen 36 Studierende einer Schweizerischen Universität teil, die 2016 ihre Hochschulzugangsberechtigung erworben und im selben Jahr ein Hochschulstudium begonnen hatten. Die Studierenden gaben eine Haarprobe ab, füllten mehrere Selbstberichtsverfahren aus und gaben für beide Eltern getrennt jeweils den akademischen Status an. Die Erhebung fand 1,5 Monate nach Studiumsbeginn statt.

Zur Analyse der Cortisolkonzentration im Haar wurden drei dünne Strähnen am Hinterkopf entnommen (ca. 2 cm unterhalb des Schädelknochens), da dort das Haarwachstum mit ca. 1 cm pro Monat am einheitlichsten ist (Harkey 1993; Pragst und Balikova, 2006). Anschließend wurden aus diesen Strähnen gemäß eines individuellen Schneideprotokolls zwei jeweils 1,5 cm lange Segmente geschnitten. Diese umfassten jeweils eine Phase von 1,5 Monaten und entstammten aus dem Zeitraum zu Studiumsbeginn (Mitte September bis Ende Oktober) bzw. aus der Zeit davor (Mitte Juli bis Ende August). Die Analyse des Haarcortisols erfolgte nach dem Protokoll von Kirschbaum et al. (2009) sowie von Stalder et al. (2012) in den Laboren der Dresden LabService GmbH.

Wir führten zwei Kovarianzanalysen durch, in denen wir jeweils drei Studierendengruppen verglichen: Kein Elternteil (n=19), ein Elternteil (n=10) oder beide Elternteile (n=7) mit Hochschulabschluss. In den Analysen kontrollierten wir für das Geschlecht, da dieses die Cortisolkonzentration beeinflussen kann (Stalder et al., 2017).

#### Ergebnisse

Wir fanden, dass sich die Gruppen nicht in der Cortisolkonzentration zum Zeitraum vor Studiumsbeginn unterschieden,  $p=.44$ . Es fand sich jedoch ein Unterschied in der Cortisolkonzentration zum Zeitraum des Studiumsbeginns,  $p=.007$ ,  $\eta^2=.27$  (die Cortisolkonzentration zum Zeitraum vor dem Studiumsbeginn wurde dabei kontrolliert, um die Veränderung zu untersuchen). Bonferroni-adjustierte paarweise Vergleiche mit Bootstrapping (Field, 2013) erbrachten, dass Studierende mit zwei studierten Elternteilen signifikant höhere Cortisolwerte aufwiesen als die Studierenden ohne bzw. mit einem studierten Elternteil,  $p=.02$  und  $.03$ . Die Cortisolkonzentrationen der Studierenden aus den letztgenannten Gruppen unterschieden sich nicht voneinander,  $p=.67$ . Die Ergebnisse liefern somit einen ersten Hinweis darauf, dass der Bildungsvorteil von Akademikerkindern auf der Kehrseite eine erhöhte Stressbelastung bedeuten kann.

## **Lernentwicklungsgespräche – Effekte auf Selbstkonzept, Attributionsmuster und Motivation**

**Sonja Dollinger, Andreas Hartinger**

Universität Augsburg, Deutschland; [sonja.dollinger@phil.uni-augsburg.de](mailto:sonja.dollinger@phil.uni-augsburg.de)

Im Poster soll eine Studie vorgestellt werden, in der die Tiefenstruktur von Lernentwicklungsgesprächen (LEG) analysiert wird – im Fokus stehen dabei Aspekte der Tiefenstruktur, von denen Effekte auf zentrale Elemente der Lernmotivation angenommen werden können.

Verschiedene Untersuchungen belegen, dass Leistungsbeurteilungen u.a. Selbstkonzept, Attributionsmuster und Motivation beeinflussen (z.B. Hattie 2015; Harks et al. 2014; Mischo & Rheinberg 1995); allerdings nur dann, wenn sie sich auf die Tiefenstrukturen des Lernens beziehen, wie z.B. eine individuelle Bezugsnormorientierung, lernförderliches Feedback etc. Dies wird durch Studien zum Vergleich von Verbalbeurteilungen und Ziffernnoten (z.B. Valtin & Wagner 2002) bestätigt, in denen sich derartige Effekte nicht nachweisen ließen, da es sich bei den Verbalbeurteilungen mehr um eine Übersetzung von Noten, als um eine Beschreibung der Lernentwicklung handelte (vgl. z.B. Schmude 2002), also lediglich um ein anderes Format.

Lernentwicklungsgespräche – Gespräche zwischen Lehrperson und Kind im Beisein der Eltern (z.B. Hardeland 2017) – stellen eine Form der Leistungsbeurteilung dar, von der Effekte auf die genannten Konstrukte zu erwarten sind, da in deren Konzeption wirksame tiefenstrukturelle Elemente Berücksichtigung finden können. Lernentwicklungsgespräche sind von bestimmten Gestaltungsmerkmalen geprägt (vgl. ebd.). Hierbei handelt es sich um Aspekte, die auch im Formativen Assessment von Bedeutung sind, wie der folgende Überblick über die jeweils zentralen Elemente zeigt:

Gestaltungselemente der LEG (z.B. Hardeland 2017):

- Selbsteinschätzung der Kinder im Vorfeld
- Fremdeinschätzung durch die Lehrperson
- Vereinbarung von Lernzielen
- Rückmeldung im Gespräch
- Überprüfung der Zielvereinbarungen

Zentrale Elemente im Formativen Assessment (vgl. Schmidinger et al. 2016; Maier 2015; Black & William 2009):

- Selbstbewertung/Selbstreflexion
- Erfassung von Lernstand sowie Lernprozess und Lernentwicklung unter Verwendung der individuellen und kriterialen Bezugsnorm
- Klarheit über (vereinbarte) Lernziele und Erfolgskriterien sowie Bewertungskriterien
- lernunterstützendes Feedback durch Kommentare und qualitative Aussagen
- Diagnosezyklen, u.a. zur Lernverlaufsdiagnostik oder zur Reflexion der Zielerreichung

Auch die Wirkung von klaren Zielen, Selbstbewertung/Selbstreflexion und wiederholten Diagnosen auf die genannten Konstrukte wurde in verschiedenen Untersuchungen nachgewiesen (z.B. Hattie 2015; Andrade 2010; Maier 2015). Die Wirksamkeit der einzelnen Elemente erklärt auch, warum insgesamt von formativen Verfahren Effekte auf zentrale Aspekte der Lernmotivation ausgehen (z.B. Hattie 2015). Da die Effektivität formativer Verfahren jedoch auch von der Qualität der Umsetzung der einzelnen Elemente abhängt (vgl. Maier 2015), also von einer tatsächlichen Umsetzung auf tiefenstruktureller Ebene, wird dies ebenso für die Lernentwicklungsgespräche erwartet.

Die Effekte von Lernentwicklungsgesprächen auf Selbstkonzept, Attributionsmuster und Motivation sind bislang noch nicht untersucht worden – ebenso, inwieweit die benannten Elemente tatsächlich umgesetzt werden.

Somit ergeben sich konkret die folgenden beiden Fragestellungen:

1. Welche Elemente Formativen Assessments werden in welcher Qualität in Lernentwicklungsgesprächen umgesetzt?
2. Sind Stärke und Stabilität der Effekte auf Selbstkonzept, Attributionsmuster und Motivation abhängig von der qualitativen Umsetzung bestimmter formativer Elemente?

Im Poster wird über die geplante Anlage zur ersten Fragestellung berichtet werden.

Die Umsetzung und die Qualität der zentralen formativen Elemente der LEG sollen über einen Mixed-Methods-Ansatz erfasst werden. Dazu werden die Lehrpersonen zu ihrem LEG-Verfahren befragt werden, die eingesetzten Bögen zur Selbst- und Fremdeinschätzung sowie die vereinbarten Ziele bezüglich ihrer qualitativen Umsetzung analysiert werden und die Kinder zu ihrer Wahrnehmung und den im Gespräch erhaltenen Informationen hinsichtlich Lernstand, Lernentwicklung sowie lernunterstützenden Rückmeldungen befragt werden. Zur Absicherung werden zudem Gespräche analysiert werden.

Vorgestellt werden dann auch die Qualitätskriterien, anhand derer die zum Einsatz kommenden Selbst- und Fremdeinschätzungsbögen, die Gesprächssituation sowie die Vereinbarung der Ziele analysiert werden.

Zum Zeitpunkt der Tagung werden erste Ergebnisse einer Pilotierungsstudie vorliegen, die ebenfalls präsentiert werden.

## **Der Einfluss von Zielorientierung und Kooperationsbereitschaft auf Lernverhalten bei adaptivem und non-adaptivem Lernen: eine Längsschnittstudie**

**Stefan Kulakow, Sabine Schweder, Tino Wulff**

Ernst-Moritz-Armdt-Universität Greifswald, Deutschland; [stefan.kulakow@uni-greifswald.de](mailto:stefan.kulakow@uni-greifswald.de)

Theoretischer Hintergrund

Leitbilder schulischer Lernprozesse sind Individualisierung und Kompetenzorientierung (KMK, 2004, 2011). Zentrum des in dieser Studie untersuchten Konzepts ist das Kompetenzraster auf deren Grundlage der Stand von Vorwissen bestimmt und damit eine individuelle Lernausgangsposition (individuelle Bezugsnorm) der Schüler/-innen ermittelt werden kann (Krielle, 2016). Von dieser Bezugsnorm ausgehend, ergeben sich für Schüler/-innen beim Lernen folgende Adaptionsmöglichkeiten: (1) Für die Schüler/-innen wird es möglich zwischen den mit dem Kompetenzraster korrespondierenden Lernaufgaben auszuwählen, wobei Lernbedarfe und Präferenzen (Interessen) die Auswahl bestimmen. Die Lernaufgaben unterscheiden sich nach Verarbeitungstiefe und einem damit verbundenen Aufwand an elaborativen und transformativen Denkstrategien. Mit zunehmender Kompetenzentwicklung steigt das Anforderungsniveau der Aufgaben (Günther et al., 2016). (2) Je nach Wunsch können die Lernenden mit anderen Schüler/-innen zusammenarbeiten, (3) es ist möglich die zur Verfügung stehende Zeit eigenständig zu strukturieren, (4) und bei Bedarf Hilfen durch begleitende Lehrer/-innen in Anspruch zu nehmen. Damit folgt das in dieser Studie untersuchte Konzept zentralen Merkmalen adaptiver Lernumgebungen (Benavides, Istance, & Dumont, 2008; Wang & Lindvall, 1984).

Mit dem Begriff der Zielorientierung werden Vorstellungen zur motivationalen Ausrichtung in Lern- und Leistungszusammenhängen beschrieben (Elliot & Hulleman, 2017). Die Studie untersucht drei Formen der Zielorientierung: (1) Personen mit einer ausgeprägten Lernzielorientierung suchen Handlungs- und Lernsituationen im Rahmen derer Kompetenzsteigerung und damit das Erweitern von Handlungsmöglichkeiten möglich wird (Ames, 1992; Nicholls, Patashnick, & Nolen, 1985). (2) Personen mit Leistungszielorientierung suchen Kontexte in denen eigene Kompetenzen an Kompetenzen anderer gemessen und unter Beweis gestellt werden (Elliot & Hulleman, 2017). (3) Personen mit einer Annäherungs-Leistungszielorientierung bemühen sich in Handlungs- und Lernkontexten Inkompetenz zu verbergen (Elliot & Harackiewicz, 1996). Vor allem Lernzielorientierung, hat positive Folgen auf tiefere Verarbeitungsprozesse und Anstrengungsbereitschaft (Hulleman & Barron, 2016; Spinath, 2012).

Kooperationsbereitschaft beschreibt das Ausmaß mit dem Personen in Handlungs- und Lernkontexten Kontakt zu anderen aufnehmen, um das Handeln und Lernen gemeinsam zu gestalten und zu bewältigen. Die damit einhergehenden Austausch- und Abstimmungsprozesse intensivieren und fördern eine tiefere Verarbeitung von Wissen und steigern Lernengagement und Qualität (Sparfeldt, Rost, Schleebusch, & Heise, 2012).

Das in dieser Studie untersuchte adaptive Konzept (AL) ist an etlichen Schulen etabliert und bezüglich seiner Wirkung nicht ausreichend untersucht. Somit ist Ziel der Studie den Einfluss von Lernzielorientierung, Leistungszielorientierung und Annäherungs-Leistungszielorientierung in einem ersten Messpunkt auf Kooperationsbereitschaft, Elaboration, Transformation und Anstrengungsbereitschaft mit Berücksichtigung der mediierenden Rolle von Kooperationsbereitschaft zu untersuchen und dann mit lehrerzentrierten und damit nicht adaptivem Unterricht (N-AL) zu vergleichen.

Hypothesen

Dafür wird folgende Hypothesen nachgegangen: H1: Es wird bei AL und N-AL von einem unterschiedlichen Zusammenspiel der Variablen ausgegangen. H2: Kooperationsbereitschaft mediiert die bestehenden Zusammenhänge in AL und N-AL unterschiedlich.

Methode

Die empirische Grundlage der Längsschnittstudie bilden Fragebogendaten von 1153 Schüler/-innen (AL/T1: MAlter1 = 13.98, SD = 1.37; AL/T2: N = 772, MAlter2 = 14.27, SD = 1.25) an sechs Schulen, wobei drei Schulen den Hauptteil der schulischen Lernzeit auf AL, umgestellt haben. Die quantitativen Analysen wurden mithilfe von Mplus 7.0 durchgeführt. Zum Nachweis von Messinvarianz wurde dem Ansatz von Brown (2015) gefolgt, zudem unterschiedlich restringierte Mehrgruppenstrukturgleichungsmodelle konzipiert.

#### Ergebnisse

H1: Der Vergleich der Multigruppenstrukturgleichungsmodelle belegt ein unterschiedliches Zusammenspiel der Variablen bei AL und N-AL. Lernzielorientierung beeinflusst bei AL alle Zielvariablen, was ein Beleg dafür ist, dass mit AL Schüler/-innen mit Lernzielorientierung gleichzeitig positives Lernverhalten nachweisen. Im Gegensatz dazu beeinflusst bei N-AL Lernzielorientierung lediglich höheres Elaborieren, wegen nicht ausreichender Kooperationsmöglichkeiten weniger Transformationsverhalten und Anstrengungsbereitschaft.

H2: Kooperationsbereitschaft mediiert in AL und N-AL auf unterschiedliche Weise die Beziehungen der Ausgangsvariablen (Lernzielorientierung, Leistungszielorientierung, Annäherungs-Leistungszielorientierung) und Elaboration, Transformation und Anstrengungsbereitschaft. Die Ergebnisse für AL können so gedeutet werden, dass die vorhanden Kooperationsmöglichkeiten Transformationsprozesse fördern. Durch die aktive Auseinandersetzung wird Abstraktion und Transfer gefördert. Pädagogische Implikationen werden weiterführend diskutiert.

### **Messung kognitiver Belastung in Laborpraktika**

**Katrin Schüßler, Maik Walpuski**

Universität Duisburg-Essen, Deutschland; [katrin.schuessler@uni-due.de](mailto:katrin.schuessler@uni-due.de)

#### Theoretischer Hintergrund

Laborpraktika sind traditionell ein charakteristisches Element naturwissenschaftlicher Studiengänge. Als eine herausragende Besonderheit wird dabei die Vermittlung zwischen der Domäne der Objekte und Beobachtungen (praktischer Umgang mit Geräten und Chemikalien) und der Domäne der Vorstellungen (konzeptuelles Verständnis von Fachinhalten) angesehen (Abrahams & Millar, 2008; Tiberghien, 2000). Daraus resultieren zwei Ebenen, auf denen Lernende beim Experimentieren effektiv sein können: Die Handlungsebene (Experimente werden wie vorgesehen durchgeführt) und die Verständnisebene (Verstehen der durch das Experiment beobachtbaren Sachverhalte). Um erfolgreich Experimentieren zu können, müssen Lernende folglich auf beiden Ebenen effektiv sein. Nur wenn beides gelingt, kann die angestrebte Verknüpfung zwischen der Domäne der Objekte und Beobachtungen und der Domäne der Vorstellungen erfolgreich sein.

Aus Sicht der Cognitive Load Theory (z. B. Paas & Sweller, 2014) bergen Laborpraktika ein hohes Risiko kognitiver Überlastung von Lernenden, weil zum einen in zwei Inhaltsbereichen gleichzeitig gelernt werden soll (1) Domäne der Objekte, (2) Domäne der Vorstellungen, sodass ein hoher intrinsic cognitive load zu erwarten ist. Zum anderen besteht durch die Laborumgebung eine vergleichsweise große Gefahr, dass Lernende durch unklare Versuchsanleitungen oder die Suche von Geräten und Chemikalien im Labor einen hohen extraneous cognitive load erfahren. Der Lernerfolg von Studierenden in Laborpraktika hängt daher vermutlich davon ab, wie stark ihre kognitive Kapazität durch die beiden Lernbereiche und durch irrelevante Verarbeitungsprozesse beansprucht wird und ob darüber hinaus noch Kapazität für tiefere Verarbeitungsprozesse (germane cognitive load) verfügbar ist (und ob Lernende ausreichend motiviert hierfür sind).

Neuere Testinstrumente zur Erfassung der kognitiven Belastung ermöglichen es intrinsic load und extraneous load getrennt voneinander zu erfassen (Leppink, Paas, van der Vleuten, van Gog, & van Merriënboer, 2013). Das Instrument erfasst darüber hinaus einen dritten Faktor, für den allerdings nicht abschließend geklärt werden konnte, ob er germane load erfasst (Leppink, Paas, van Gog, van der Vleuten, & van Merriënboer, 2014). Ziel der hier vorgestellten Studie ist es zu prüfen, inwiefern das Testinstrument, das bisher nur im Rahmen klassischer Lernumgebungen (paper pencil, multimedia learning) eingesetzt wurde, auch in Laborpraktika erfolgreich eingesetzt werden kann, um zwischen intrinsic load und extraneous load zu differenzieren. Darüber hinaus soll überprüft werden, inwiefern es möglich ist im Rahmen von Laborpraktika zusätzlich die kognitive Belastung, die durch die beiden Lernbereiche (Domäne der Objekte, Domäne der Vorstellungen) unterschieden werden kann. Wenn dies gelingt, kann mithilfe des Testinstruments überprüft werden, ob die Effektivität von Laborpraktika durch Interventionsmaßnahmen, die auf einen der beiden Lernbereiche abzielen, verbessert werden kann.

#### Fragestellung

1. Inwiefern können im Rahmen eines Laborpraktikums intrinsic und extraneous load getrennt voneinander erfasst werden?
2. Inwiefern kann im Rahmen eines Laborpraktikums die kognitive Belastung, die durch die Domäne der Objekte verursacht wird, von der kognitiven Belastung, die durch die Domäne der Vorstellungen verursacht wird, unterschieden werden?

#### Methode

Im Rahmen eines Praktikums für Chemiestudierende (Lehramt) des ersten Semesters werden Studierenden nach der Durchführung jedes Versuchs Fragen zur kognitiven Belastung durch die beiden Lernbereiche und unnötige Verarbeitungsprozesse, die durch die Lernumgebung verursacht werden, gestellt. Hierzu wurden das Instrument von Leppink und Kollegen (Leppink et al., 2013; Leppink et al., 2014; Leppink, van Gog, Paas, & Sweller, 2015) ins Deutsche übersetzt und an die spezifische Lernsituation angepasst. Auf Basis der gewonnenen Daten soll geprüft werden, inwiefern die verwendeten Items geeignet sind, um intrinsic load und extraneous load im Rahmen eines Laborpraktikums getrennt voneinander zu erfassen und ob es darüber hinaus gelingt, die durch die beiden Inhaltsbereiche (Domäne der Objekte, Domäne der Vorstellungen) verursachte kognitive Belastung getrennt voneinander zu erfassen.

#### Ergebnisse

Die Durchführung der Studie ist für den Beginn des Wintersemesters 2017/18 geplant. Erste Ergebnisse (von N = 80 Studierenden) werden bis zum Beginn der Tagung vorliegen.

### **„Ich arbeite danach und ich hoffe, dass alle anderen das genauso tun.“ Leistungsbeurteilung in zentralen Abschlussprüfungen aus der Sicht von Lehrkräften**

**Alexandra Kranz, Svenja Mareike Kühn**

Universität Koblenz-Landau, Deutschland; [akranz@uni-koblenz.de](mailto:akranz@uni-koblenz.de)

#### Theoretischer Hintergrund

Zentrale Prüfungen haben das Ziel, einheitliche Leistungsanforderungen zu sichern und Schulabschlüsse vergleichbarer zu machen (vgl. z.B. Klein et al., 2014). Landesweit einheitliche Korrektur- und Bewertungsvorgaben sollen in diesem Kontext eine möglichst objektive, transparente und vergleichbare Leistungsbeurteilung sicherstellen. Die wenigen bislang dazu vorliegenden Studien weisen allerdings darauf hin, dass Korrektur- und Bewertungsvorgaben sehr unterschiedlich gestaltet sind (vgl. Kühn, in Begutachtung) und deren Qualität und Nützlichkeit von den beurteilenden Lehrkräften unterschiedlich eingeschätzt wird (vgl. Appius & Holmeier, 2012). Für die Anwendung dieser Vorgaben bei der Beurteilung von Schülerleistungen im Kontext zentraler Prüfungen dürfte insofern deren Qualität (z.B. eindeutige und inhaltlich ausdifferenzierte Beurteilungskriterien, klare Vorgaben für die Notenvergabe etc.) entscheidend sein (vgl. Köster, 2010). Im Rekurs auf schulische Governance-Ansätze (z.B. Altrichter, Brüsemeister & Wissinger, 2007; Altrichter & Maag Merki, 2016) ist zudem davon auszugehen, dass die tatsächliche Anwendung dieser Vorgaben nicht nur durch deren konkrete Ausgestaltung bedingt wird, sondern durch vielfältige „Wahrnehmungs-, Verarbeitungs- und Übersetzungsleistungen“ (Altrichter & Maag Merki, 2016, S. 5) der beurteilenden Lehrkräfte sowie organisational-situative Handlungsbedingungen (z.B. einzelschulische „Benotungspolitik“, vgl. Terhart, 2014, S. 890ff.) multifaktoriell beeinflusst wird. Die Fragen, inwieweit Lehrkräfte die Korrektur- und Bewertungsvorgaben nutzen und welche Faktoren sich diesbezüglich als förderlich bzw. hinderlich erweisen, sind bislang eine Leerstelle in der Forschung.

#### Fragestellung

Im Posterbeitrag werden erste Ergebnisse einer explorativen Studie dargestellt, die zentrale Prüfungen zum Erwerb des Mittleren Schulabschlusses nach Klasse 10 fokussiert. Vor dem Hintergrund der skizzierten Forschungsdesiderata war folgende Forschungsfrage handlungsleitend: Inwieweit nutzen Lehrkräfte die Korrektur- und Bewertungsvorgaben und welche Faktoren beeinflussen die Nutzung?

#### Methode

In dieser Explorationsstudie wurden leitfadengestützte Interviews mit Lehrkräften der Fächer Deutsch und Mathematik durchgeführt. Befragt wurden pro Fach drei Lehrkräfte in drei Ländern (N=18); Kriterien für die Auswahl der Befragten innerhalb eines Landes waren primär vorhandene Erfahrungen als beurteilende Lehrkraft in zentralen Abschlussprüfungen. Die Länder wurden auf der Basis einer empirisch ermittelten Typologie von Korrektur- und Bewertungsvorgaben aller Bundesländer mit zentralen Prüfungen ausgewählt (vgl. Kühn, in Begutachtung), die in beiden Fächern jeweils drei Typen umfasst. Die Typen sind durch unterschiedliche Gestaltungsmerkmale der Korrektur- und Bewertungsvorgaben charakterisiert – die Vorgaben der drei ausgewählten Länder unterscheiden sich erheblich (z.B. konkrete Musterlösungen mit eindeutiger Punktezuweisung und Punkte-Noten-Schlüssel vs. abstrakte Kriterien ohne Hinweise zur Notenvergabe). Die Interviewdaten wurden inhaltsanalytisch ausgewertet, Grundlage dafür war ein deduktiv-induktiv entwickeltes Kategoriensystem (vgl. Kuckartz, 2012). Die Qualität des Kategoriensystems wurde über einen mehrstufigen Prozess der Erprobung, Diskussion, Ausdifferenzierung und Reduktion sichergestellt.

#### Ergebnisse

Insgesamt zeigt sich unter den befragten Lehrkräften eine heterogene Nutzungspraxis, die von einer sehr genauen Verwendung der Korrektur- und Bewertungsvorgaben über eine grobe Orientierung an diesen Vorgaben bis hin zu deren bewussten Nicht-Nutzung reicht. Die Nützlichkeit der landesweit einheitlichen Korrektur- und Bewertungsvorgaben wird dabei sehr unterschiedlich eingeschätzt, wobei subjektiv wahrgenommene Schwächen der Vorgaben (z.B. nicht hinreichend konkrete Kriterien) in unterschiedlichem Maße durch Erfahrungswerte mit zentralen Prüfungen und selbst zugeschriebene Beurteilungskompetenz kompensiert werden. Es finden sich daher Hinweise, dass der Umgang mit den Korrektur- und Bewertungsvorgaben gerade für jüngere Lehrkräfte eine Herausforderung darstellt. Auffällig ist weiterhin, dass die je nach Bundesland unterschiedlich großen Beurteilungsfreiräume in der Regel genutzt werden, um eigene Bewertungs- und Leistungsansprüche in die Korrektur und Beurteilung einfließen zu lassen. Neben sachbezogenen Kriterien gehen dabei auch nicht-prüfungsleistungsbezogene Aspekte (z.B. bisheriges Leistungsniveau des/der Schülers/-in) in die Leistungsbeurteilung ein. Neben diesen individuellen Faktoren haben sich zudem sowohl schulexterne Unterstützungsangebote (z.B. gemeinsame Korrekturbesprechungen) als auch (formal-organisierte und informelle) schulinterne kollegiale Absprachen als relevante Einflussfaktoren herausgestellt, die in unterschiedlichem Maße einen Konsens in der Beurteilungspraxis herbeiführen, der nicht zwangsläufig der intendierten Zielsetzung zentraler Prüfungen entspricht.

### Zum Zusammenhang von epistemischen Überzeugungen und kognitiven Anforderungen von Multiple Documents Literacy

**Nina Mahlow<sup>1</sup>, Cornelia Schoor<sup>2</sup>, Carolin Hahnel<sup>3,4</sup>, Ulf Kröhne<sup>3</sup>, Cordula Artelt<sup>1,2</sup>, Frank Goldhammer<sup>3,4</sup>**

<sup>1</sup>Leibniz-Institut für Bildungsverläufe, Deutschland; <sup>2</sup>Otto-Friedrich-Universität Bamberg, Deutschland; <sup>3</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt am Main; <sup>4</sup>Zentrum für internationale Vergleichsstudien (ZIB), Frankfurt am Main; [nina.mahlow@lifbi.de](mailto:nina.mahlow@lifbi.de)

Multiple Documents Literacy (MDL) – die Fähigkeit, relevante Informationen aus mehreren Quellen zu integrieren (vgl. Bråten & Strømsø, 2010) – stellt heutzutage eine zentrale Schlüsselkompetenz für Studierende dar, da sie im Laufe ihres Studiums immer wieder mit redundanten, komplementären oder widersprüchlichen Informationen zu einem bestimmten Inhaltsbereich umgehen müssen (Bråten et al., 2014). Aktuelle Modellvorstellungen zu MDL postulieren, dass zur kognitiven Repräsentation multipler Dokumente ein sogenanntes Documents Model aufgebaut wird (Documents Model Framework; Britt & Rouet, 2012). Dieses umfasst eine integrierte Darstellung von Inhalten ("integriertes Situationsmodell") und Metainformationen einzelner Dokumente ("Intertext-Modell"). Britt und Kollegen (1999) konnten ein solches Documents Model für Personen mit hoher MDL empirisch belegen.

Der vorliegende Beitrag untersucht, ob das Documents Model, das Studierende beim Lesen multipler Dokumente erarbeiten, durch ihre epistemischen Überzeugungen (EÜ) beeinflusst wird. EÜ bezeichnen subjektive Theorien über das Wissen und den Wissenserwerb (Bråten, 2010). Entsprechend aktueller Annahmen (Hofer & Pintrich, 1997; Ferguson et al., 2013; Greene et al., 2008) können folgende EÜ-Dimensionen unterschieden werden: simplicity of knowledge (Komplexität), certainty of knowledge (Sicherheit), personal justification (persönliche Rechtfertigung), justification by authority (Glaube an Experten) und justification by multiple sources (Überprüfung des Wahrheitsgehaltes durch andere Quellen). Bråten und Kollegen (2011) nehmen hierbei an, dass die unterschiedlichen EÜ-Dimensionen differentielle Effekte auf MDL entsprechend der zugrundeliegenden Aufgabenanforderungen haben können. Demnach seien die Dimensionen simplicity und justification mit dem integrierten Situationsmodell und die certainty-Dimension mit dem Intertext Model verknüpft. Die empirische Prüfung dieser Annahme stand bislang aus.

Entsprechend lassen sich folgende Hypothesen ableiten:

H1. Je stärker die EÜ einer Person ausgeprägt sind, desto höher ist ihre MDL.

H2. Sophistiziertere simplicity-Ansichten erhöhen die Wahrscheinlichkeit MDL-Aufgaben zu lösen, die eine Integration von Inhalten über multiple Dokumente erfordern („integriertes Situationsmodell“).

H3. Sophistiziertere certainty-Ansichten erhöhen die Wahrscheinlichkeit MDL-Aufgaben zu lösen, die einen Vergleich von Quellen erfordern („Intertext Model“).

H4. Sophistiziertere Ansichten auf den drei justification-Skalen erhöhen die Wahrscheinlichkeit MDL-Aufgaben zu lösen, die eine Integration von Inhalten über multiple Dokumente erfordern („integriertes Situationsmodell“).

Zur Überprüfung dieser Hypothesen wurden die Daten von 310 Studierenden (79.4% weiblich) der Geistes- und Sozialwissenschaften im ersten Bachelor- oder Master-Semester zweier deutscher Universitäten herangezogen. Die Studierenden bearbeiteten einen Fragebogen zur Erfassung ihrer EÜ mit 26 Items auf den oben genannten 5 Dimensionen (5-stufige Likert-Skala; adaptiert nach Bråten & Strømsø, 2010; Ferguson et al., 2013) sowie einen Test zur Erfassung ihrer MDL (67 Items; WLE Reliabilität = .69; Schoor et al., 2016). Entsprechend eines Multi-Matrix-Designs erhielt jeder Proband 3 von 6 MDL-Units (d.h. 2-3 Texte mit bis zu 16 Items). Die Items ließen sich den Anforderungsbereichen (1) Vergleich von Informationen (2) Verknüpfung von Informationen („integriertes Situationsmodell“), (3) Einschätzung und Vergleich von Quellen („Intertext Model“) und (4) Vergleich von Quelle-Inhalt-Kombinationen zuordnen.

Aufgrund mangelnder interner Konsistenz musste u.a. die simplicity-Skala aus den Analysen ausgeschlossen werden, wodurch die H2 nicht geprüft werden konnte. Eine konfirmatorische Faktorenanalyse der verbleibenden EÜ-Items zeigte einen zufriedenstellenden Fit ( $\chi^2 = 221.52$ ,  $p < .001$ ; RMSEA = .05; CFI = .94). Eine anschließende SEM-Analyse erbrachte keine bedeutsamen Zusammenhänge zwischen den EÜ-Dimensionen und der MDL (H1). Anhand von generalisierten linearen gemischten Modellen (GLMMs) wurden in Referenz zum ersten Anforderungsbereich (Vergleich von Informationen) die Lösungswahrscheinlichkeiten einer Aufgabe überprüft. Im Widerspruch zur H3 erhöhte die certainty-Ausprägung nicht die Lösungswahrscheinlichkeit einer MDL-Aufgabe zum Intertext-Modell, stand aber mit Aufgaben zum integrierten Situationsmodell in Verbindung ( $b = 0.33$ ,  $p < .01$ ). Entsprechend der H4 erhöhte sich die Lösungswahrscheinlichkeit von Aufgaben zum integrierten Situationsmodell mit der Ausprägung in justification by multiple sources ( $b = 0.23$ ,  $p < .01$ ) und justification by authority ( $b = 0.22$ ,  $p < .05$ ). Es fand sich jedoch keine Interaktion mit der personal justification-Skala. Der Beitrag wird mit einer kritischen Bewertung der überwiegend erwartungswidrigen Befunde schließen.

## Was ermöglicht angehenden Physiklehrkräften erfolgreiches Problemlösen?

**Martina Brandenburger<sup>1</sup>, Silke Mikelskis-Seifert<sup>1</sup>, Peter Labudde<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Freiburg; <sup>2</sup>Pädagogische Hochschule Hochschule Nordwestschweiz;  
[martina.brandenburger@ph-freiburg.de](mailto:martina.brandenburger@ph-freiburg.de)

### Theoretischer Hintergrund

Problemlösen ist eine wesentliche Voraussetzung für das Handeln in allen Bereichen des Lebens: In einfachen Alltagssituationen, bei (natur-) wissenschaftlichen Fragestellungen oder bei komplexeren gesellschaftlich relevanten Problemen spielt das Lösen von Problemen eine Rolle. Im Rahmen der Lehrprofessionalisierungsforschung ist das Problemlösen innerhalb der Förderung der fachbezogenen Kompetenzen von Lehrkräften zu verorten (vgl. Riese, 2009 in Anlehnung an Baumert / Kunter, 2006 und Blömeke et al., 2008). Von angehenden Lehrkräften ist zu erwarten, dass sie selbst erfolgreich Probleme lösen, um dies den Schülern vermitteln zu können. Die täglichen Erfahrungen des Lehrbetriebs an Hochschulen machen jedoch deutlich, dass Teile der Lehramtsstudierenden Schwierigkeiten haben, Probleme erfolgreich zu bearbeiten.

„Probleme“ sind alle Anforderungen, die das Analysieren und Schlussfolgern auf ein Ziel (oder eine „Lösung“) hin benötigen und Verständnis des bearbeiteten Themenbereichs voraussetzen (vgl. Smith, 1991). „Problemlösen“ als Prozess wird in der vorliegenden Arbeit mit Hilfe des domänenspezifischen Modells des wissenszentrierten Problemlösens (nach Friege, 2001) beschrieben

Aus der Forschung zum Problemlösen ist vor allem die Bedeutung des domänenspezifischen Fachwissens bekannt (vgl. z.B. Smith, 1991; Friege, 2001). Aber auch kognitiv-emotionale Aspekte, wie z.B. das Selbstkonzept beeinflussen, wie gut Probleme gelöst werden können (vgl. z.B. Jonassen, 2000; Laukenmann et al., 2000). Von den genannten Einflussfaktoren ist aus der Perspektive der Hochschullehre vor allem das Fachwissen an bestimmte Veranstaltungen geknüpft. Die übrigen Einflussfaktoren sind, auch aus dem Grund, dass das Problemlösen einen übergeordneten Charakter besitzt (vgl. z.B. auch Klieme et al., 2001) nicht an eine konkrete Veranstaltung gebunden. Somit wird die Notwendigkeit einer (horizontal) kohärenten Lehre deutlich, die beispielsweise das Selbstkonzept oder die Erfahrung der Studierenden über unterschiedliche Lehrveranstaltungen hinweg fördert.

### Fragestellung

Die Diskrepanz zwischen der Wichtigkeit des erfolgreichen Problemlösens und den Schwierigkeiten, die Studierende haben, ist der Ausgangspunkt für die zentrale: Was unterscheidet „gute“ von „schlechten“ Problemlösern und wie können angehende Lehrkräfte angemessen und lehrveranstaltungsübergreifend gefördert werden?

### Methode

Mit Hilfe einer empirischen Untersuchung wird dazu beigetragen, bekannte Erkenntnisse aus der traditionsreichen Forschung zum Problemlösen zu bestätigen, zu quantifizieren und zueinander in Beziehung zu setzen. Es werden sowohl quantitative als auch qualitative Eigenschaften „guter“ und „schlechter“ Problemlöser herausgearbeitet, die als Grundlage zur gezielten Förderung innerhalb der Lehramtsausbildung herangezogen werden können. Zentrales Instrument der Untersuchung ist ein speziell entwickelter Problemlösetest, der auf dem oben angerissenen Modell des wissenszentrierten Problemlösens basiert und die verschiedenen Phasen des Problemlösens getrennt voneinander erhebt. Des Weiteren wurden erobert: Angaben zur Person (z.B. Abiturnote, Erfahrung), Likert-Skalen zum Selbstkonzept und Interesse, Rasch-homogene Fachwissenstests zur Mathematik und zur Mechanik. Die Testzeit beträgt 90 Minuten. Die Instrumente und Ergebnisse sind in Brandenburger (2016) vollständig veröffentlicht.

### Ergebnisse

Zum Vergleich von weniger erfolgreichen Problemlösern mit erfolgreicheren (Experten-Novizen-Vergleich) wurde aus der Rasch-Modellierung des Problemlösetests ein Kompetenzmodell entwickelt, das drei Stufen der Problemlösefähigkeit identifiziert. Die Stufen unterscheiden sich in quantitativer („Probleme besser lösen können“) und in qualitativer Hinsicht („andere Arten von Problemen lösen können“) voneinander. Die Unterschiede zwischen guten und schlechten Problemlösern



stehen im Einklang mit den Erwartungen aus der Theorie. Es wird vor allem die Bedeutung des vorhandenen Fachwissens in Mechanik deutlich.

Um indirekte Effekte aufzudecken, wurden Mediationsanalysen durchgeführt. Von den genannten Prädiktoren behält lediglich das einen direkten Einfluss. Ausgehend davon konnte durch hierarchische lineare Regression quantitativ herausgearbeitet werden, inwiefern neben dem Fachwissen das Selbstkonzept als unabhängige Variable den Erfolg beim Problemlösen bedeutsam vorhersagt ( $R^2 = .25$ ;  $F(1, 125) = 20.2$ ;  $p = .000$ , mittlerer Effekt).

Die Ergebnisse zur Unterscheidung von „guten“ und „schlechten“ Problemlösern unterstreichen die Bedeutung des Fachwissens für erfolgreiches Problemlösen. Durch Regression kann das Selbstkonzept der Studierenden als unabhängiger Faktor identifiziert werden. Aus der Perspektive einer kohärenten Lehre bietet sich entsprechend die Förderung des Selbstkonzeptes über mehrere Lehrveranstaltungen hinweg an, um Studierende in der Ausbildung zu unterstützen.

## Beitrag wurde zurückgezogen!

### Faktoren multiplen Sprachen- und interkulturellen Lernens – Eine quasi-experimentelle Studie zur Mehrsprachigkeitsorientierung im Französischunterricht

**Kerstin Göbel<sup>1</sup>, Lars Schmelter<sup>2</sup>, Julie Buret<sup>2</sup>, Hannah Yola Tziotzios<sup>1</sup>**

<sup>1</sup>Universität Duisburg-Essen, Deutschland; <sup>2</sup>Bergische Universität Wuppertal, Deutschland; [kerstin.goebel@uni-due.de](mailto:kerstin.goebel@uni-due.de), [lars.schmelter@uni-wuppertal.de](mailto:lars.schmelter@uni-wuppertal.de), [buret@uni-wuppertal.de](mailto:buret@uni-wuppertal.de), [yola.tziotzios@uni-due.de](mailto:yola.tziotzios@uni-due.de)

### Entwicklung und Validierung einer videobasierten Simulationsumgebung zur Messung von Kompetenzen der Unterrichtsdiagnose angehender Biologielehrkräfte

**Julia Schaubberger, Christian Förtsch, Birgit J. Neuhaus**

Ludwig-Maximilians-Universität München, Deutschland; [julia.schaubberger@bio.lmu.de](mailto:julia.schaubberger@bio.lmu.de)

#### Theoretischer Hintergrund

Diagnostische Kompetenzen als Teilaspekt der professionellen Kompetenz von Lehrkräften stehen in den letzten Jahren verstärkt im Fokus der empirischen Bildungsforschung. Dabei gelten kognitive, kontextspezifische Leistungsdispositionen und somit die professionelle Wissensbasis eines Lehrers als Grundlage von Diagnosekompetenz (Koeppen et al., 2008). Diese drückt sich im unterrichtlichen Handeln im Rahmen des Diagnoseprozesses durch konkrete epistemisch-diagnostische Aktivitäten, wie Problemidentifikation, Generieren von Evidenzen, Evidenzevaluation und Ziehen von Schlussfolgerungen, aus (Fischer et al., 2014b). Häufig wird die diagnostische Kompetenz auch mit der Fähigkeit, genaue diagnostische Urteile über den Diagnosegegenstand abgeben zu können, gleichgesetzt (Schrader, 2011). In Simulationen können Diagnosesituationen mittels Videovignetten dargestellt werden. Innerhalb dieser Lernumgebungen wird es Probanden ermöglicht durch das Auswählen von Handlungsalternativen in die Situation einzugreifen (Fischer et al., 2014a). Somit dienen videobasierte Simulationsumgebungen als geeignetes Instrument zur Förderung und Messung von Diagnosekompetenzen angehender Biologielehrkräfte. Diagnostiziert wird hierbei die Qualität der unterrichtlichen Prozesse, so dass der Fokus dieses Projektes auf der Unterrichtsdiagnose liegt. Untersuchungen zu allgemeinen und biologiespezifischen Unterrichtsqualitätsmerkmalen (Wüsten, 2010) sowie Ergebnisse aus quantitativen Videostudien, wie ProWiN und LerNT (Förtsch et al., 2017; von Kotzebue et al., 2015; Werner et al., 2017) bilden dabei theoretische Grundlagen für Diagnosesituationen im Biologieunterricht.

Im Rahmen dieser Teilstudie sollen videobasierte Simulationsumgebungen zur Diagnose von Unterrichtsqualität im Biologieunterricht erstellt und validiert werden. Bei der Validierung sollen Teilaspekte von Presence, wie die wahrgenommene Authentizität der Videovignetten, sowie die kognitive Involviertheit innerhalb der Simulation überprüft werden (vgl. Schubert et al., 2001), da diese ausschlaggebend für den Transfer von gezeigtem Verhalten in der Simulation auf reale Diagnosesituationen sind (Frank, 2015).

#### Fragestellungen

1. Werden die konstruierten diagnostisch-relevanten Situationen innerhalb der Videosequenzen von Experten und Novizen als solche identifiziert?
2. Werden die Videosequenzen innerhalb der Simulationen von Experten und Novizen als authentisch eingeschätzt und ermöglicht die Simulation ein Gefühl von Involviertheit?
3. Setzen Experten und Novizen unterschiedliche Strategien bei der Umsetzung der epistemisch-diagnostischen Aktivitäten ein?
4. Unterscheiden sich Experten und Novizen hinsichtlich des Diagnoseergebnisses?

#### Methode

Als Teil der DFG-Forschergruppe Cosima zur Analyse und Förderung von Diagnosekompetenzen im schulbezogenen und medizinischen Kontext fokussiert dieses Teilprojekt auf eine interaktionsbasierte Diagnose von Unterrichtsqualität mittels videobasierter Simulationsumgebungen.

Hierfür werden Drehbücher zu Standardversionen von Unterricht mit enthaltenen diagnostisch relevanten Situationen zur Unterrichtssequenz „Haut“ (5. Jgst.) entwickelt. Anhand der ausgearbeiteten Drehbücher werden anschließend fiktionale Videovignetten mit einer Schauspielerklasse produziert. Experten (N=6 Biologielehrkräfte) und Novizen (N=6 Lehramtsstudierende des Fachs Biologie) beobachten anschließend die Unterrichtsausschnitte. Mittels think aloud Interviews werden Rückschlüsse über die von den Probanden identifizierten diagnostisch-relevanten Situationen gewonnen. Zusätzlich bearbeiten die Probanden eine Aufgabe zur Unterrichtsdiagnose, um exemplarisch die von den Probanden genutzten Strategien hinsichtlich des Diagnoseprozesses zu untersuchen. Die Ergebnisse werden genutzt, um die Aufgabenstellungen zu überarbeiten und diagnostisch-relevante Videovignetten auszuwählen, für welche jeweils Optimierungsvarianten entwickelt werden. Die ausgeschärften Videovignetten mit den alternativen Handlungsoptionen dienen als Grundlage für die Konstruktion der finalen videobasierten Simulationsumgebung. In einem zweiten Teil der Validierungsstudie soll anhand von freien Textfeldern untersucht werden, inwiefern sich Experten (N=15 Biologielehrkräfte) und Novizen (N=15 Lehramtsstudierende des Fachs Biologie) hinsichtlich des Diagnoseergebnisses unterscheiden. Außerdem werden die wahrgenommene Authentizität (drei Items), sowie die kognitive Involviertheit (vier Items) über Fragebögen mit fünfstufiger Likertskala erhoben.

#### Erwartete Ergebnisse

Insgesamt wird erwartet, dass die Videovignetten innerhalb der Simulationen als authentisch wahrgenommen werden und eine kognitive Involviertheit gewährleistet ist. Zudem wird damit gerechnet, dass die diagnostisch-relevanten Situationen von den Probanden identifiziert werden, wobei Experten hierbei bessere Ergebnisse erzielen sollten. Allgemein wird somit von einer höheren Diagnosekompetenz bei Experten ausgegangen. Die Ergebnisse werden zur Überarbeitung der Videovignetten genutzt.

## Wie lassen sich kindliche Interessen messen? Entwicklung und Validierung eines bilderbasierten Interessentests

**Nadine Schneider, Katja Pässler**

Hochschule für Angewandte Psychologie FHNW, Schweiz; [nadine.schneider@fhnw.ch](mailto:nadine.schneider@fhnw.ch)

Berufliche Interessen sind wichtige Determinanten der Ausbildungs- und Studienwahl (Humphreys, Lubinski, & Yao, 1993) sowie bedeutsame Prädiktoren akademischer und beruflicher Leistung (Nye, Su, Rounds, & Drasgow, 2012) und Zufriedenheit (Nye, Su, Rounds, & Drasgow, 2017; Tsabari, Tziner, & Meir, 2005). Obwohl sich individuelle Unterschiede in Aktivitäts- und Tätigkeitspräferenzen bereits sehr früh zeigen, weiß man bisher relativ wenig darüber, wie sich Interessen bei Kindern ausbilden und entwickeln (Hartung, Porfeli, & Vondracek, 2005). Ein Grund dafür liegt im Fehlen geeigneter Instrumente zur Erfassung kindlicher Interessen (Tracey & Ward, 1998). Mit dem Inventory of Children's Activities wurde von Tracey und Ward (1998) zwar erstmals ein Fragebogeninstrument zur Erfassung der allgemeinen Interessenorientierung nach Hollands (1997) RIASEC Modell zur Verfügung gestellt, dieses lässt sich aber erst bei Kindern ab der 4. Klasse sinnvoll einsetzen. Im Rahmen einer Längsschnittstudie zur Entwicklung geschlechtsspezifischer MINT-Interessen wurde deshalb der Kinder-Bilderinteressentest (K-BIT) entwickelt, der auf einer aktualisierten Version des Inventory of Children's Activities (ICA-3) von Tracey und Caulum (2015) basiert. Der K-BIT ermöglicht die Messung kindlicher Interessen basierend auf dem RIASEC-Modell von Holland (1997) anhand von Tätigkeiten aus der Alltagswelt, die den Kindern in Form eines Bilderbuchs präsentiert werden. Der K-BIT umfasst insgesamt 30 Bilder, jeweils fünf Bilder pro Interessendimension. Je nach Alter der Kinder kann der K-BIT im Einzel- oder Gruppensetting angewandt werden.

An zwei Stichproben wurde die Reliabilität und Validität des Instruments untersucht. Kohorte 1 besteht aus 371 Kindergartenkindern und Primarschülerinnen und -schülern zwischen 4 und 8 Jahren (MW = 6.1, SD = .98). Kohorte 2 besteht aus 349 Primarschülerinnen und -schülern zwischen 9 und 13 Jahren (MW = 11.3, SD = .75). Analysen zur internen Konsistenz verweisen auf gute bis zufriedenstellende Skalenreliabilitäten (Kohorte 1:  $\alpha = .47 - .60$ ; Kohorte 2:  $\alpha = .75 - .82$ ). Die Ergebnisse der Retest-Reliabilität (Kohorte 2:  $rtt = .44 - .63$ ) sind am Ende der Primarschule mit den bei Jugendlichen und jungen Erwachsenen gefundenen Koeffizienten vergleichbar. In der jüngeren Kohorte fallen sie deutlich niedriger aus ( $rtt = .14 - .29$ ) und deuten auf noch instabile Interessenprofile in der Altersgruppe der 4- bis 8-Jährigen hin. Die Korrelationskoeffizienten der jeweils konvergenten Dimensionspaare von ICA-3 und K-BIT liegen zwischen  $r = .76$  und  $r = .86$ . Explorative Faktorenanalysen bestätigen in der Kohorte der 9- bis 13-Jährigen eine zugrundeliegende 6-Faktoren-Struktur. Allerdings deuten die Resultate konfirmatorischer Analysen (RTOR- und CIRCUM-Analysen) darauf hin, dass die Interessenstruktur von Primarschülerinnen und -schülern noch nicht vollumfänglich den strukturellen Annahmen des RIASEC-Modells entspricht. Geschlechtsspezifische Vergleiche zeigen, dass sich dies vor allem auf den schlechten Fit der Jungenstichprobe zurückführen lässt (C1Jungen = .06,  $p = .42$ ; C1Mädchen = .65,  $p = .02$ ). Diese Ergebnisse gehen einher mit bisherigen empirischen Befunden, dass die Interessenstruktur erst im Jugendalter (ab etwa 14 Jahren) jener von Erwachsenen entspricht (Tracey, 2002). Die geschlechtsspezifischen Analysen bestätigen, dass die Modellpassung bei Mädchen deutlich besser ausfällt als bei Jungen (Lent, Tracey, Brown, Soresi, & Nota, 2006)

## Der Erwerb professionellen Wissens im Lehramtsstudium – Konzeption eines Messinstrumentes –

**Müser Sinja, Dorothee Gronostay, Oliver Kraft, Sebastian Olschak, Andreas Büchter, Inga Gryl, Detlev Leutner, Sabine Manzel, Stefan Rumann, Vanessa Pupkowski, Maik Walpuski**

Universität Duisburg-Essen, Deutschland; [sinja.mueser@uni-due.de](mailto:sinja.mueser@uni-due.de)

Theoretischer Hintergrund

Die Modulhandbücher der lehramtsbildenden Studiengänge an Universitäten und Pädagogischen Hochschulen beschreiben Kompetenzen, die die angehenden Lehrkräfte erreichen sollen. Zu ihnen zählen fachwissenschaftliche, fachdidaktische und bildungswissenschaftliche Kompetenzen (Blömeke, 2009), die das Professionswissen von Lehrkräften operationalisieren (u.a. Shulman, 1986; Baumert & Kunter, 2006; Abell, 2007; Borowski et al., 2010). Das Fachwissen (content knowledge, CK) wird definiert durch ein vertieftes Verständnis schulrelevanter fachlicher Inhalte (Krauss et al., 2008). Das fachdidaktische Wissen (pedagogical content knowledge, PCK) lässt sich nach Park & Oliver (2008) in fünf weitere Facetten unterteilen: „Knowledge of Students' Understanding“, „Knowledge of Curriculum“, „Knowledge of Instructional Strategies“, Knowledge of Assessment of Students' Learning of Subject Matter“, „Orientation to Teaching Subject Matter“. Das Konzept des pädagogischen Wissens (pedagogical knowledge, PK) wird ergänzt um erziehungswissenschaftliches, psychologisches und soziologisches Wissen, das im Sinne bildungswissenschaftlichen Wissens über die reine Unterrichtsgestaltung hinausgeht und im Sinne der KMK-Standards (KMK, 2014) für die Lehrerbildung definiert wird.

Ziel dieser Studie ist es, die Kompetenzen der Studierenden am Ende der universitären Ausbildungsphase in den drei Facetten des Professionswissen (in den Fächern Mathematik, Sozialwissenschaften, Sachunterricht sowie den fächerübergreifenden Bildungswissenschaften) zu untersuchen und mit den Vorgaben aus den Modulhandbüchern sowie extern formulierten Standards zu vergleichen.

Fragestellungen

- 1) Erwerben die Lehramtsstudierenden die geforderten Kompetenzen in den bildungswissenschaftlichen Studiengängen?
- 2) Erwerben die Lehramtsstudierenden die geforderten fachlichen und fachdidaktischen Kompetenzen in ihrem jeweiligen Studienfach
  - a. Mathematik?
  - b. Sachunterricht?
  - c. Sozialwissenschaften?

Methode

Auf Grundlage der Kompetenzformulierung in den Modulhandbüchern sowie externen Standards und Vorgaben (KMK Standards für die Lehrerbildung, Kerncurriculum für die Ausbildung im Vorbereitungsdienst, Schulcurricula) wurde das Professionswissen fächerabhängig kategorisiert und operationalisiert. Zu den jeweiligen Kategorien wurden Testinstrumente entwickelt (Paper-Pencil, Multiple-choice Single-select, True-false-, sowie Kurzantwort-Aufgaben), die die Facetten des Professionswissens angehender Lehrkräfte erfassen sollen. Für die drei Bereiche Fachwissen, fachdidaktisches Wissen sowie bildungswissenschaftliches Wissen wurden dabei Testhefte im Umfang von 44 bis 115 Items generiert, welche jeweils in einem Zeitraum von 90 bis 120 Minuten bearbeitet wurden.

Die Stichprobe der Pilotstudie bilden Lehramtsstudierende verschiedener universitärer Standorte am Ende des Bachelor-Studiengangs sowie zu Beginn des Master-Studiengangs. Ziel der Pilotstudie ist es, die neu entwickelten Testinstrumente zu validieren, sie hinsichtlich der Facetten des Professionswissens zu analysieren und anschließend für die Hauptstudie zu optimieren. Die Auswertung der Daten erfolgt mit Hilfe klassischer und probabilistischer Testtheorien.

#### Ergebnisse der Pilotierung

Die Auswertung der Pilotstudie erfolgte mit Hilfe von IRT-Analysen. Die Daten berufen sich auf die Analyse der Testinstrumente aus welcher Implikationen für die Hauptstudie abgeleitet werden sollen.

Das Testinstrument zur Messung des bildungswissenschaftlichen Wissens weist in einer ersten Pilotierungsstudie bei einer Stichprobe von 177 Studierenden eine gute Reliabilität (EAP/PV .768) auf. Die eingesetzten Items (N = 60) zeigen zufriedenstellende Kennwerte ( $0.9 < \text{Infit} < 1.1$ ). Die Lösungswahrscheinlichkeit der Items liegt zwischen 15.6 % und 87.4 %. Die Varianz liegt in einem eher unteren Bereich (0.254), jedoch wird erwartet, dass sich diese bei steigender Anzahl der Items in einer weiteren Pilotierungsstudie erhöhen wird.

Die Testinstrumente zur Messung des fachdidaktischen Wissens und des Fachwissens mit dem Schwerpunkt Sachunterricht zeigen bei einer Stichprobe von 43 Studierenden gute Reliabilitäten (EAP/PV fachdidaktisches Wissen .695, EAP/PV Fachwissen .791). Die eingesetzten Items (N = 90) zeigen zufriedenstellende Kennwerte ( $0.8 < \text{Infit} < 1.2$ ). Zudem konnte gezeigt werden, dass es sich bei den eingesetzten Testinstrumenten um 2 trennbare Skalen handelt (Chi-Quadrat:  $\Delta\text{Deviance} = 53.81, p < .001$ ). Die Testinstrumente für die Schwerpunkte Mathematik und Sozialwissenschaften werden in einer zweiten Pilotierungsstudie eingesetzt. Die Ergebnisse psychometrischen Qualität aller neu entwickelten Testinstrumente werden auf der Tagung präsentiert.

## **Befürchtungen und Stresserfahrungen von Lehramtsstudierenden in der Schul-Universitäts-Partnerschaft „Schuladoption“**

**Andreas Bach**

Paris Lodron Universität Salzburg, Österreich; [andreas.bach@sbg.ac.at](mailto:andreas.bach@sbg.ac.at)

**\*Theoretischer Hintergrund:\*** Partnerschaftsmodelle zwischen Schulen und Lehrerbildungsinstitutionen gelten als ein Schlüsselkonzept in den aktuellen Lehrerbildungsprogrammen vieler Länder (Kruger et al., 2009; Tsui, Edwards & Lopez-Real, 2009; White & Forgasz, 2016). In Norwegen wurde das Partnerschaftsmodell „Schuladoption“ entwickelt (Haugaløkken & Ramberg, 2007), das inzwischen in mehreren europäischen Ländern als Praktikumskonzept durchgeführt wird (Bach, 2017). Kernelement der Schuladoption ist eine Praktikumswoche („Adoptionswoche“), in der Lehramtsstudierende den Unterricht an einer Praktikumschule übernehmen, während das Lehrerkollegium in dieser Zeit an einem für die Schulentwicklung der Schule spezifischen Fortbildungsprogramm teilnimmt. Die Schuladoption gibt somit die typische Trainingsstruktur von Schulpraktika mit einem Mentor im Klassenzimmer auf und spricht den Lehramtsstudierenden eine im Vergleich zu herkömmlichen Praxisphasen hohe Verantwortung für den Unterricht und die Schule zu. Gleichzeitig intensiviert die Universität ihre Lernbegleitung in der Schule, um Lehramtsstudierende effektiver auf die komplexe Arbeit vorzubereiten. Bislang wurde nicht untersucht, wie die hohe Komplexität von Unterricht und Schule in derartigen Praktikumsformen erlebt wird. Der Fokus der Studie ist es deshalb, zu untersuchen, inwiefern Lehramtsstudierende die Schuladoption als stressreich erleben, und welche Faktoren relevant sind, um mit der hohen Komplexität umgehen zu können.

**\*Fragestellung:\*** Folgende Fragestellungen werden im Einzelnen untersucht:

- (1) Welche Befürchtungen haben Lehramtsstudierende im Hinblick auf die Adoptionswoche und inwiefern sind die Befürchtungen aus ihrer Sicht eingetreten?
- (2) Wie verändert sich das Herausforderungs- und Bedrohungsleben von Lehramtsstudierenden in der Adoptionswoche?
- (3) Welche Faktoren sind aus Sicht von Lehramtsstudierenden bedeutsam, um mit der Komplexität im Klassenzimmer während der Schuladoption umgehen zu können?

**\*Methode:\*** Die qualitativen und quantitativen Analysen basieren auf einer schriftlichen Befragung von Lehramtsstudierenden (N = 11) im Prä-Post-Design an der Universität Flensburg, die das Projekt im Rahmen eines Praxissemesters im dritten Fachsemester MA Lehramt an Grundschule absolvierten. Die Befürchtungen der Lehramtsstudierenden, die retrospektive Einschätzung, ob diese eingetreten sind, und die Faktoren, die bedeutsam waren, um mit der Komplexität im Klassenzimmer und in der Schule während des Schulpraktikums umgehen zu können, konnten in einem freien Antwortfeld beschrieben werden. Die qualitativen Daten wurden inhaltsanalytisch ausgewertet, indem die Antworten manuell kodiert und entsprechend der Fragestellungen der Studie in Kategorien zu zentralen Themen gruppiert wurden (Mayring, 2015). Das Stresserleben der Lehramtsstudierenden wurde mit zwei Skalen zum Herausforderungsleben und Bedrohungsleben von Jerusalem (1999) erfasst. Die Mittelwertunterschiede im Herausforderungs- und Bedrohungsleben vor und nach der Schuladoption wurden mit Hilfe des T-Tests für abhängige Stichproben mit dem Programm SPSS analysiert.

**\*Ergebnisse:\*** Die Ergebnisse zeigen, dass sich die Befürchtungen den drei Hauptkategorien „Classroom Management“, „Anerkennung als Lehrkraft“ und „erhöhter Arbeitsaufwand“ zuordnen ließen. Diese drei Hauptkategorien stimmen mit den bisherigen Studien überein, die diese Bereiche ebenfalls als zentrale Stressfaktoren von Lehramtsstudierenden in Schulpraktika identifizierten (Chaplain, 2008; Danyluk, 2013; Ekşi & Yakışık, 2016; Kyriacou & Stephens, 1999). Die zu Beginn der Praktikumsphase geäußerten Befürchtungen einer möglichen Überforderung zum Beispiel durch störendes Schülerverhalten haben sich aus Sicht der Lehramtsstudierenden nicht bewahrheitet. Die Befunde der Studie zeigten einen signifikanten Anstieg im Herausforderungsleben von großer Effektstärke und einen tendenziellen Rückgang des Bedrohungslebens. Die Lehramtsstudierenden waren nach dem Schulpraktikum statistisch bedeutsam stärker davon überzeugt, berufliche Anforderungen bewältigen zu können. Die Ergebnisse deuten darauf hin, dass in Schulpraktika komplexe Lernsituationen arrangiert werden können, sofern diese hinreichend strukturiert vorbereitet und begleitet werden. In dieser Studie konnten drei zentrale Faktoren identifiziert werden, die dazu beitragen, dass die Komplexität der beruflichen

Anforderungen im Schulpraktikum als bewältigbar erlebt wird. Für den Umgang mit der Komplexität waren neben der Vorbereitung auf das Schulpraktikum vor allem die Begleitung der Praxiserfahrungen in einer Lerngemeinschaft und das Team-Teaching der Lehramtsstudierenden entscheidend. Insbesondere der Austausch mit Kommilitonen in einer universitär organisierten Lerngemeinschaft scheint ein bedeutender stressreduzierender Faktor in Schulpraktika zu sein.

## **Kompetenzdiagnostik im Hochschulpraktikum**

**Pascal Engler**

Universität Bern, Schweiz; [pascal.engler@bfh.ch](mailto:pascal.engler@bfh.ch)

Ein Dissertationsprojekt zur Überprüfung professioneller Kompetenzen der Sozialen Arbeit an der Schnittstelle von Studium und Praxis

Gegenstand der Dissertation

Studierende der Sozialen Arbeit absolvieren während dem Bachelor-Studium zwei Praxismodule (Praktika). Während der Praxisausbildung werden die Studierenden von einer ihnen zugewiesenen Fachperson der Sozialen Arbeit (Praxisausbildende) begleitet und beurteilt. Fällt die summative Beurteilung ungenügend aus, kann das Praxismodul einmal wiederholt werden. Eine erneute ungenügende Bewertung führt zur Exmatrikulation. Die Bewertung der Praxismodule hat damit eine wichtige Gatekeeping-Funktion, indem sie promotionsrelevant ist (Miller & Koerin, 2002).

Umso überraschender ist deshalb die Tatsache, dass im deutschsprachigen Raum die Diagnostik der Praxiskompetenzen weitgehend der Praxis überlassen wird und die Hochschule dafür keine zuverlässigen Messinstrumente zur Verfügung stellt. Das Fehlen eines validierten Instrumentes für die Kompetenzdiagnostik ist aus unterschiedlichen Gründen bedeutsam:

1) Die Hochschule delegiert während dem Praxismodul einen Teil des Kompetenzerwerbs an die Praxisorganisationen. Weil aber die Hochschule für die Professionskompetenz verantwortlich bleibt (Regehr, Bogo, & Regehr, 2011) und am Schluss des Studiums über die Diplomierung entscheidet, muss die Hochschule sicherstellen, dass die Bewertung der delegierten Kompetenzen valide, reliabel und objektiv ist.

2) In der Literatur wird wiederholt darauf aufmerksam gemacht, dass die verzerrenden Einflussfaktoren, wie z.B. die persönliche Beziehung bei der Bewertung gewichtig sind (u.a. Müller Fritschi & Roth, 2014; Güntert, 2014). Dies wird auch mit der fehlenden Validität und Reliabilität der bestehenden Bewertungsinstrumente (Bogo, 2010) begründet.

3) Die Kompetenzdiagnostik und insbesondere die Notensetzung ist für Praxisausbildende eine stressvolle Angelegenheit (Bogo, Regehr, Power, & Regehr, 2007). Dieses Unbehagen hat auch damit zu tun, dass die Praxisausbildenden ihre Beurteilung als zu wenig abgestützt und somit objektiv empfinden (Güntert, 2014).

4) Bei Studierenden, welche die Beurteilung anzweifeln und als nicht aussagekräftig empfinden, kann dies zu einem Motivationsverlust führen (Güntert, 2014). Die Bewertung ist aber nur dann erfolgreich und für den Lernprozess sinnvoll, wenn die Studierenden diese auch verstehen und annehmen können (Bogo, 2010).

Fragestellung

Im Zentrum der Dissertation steht folgende Fragestellung:

Wie können professionelle Kompetenzen der Sozialen Arbeit an der Schnittstelle von Studium und Praxisfeld operationalisiert und beurteilt werden?

Ziel des Vorhabens ist es, ein valides Beurteilungsraster für die Praxisausbildung der Sozialen Arbeit zu entwickeln. Ein Beurteilungsraster setzt einen Konsens von Hochschule und der Praxisorganisationen voraus. Dieser inhaltliche Dialog ist zentral um ein gemeinsames Verständnis von Professionalität und Berufsbefähigung zu erarbeiten.

Methodisches Vorgehen

Im empirischen Teil rücken zunächst validierte diagnostische Instrumente (PBE und CBE)\*, die im englischsprachigen Raum für die Kompetenzdiagnostik während der Praxisausbildung der Sozialen Arbeit eingesetzt werden, in den Mittelpunkt. Die ausgewählten Instrumente werden auf Deutsch übersetzt und für eine Replikation von Studien aus Kanada und den USA verwendet. Dabei setzen erfahrene Praxisausbildende die diagnostischen Instrumente bei der Bewertung von Studierenden ein. Die Resultate aus dieser Testphase sowie die ergänzenden Rückmeldungen der Praxisausbildenden, welche im Rahmen von Interviews erhoben werden, dienen dazu, ein eigenes Instrument zu entwickeln.

\*=PBE: Practice-based evaluation tool / Regehr et al. (2011) / CBE: Competency-based evaluation tool / Bogo et al. (2004)

Ausblick / Ergebnisse (Hypothetisch)

Sollten die empirischen Ergebnisse eine deutliche Konvergenz zwischen den Forschungen aus dem englischsprachigen Raum und der vorliegenden Forschung aufzeigen, so könnte zumindest vermutet werden, dass es eine Generalisierbarkeit von Meta-Kompetenzen der Sozialen Arbeit gibt, welche unabhängig vom Untersuchungsort nachgewiesen und beurteilt werden können. Mit dem neuen Beurteilungsraster wird zum ersten Mal ein validiertes deutschsprachiges Instrument zur Verfügung stehen, welches mit nachfolgenden Evaluationsstudien weiterentwickelt werden kann. Für die Gewährleistung von Praxisrelevanz, Anwenderfreundlichkeit und einem gelingenden Implementierungsprozess sind für das Projekt der Dialog zwischen den Praxisorganisationen und der Hochschule zentral.

## **Forschendes Lernen in der frühen und mittleren Adoleszenz. Der Einfluss von Zielbindung durch Interesse und Selbstwirksamkeit und die medierende Rolle von Selbstkontrolle**

**Sabine Schweder, Judith Weckwerth**

Universität Greifswald, Deutschland; [schweders@uni-greifswald.de](mailto:schweders@uni-greifswald.de)

Theoretischer Hintergrund

Ziele sind bedeutende Motive und geben Handlungen und Lernen Richtung (Bandura, 1988; Locke & Latham, 1990, 2013). Je enger die Zielbindung, desto entschlossener und kreativer agieren Personen in darauf bezogenen Handlungs- und Lernprozessen. Zielbindung nimmt in dem Maße zu, wie Personen bei der Zielentwicklung beteiligt werden und steigert sich bei selbstentwickelten Zielen bis zu einem Maximum (Deci & Ryan, 1991; Hollenbeck, Williams, & Klein, 1989; Zimmerman, 2008). Grund dafür ist die bei partizipativer Zielentwicklung mögliche Bezugnahme zu intrinsischem Interesse und

Selbstwirksamkeitserfahrungen (Zimmerman, Schunk, & Boucher, 2017). Diese Komponenten sind im schulischen Kontext wegen positiver Effekte auf Selbstregulation und Selbstkontrolle (Hidi & Ainley, 2008; Hidi & Renninger, 2006; Pintrich & Zusho, 2002; Zimmerman & Cleary, 2006), Lernanstrengung (Ainley, 2006; Bandura, 2013), Elaboration (Ainley, Corrigan, & Richardson, 2005; Baumert & Köller, 1998) und Absorption (Csikszentmihalyi, 1990, 2013; Fullagar & Fave, 2017; Hidi & Ainley, 2008) bedeutsam.

Die an der Studie beteiligten Schulen haben ein Konzept unter dem Namen „Forschendes Lernen“ eingeführt, was Adoleszente die Entwicklung von eigenen Lern- und Explorationszielen und die Umsetzung und Reflexion darauf bezogener Strategien ermöglicht. Begleitung und strukturelle Unterstützung beim Aufbau einer Zielbindung ist zentrales Merkmal und fördert bei den Schüler/-innen die Bezugnahme zu individuellen Interessen und Selbstwirksamkeitserfahrungen.

Ausgehend von Ergebnissen neurowissenschaftlicher Studien (Braams, van Duijvenvoorde, Peper, & Crone, 2015; Telzer, 2016) werden Adoleszente vorzugsweise durch personalisierte und explorierende Lernbedingungen angesprochen (Harackiewicz & Knogler, 2017; Yeager, Lee, & Dahl, 2017). Interpretationsrahmen des Konzepts ist das Drei-Phasen-Modell zum selbstregulierten und problemösenden Lernen von Zimmerman und Campillo (2003). Die Studie hat für dieses Konzept altersbezogene Wirkungen zum Einfluss von Interesse und Selbstwirksamkeit, unter Mediation von Selbstkontrolle, auf Anstrengungsbereitschaft, Absorption und Elaboration untersucht, da davon ausgegangen wird, dass (1) sich das Bewusstsein für motivationale, metakognitive und kognitive Prozesse in der Adoleszenz generell verändert (Demetriou, 2000; Gestsdottir & Lerner, 2008) und (2) Schulalter und Motivation sich negativ bedingen (Gottfried, Flemming, & Gottfried, 2001; Harackiewicz & Knogler, 2017). Gestalter schulischer Lehr- und Lernkonfiguration erhalten Informationen zu Wirkung und Rahmenbedingungen des Konzepts. Folgende Hypothesen wurden dafür untersucht:

H1: Es gibt keine altersspezifischen Unterschiede beim Zusammenwirken der Variablen.

H2: Selbstkontrolle fungiert in allen Zusammenhängen als Mediator.

H3: Interesse und Selbstwirksamkeit beeinflussen Selbstkontrolle, Anstrengungsbereitschaft, Elaboration und Absorption zu einem späteren Zeitpunkt positiv.

H4: Es gibt altersspezifische Mittelwertsunterschiede bei Interesse, Selbstwirksamkeit, Selbstkontrolle, Anstrengungsbereitschaft, Absorption und Elaboration.

#### Methoden

Die empirische Grundlage der Studie bilden längsschnittliche quantitative Fragebögen von 754 Schülerinnen und Schüler aus sechs Sekundarschulen Mecklenburg-Vorpommerns und Schleswig-Holsteins. Ziel der Studie war es, zwei Altersgruppen zu untersuchen. In Gruppe 1 (t1: n = 395; Mage = 12,7; SD 0,66; 54,9 % Mädchen; t2: n = 401; Mage = 12,8; SD = 0,67; 54,6 % Mädchen) befinden sich Schülerinnen und Schülern der Klassenstufe 6 und 7 und in zweiter Gruppe (t1: n = 359; Mage = 14,4; SD = 0,66; 44,8 % Mädchen; t2: n = 347; Mage = 14,6; SD = 0,65; 45,5% Mädchen) der 8. und 9. Klassenstufe. Um zunächst die Unterschiede zwischen beiden Gruppen in Bezug auf die postulierten Zusammenhänge und Unterschiede in den Mittelwerten nachzuweisen, wurden Multigruppen-Strukturgleichungsmodelle mit Mplus 7.3 und konfirmatorische Faktorenanalysen für einen latenten Mittelwertvergleich konzipiert.

#### Ergebnisse

Wie mit H1 angenommen, wurden keine signifikanten altersbedingten Unterschiede im Zusammenwirken der untersuchten Variablen nachgewiesen. H2 bestätigt sich ebenfalls. Selbstkontrolle mediert alle bestehenden Beziehungen. Die Beziehung zwischen Selbstwirksamkeit und Absorption mediert Selbstkontrolle überraschenderweise vollständig, womit H3, die in allen Fällen von signifikant positiven Effekten in den Beziehungen ausgeht, zu Teilen widerlegt wurde. H4 kann teilweise bestätigt werden. Jüngere Jugendliche weisen eine höhere Bezugnahme zu Interessen, mehr Bereitschaft zur Selbstkontrolle, ein höheres Ausmaß an Anstrengungsbereitschaft und Absorption auf. Selbstwirksamkeit und Elaboration unterscheiden sich dagegen nicht. Die Ergebnisse werden vor dem Hintergrund der bestehenden Forschung diskutiert.

## **Energiekonzepte angehender Chemielehrkräfte**

**Annabel Pauly, Arnim Lühken**

Goethe-Universität Frankfurt, Deutschland; [a.pauly@chemie.uni-frankfurt.de](mailto:a.pauly@chemie.uni-frankfurt.de)

#### Einführung und theoretischer Hintergrund

In der Alltagssprache ist der Begriff Energie in vielerlei Kontexten vorhanden. In der medialen Aufmerksamkeit stehen die klassischen Misskonzepte des Energieverbrauchs und der Energiegewinnung. Im Kontrast dazu stehen die naturwissenschaftlichen Konzepte der Energieerhaltung und der Energieumwandlung, die sich auch beispielsweise im Basiskonzept Energie wiederfinden (Schmidkunz, Parchmann 2011, KMK 2004). Schon alleine durch diese begrifflichen Unklarheiten scheint es kaum verwunderlich, dass im Rahmen der länderübergreifenden Energiebildungsstudie (Euler 2013) festgestellt wurde, dass die „Energiebildung“ in allen Schulstufen intensiviert werden muss. Für die LehrerInnenbildung ist dabei bemerkenswert, dass befragte LehrerInnen die Notwendigkeit einer besseren fachlichen und fachdidaktischen Ausbildung im Bereich der Energie an der Hochschule sehen. Im Hinblick auf die Wirkung von LehrerInnenbildung (Terhart 2012) sowie auf professionelles Handeln ist es notwendig, dass die Lehramtsstudierenden schon in der Phase der universitären Lehramtsausbildung ein Energiekonzept aufbauen. Im Bereich der Physikdidaktik liegen Forschungsergebnisse zum Energieverständnis von SchülerInnen sowie LehrerInnen vor (vgl. u.a. Duit 2014, Viering 2012, Pahl 2013). In der Chemiedidaktik gibt es ebenfalls Ergebnisse zu SchülerInnenvorstellungen (vgl. u.a. Barke 2006), jedoch stellen ähnliche Untersuchungen im Bereich der Hochschuldidaktik weitestgehend ein Desiderat dar.

Aus diesem Grund sollen Energiekonzepte angehender Chemielehrkräfte, die sich am Ende ihrer universitären Ausbildung befinden, erfasst werden.

Die Beschreibung der Energiekonzepte der Studierenden geschieht auf folgenden Ebenen:

- Fachlicher Hintergrund im Chemieunterricht
- Fachdidaktisches Wissen
- Versprachlichung

#### Fragestellung

- Über welches fachliche und fachdidaktische Wissen im Themengebiet Energie verfügen Lehramtsstudierende der Chemie?

- Inwiefern versprachlichen die Studierenden ihr Energiekonzept?

#### Methode

Um den relevanten fachlichen Hintergrund der Energiethematik im Chemieunterricht zu erfassen, wurde eine Schulbuchanalyse durchgeführt. Es wurde nach dem Prinzip der inhaltlichen Strukturierung (Mayring 2014) vorgegangen. Als Grundlage diente ein deduktives Kategoriensystem, das aus den theoretischen Erkenntnissen zum physikalischen Energiekonzept sowie deren Bearbeitungen in der Biologiedidaktik abgeleitet wurde (vgl. dazu Duit 2014, Viering 2012, Wernecke 2016). Um dieses Energiekonzept auf die Chemiedidaktik zu erweitern, wurden die Inhalte, die sich nicht dem physikalischen Konzept zuordnen lassen, mittels induktiver Kategorienbildung gegliedert. Die deduktiv/induktiv erarbeiteten Kategorien sind die Basis zur Erfassung des fachlichen Hintergrunds der Energiethematik im Chemieunterricht. Aus den Elementen des fachdidaktischen Wissens für die universitäre Ausbildung angehender Chemielehrkräfte (Großbrahm 2013) wurden diejenigen Facetten ausgewählt, die sich themenspezifisch auf die Energie beziehen lassen. Nicht beachtet wurden Facetten, die einen allgemeinen Aspekt des Chemieunterrichts abdecken.

Um alle drei Ebenen des Energiekonzepts von Lehramtsstudierenden der Chemie wurde eine Interviewstudie basierend auf jeweils zwei Leitfadenterviews durchgeführt. In einer ersten Befragung definieren die Studierenden die in der Schulbuchanalyse entwickelten Facetten des relevanten fachlichen Wissens und geben jeweils Beispiele (fachchemisch, alltagsbezogen) an. Die qualitative Auswertung (induktive Kategorienbildung (Mayring 2014)) dieser Befragung stellt die Grundlage für die folgenden Interviews dar. Neben der Abfrage des fachlichen Hintergrunds wird in diesen ebenfalls auf die grundlegenden, fachinhaltsbezogenen Facetten des fachdidaktischen Wissens eingegangen. Die Interviews werden durch visual elicitation (Zehe, Belz 2016) mit Leitfragen aufgrund von Schulbuchseiten und Sequenzen aus Unterrichtsvideos durchgeführt. Die Auswertung beider Interviewdurchgänge erfolgt ebenfalls mittels induktiver Kategorienbildung nach Mayring. Dabei kann es durchaus passieren, dass fachliches Wissen und Versprachlichung desjenigen nicht korrelieren. Um diese Aspekte in die Analyse miteinzubeziehen, wird an einzelne interpretationsbedürftige Textstellen eine Kontextanalyse (Explikation (Mayring 2014)) durchgeführt.

#### Erste Ergebnisse

Erste Ergebnisse der Auswertung der ersten Forschungsfrage nach dem fachlichen Wissen der Lehramtsstudierenden sowie dem Bezug zur Versprachlichung desjenigen in den Befragungen lassen erkennen, dass eine deutliche Diskrepanz zwischen den kurzen schriftlichen Definitionen der Studierenden der fachlichen Facetten sowie der Ausformulierung und Explikation in den Interviews besteht. Diese Differenzen beziehen sich vor allem darauf, dass die Studierenden fachlich viel wissen, dieses Wissen aber nicht in allen Facetten in angemessener Fachsprache auszudrücken vermögen. Im Tagungsbeitrag ausgewählte Facetten des fachlichen Wissens im Hinblick auf die Explikation der Studierenden vorgestellt und interpretiert.

### **FIS-Studie Fähigkeiten Inklusion Selbsteinschätzung**

**Camilla Haas, Ellen Aschermann, Charlotte Brenk**

Universität zu Köln, Deutschland; [camillaahaas1990@gmail.com](mailto:camillaahaas1990@gmail.com)

Das akademische Fähigkeitsselbstkonzept besitzt eine bedeutsame Rolle in Lehr- und Lernkontexten (Möller & Trautwein, 2009; Hasselhorn & Gold, 2014). Ein positives Fähigkeitsselbstkonzept gehört nicht nur zu den Zielen gelungener Bildungsprozesse, sondern steht auch in positiver Korrelation mit der akademischen Leistung (Cimeli, Neuenschwander, Röthlisberger & Roebers, 2013). Veränderungen des Selbstkonzepts sind unterschiedlichen Vergleichsprozessen geschuldet. Die Wirkung von Referenzgruppen auf die Entwicklung des Fähigkeitsselbstkonzepts im akademischen Kontext konnte in vielen Studien belegt werden (Kelley, 1952; Marsh, 2005). Es wurde bisher oft gezeigt, dass leistungsstarke Referenzgruppen einen negativen Effekt auf das Selbstkonzept bei Schülerinnen und Schülern haben sowie umgekehrt, dass Schülerinnen und Schüler vom Vergleich mit einer relativ schwächeren Vergleichsgruppe profitieren (Aust, Watermann, & Grube, 2010; Gerlach, Trautwein, & Lütke, 2007). Dieser Effekt wird BFLPE (Big Fish Little Pond Effect) genannt und konnte in etlichen Studien und unterschiedlichen akademischen Kontexten repliziert werden. In bisherigen Studien wurden jedoch hauptsächlich negative Effekte des Vergleiches mit einer relativ leistungsstärkeren Referenzgruppe geprüft, weshalb in dieser Untersuchung auch ein möglicher, positiver Effekt der Zugehörigkeit zu einer prestigeträchtigen Gruppe auf die Leistung (Basking in Reflected Glory Effect, Marsh, Kong & Hau, 2000; Trautwein & Lütke, 2005) von Interesse ist. Bisher wurde dies nur mit Schülern der Mittelstufe (Jhg. 6 und 8) an einer leistungshomogenen Schulform untersucht.

Im Rahmen der Inklusion werden die Schulklassen in Bezug auf das Leistungsniveau heterogener. Im Sinne des BFLPE sollte sich diese gesteigerte Heterogenität positiv auf das Fähigkeitsselbstkonzept leistungsstärkerer Schülerinnen und Schüler und negativ auf das Fähigkeitsselbstkonzept der relativ leistungsschwächeren Schülerinnen und Schüler auswirken (Trautwein, Köller, & Baumert, 2004). Gleichzeitig ist aufgrund der empirischen Forschungslage zu erwarten, dass die inkludierten Schülerinnen und Schüler von der Teilhabe am Regelunterricht im Vergleich zur gesonderten Beschulung aufgrund des Basking in Reflected Glory Effects und aufgrund höherer kognitiver Aktivierung im Unterricht profitieren.

Ziel des Projekts ist es, den Einfluss der inklusiven Beschulung auf das akademische Selbstkonzept der eigenen Begabung zu untersuchen. Dies wurde anhand dreier Erhebungszeitpunkten des Fähigkeitsselbstkonzeptes in Deutsch (Deutsch, Lesen und Schreiben) sowie der Leistung im Lesen und Schreiben untersucht. Ein Vergleich einer Experimentalgruppe (inkludierte Schülerinnen und Schüler mit Sonderpädagogischem Förderschwerpunkt Lernen) mit einer Kontrollgruppe (Förderschülerinnen/-schüler mit SPF Lernen) wurde durchgeführt.

#### Fragestellung:

Welche langfristigen Auswirkungen hat die inklusive Beschulung (d.h. mit relativ leistungsstärkerer Vergleichsgruppe) von Schülerinnen und Schülern mit dem Förderschwerpunkt „Lernen“ an Regelschulen in der 5. Jahrgangsstufe im Vergleich zu einer Kontrollgruppe bestehend aus Schülerinnen und Schülern mit dem Förderschwerpunkt „Lernen“ an Förderschulen in der 5. Jahrgangsstufe auf das akademische Fähigkeitsselbstkonzept und die Leistung?

#### Hypothese 1:

Die inklusive Beschulung führt langfristig zu einer signifikanten Verschlechterung des akademischen Fähigkeitsselbstkonzepts. Nach einem Schuljahr Teilnahme an der inklusiven Beschulung hat die Experimentalgruppe langfristig ein signifikant niedrigeres Fähigkeitsselbstkonzept als die Kontrollgruppe.

#### Hypothese 2:

Nach einem Schuljahr Teilnahme an der inklusiven Beschulung haben sich die Schulleistungen der Experimentalgruppe (an inkludierenden Schulen) im Vergleich zu der Kontrollgruppe langfristig signifikant verbessert.

Hypothese 3:

Ein positiv ausgeprägtes Fähigkeitsselbstkonzept wirkt sich positiv auf die Leistung aus.

Ergänzend wurden an den fünf Schulen jeweils acht Unterrichtshospitationen mit Videografie durchgeführt. Ebenso fanden Elternbefragungen und LehrerInneninterviews statt, die jedoch nicht im Rahmen des Promotionsprojekts ausgewertet werden sollen.

Aktuell befindet sich das Projekt nach dem dritten Messzeitpunkt der Hauptstudie, nachdem die Ergebnisse der Vorstudie die erwarteten Tendenzen bestätigen konnten und das entwickelte Instrument erprobt werden konnte.

### **Die Einstellung angehender Gymnasiallehrkräfte zu Mehrsprachigkeit im Klassenzimmer – Erkenntnisse zur Entwicklung im Verlaufe des Studiums und Professionalisierungschancen.**

**Maximilian Astfalk**

Universität Mannheim, Deutschland; [astfalk@uni-mannheim.de](mailto:astfalk@uni-mannheim.de)

Theoretischer Hintergrund:

An deutschen Gymnasien setzt sich die Schülerschaft insbesondere im urbanen Raum aus einer Fülle an Schülerinnen und Schülern mit unterschiedlichsten Erst- und Herkunftssprachen zusammen. Eine der Aufgaben der Lehrerschaft ist daher, dieser Mehrsprachigkeit Rechnung zu tragen und Ansätze zu entwickeln, sie produktiv in das Unterrichtsgeschehen einzubeziehen (Krumm/Reich 2013). Mehrsprachigkeit wird hierbei insbesondere unter dem Aspekt der Migration betrachtet, wobei es gilt Kinder mit einer anderen Erstsprache als dem Deutschen über ihre Herkunftssprache in den Unterricht zu integrieren (Wegner 2014). Wichtige Voraussetzung für das Gelingen ist somit die Bereitschaft und Kompetenz der Lehrkräfte diese Mehrsprachigkeit wahrzunehmen, mit ihr zu arbeiten und sie gewinnbringend in den Unterricht zu integrieren (Ekinci/Günesli 2016). Eine fundierte Vorbereitung auf diese Aufgabe, findet im Rahmen der Gymnasiallehrerausbildung an den Universitäten in Baden-Württemberg (noch) nicht statt. Es ist daher interessant zu klären, welche Grundlagen die Studierenden hier mitbringen und wie man bestenfalls darauf aufbauen kann.

Fragestellung:

Die Studie untersucht die Einstellung von Studierenden mit dem Studienziel Gymnasiallehramt an der Universität Mannheim zum Thema Mehrsprachigkeit. Neben der Erfassung der Kompetenz der Studierenden, steht insbesondere die Frage im Mittelpunkt, wie sich Studierende unterschiedlicher Fächer (insbesondere zwischen Deutsch und Fremdsprachen) und zu verschiedenen Studienzeitpunkten unterscheiden. Folgende 3 Thesen bilden den Kern der Studie:

1. Im Studienverlauf entwickeln die Studierenden zunehmend eine positive Grundeinstellung gegenüber dem mehrsprachigen Klassenzimmer.
2. Studierende einer Fremdsprache zeigen eine positivere Einstellung gegenüber Mehrsprachigkeit als jene, die keine Fremdsprache studieren.
3. Die Kompetenz im Umgang mit Mehrsprachigkeit entwickelt sich bei allen Studierenden im Verlauf des Studiums deutlich positiv.

Langfristige Zielsetzung ist es, auf der Basis dieser Fragen, Ideen zu entwickeln, wie die Lehramtsausbildung verbessert werden kann, um dem mehrsprachigen Klassenzimmer gerecht zu werden.

Methode:

Die Daten zur Beantwortung der Fragestellung werden mittels Onlineerhebung im Zeitraum August – Oktober 2017 unter den Studierenden des Gymnasiallehramts an der Universität Mannheim gewonnen. Die angestrebte Teilnehmeranzahl von N=100 ist zum jetzigen Zeitpunkt fast erreicht, sodass sich eine gute Datengrundlage ergeben sollte, um Aussagen über den Einfluss von Studienfach und Fachsemester sowohl auf die Einstellung als auch die Kompetenz der Studierenden zu ermöglichen.

Im ersten Schritt werden die Kompetenzen der Studierenden zum Thema migrationsbedingte Mehrsprachigkeit erhoben, angelehnt an das „Dynamic model of multilingualism“ von Herdina und Jessner (2002) sowie aktuelle Erkenntnisse der Mehrsprachigkeitstheorie (insb. Roche 2013). Darauf aufbauend werden Einstellungen der Studierenden zur Mehrsprachigkeit selbst, ihrer Rolle im Unterricht und eine Selbsteinschätzung der Studierenden hinsichtlich ihrer Kompetenz im Umgang damit abgefragt (angelehnt an Heyder/Schädlich 2015). Letzteres mit dem Ziel, die gemessene Kompetenz mit den Einstellungen und der Selbsteinschätzung zu vergleichen.

Die Pilotierung der Studie erfolgte im FS 2017 im Rahmen zweier Fachdidaktik Seminare mit Studierenden an der Universität Mainz.

Ergebnisse:

Aktuell liegen noch keine finalen Ergebnisse vor, da die Erhebung noch nicht beendet ist. Erste Zwischenergebnisse lassen aber darauf schließen, dass Thesen 1 und 2 sich bestätigen lassen, wohingegen für These 3 keine hinreichenden Belege gefunden werden.

Zum Zeitpunkt der Tagung werden die finalen Ergebnisse vorliegen.

### **Who is “German” and who is a “migrant?” A review of terminology used in psychology and education research on diverse populations**

**Ursula Moffitt, Linda P. Juang**

Universität Potsdam, Deutschland; [moffitt@uni-potsdam.de](mailto:moffitt@uni-potsdam.de)

Background

Researchers may spend weeks finding a model with the best fit possible, but not give more than cursory thought to terminology describing their participants. Yet, descriptive language holds real power, and can work to further marginalize members of minority groups, even if unintended (El-Tayeb, 2014). In Germany, as in most of continental Europe, race and ethnicity are taboo, so group labels focus on nationality and/or heritage (e.g., Simon, 2017). Thus, citizenship, place of birth, and parental

national origin are often self-reported in empirical research, but participants' perceived group belonging is not. Researchers therefore self select terminology to describe participants, resulting in high variation in group labels, conflation across generations, and the potential for reinforcing exclusionary notions of national belonging, even in research explicitly aimed at promoting inclusion.

#### Objectives

This review examines research conducted in Germany focusing on individuals both with and without migration background. We are interested in which labels researchers choose to describe their participants, as well as the populations under investigation. The aim is to offer an overview of current practice while highlighting the importance of reflexivity in research. We were guided by the following research question: Among recent studies published in the fields of psychology and education, which terminology is used to describe diverse populations and how do the labels used relate to notions of exclusion, belonging, and who is understood to be German?

#### Method

The review was conducted using PsycINFO and ERIC following the five-step guide by Arksey & O'Malley (2005). The steps include: (1) identifying a research question; (2) identifying relevant studies; (3) selecting studies; (4) coding the data; (5) synthesizing and summarizing the results (Arksey & O'Malley, 2005). Peer-reviewed articles published between 2001 and 2017 were examined. Search terms were: German OR Germany AND immigrant OR migrant OR migration background OR migrant background OR immigrant background. Articles published in German and English were included.

#### Results

A total of 197 articles were screened, of which 88 were retained based on relevance. A thematic analysis was then conducted (Braun & Clarke, 2012) and four main themes were chosen: (1) consistent usage of "German" vs. "immigrant/other nationality"; (2) consistent usage of "host/majority" vs. "immigrant/other nationality" despite participants of multiple generations; (3) inconsistent usage of "German/host/majority" vs. "migrant/immigrant" and "migration background" (4) consistent usage of "majority" vs. "minority/migration background". Most studies inconsistently used terminology, often using "migration background" in the introduction and "migrant/immigrant" in later sections. No differences were found between articles published in German vs. English.

#### Conclusions

Researchers hold responsibility as experts in their fields to represent their participants as accurately and equitably as possible (Stenius, 2016). Reflexivity is particularly important in the realm of education, where findings can directly inform practice and legitimize norms perpetuated in school settings. As this review makes clear, there is little internal consistency in how participants of non-German heritage are described in contemporary research. In many studies, individuals born and raised in Germany are referred to as "immigrants," in opposition to their peers without migration background who are given the label "German." At a time when over half of Turkish heritage participants in a recent study reported feeling like second-class citizens who will never be accepted as German (Pollack, Müller, Rosta, & Dieler, 2016), it is imperative that we as scholars take the time to reflect on the impact of such labels in our research. We do not offer strict terminological guidelines, as labels are highly context specific, but we do suggest clear differentiation between participant generation and the avoidance of using "German" only for those without ancestry from elsewhere.



## C01

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus S103

### Teamlernen und innovatives Arbeitsverhalten an beruflichen Schulen

**Andreas Widmann, Regina H. Mulder**

Universität Regensburg, Deutschland; [andreas.widmann@ur.de](mailto:andreas.widmann@ur.de)

#### Teamlernen und innovatives Arbeitsverhalten an beruflichen Schulen

In der beruflichen Bildung wurden interdisziplinäre Arbeitsteams eingeführt, die gemeinsam arbeiten und lernen sollen, um gesellschaftliche Herausforderungen (z.B. Integration von Flüchtlingen oder Qualitätssicherung) zu meistern (Truijen et al., 2013). Dies bedarf innovativer Lösungen und somit innovatives Arbeitsverhalten. Ergebnisse eines der Studien vorausgegangenen Literaturreviews (Widmann et al., 2016) zeigen domänenspezifische Unterschiede, wobei die Domäne der beruflichen Bildung unterrepräsentiert ist. Zudem zeigt sich, dass die Verbindung zwischen manchen Teamlernprozessen und innovativen Arbeitsverhalten vernachlässigt wurde, meist nur Teilbereiche von innovativen Arbeitsverhalten berücksichtigt und nicht als multi-dimensionales, dynamisches und kontextabhängiges Konstrukt untersucht wurde. Schließlich fehlen Längsschnittstudien. Diese Forschungslücken werden mit den beiden quantitativen Studien geschlossen.

Ziel ist es deshalb, wissenschaftliche Erkenntnisse über den Zusammenhang von Teamlernprozessen, innovativen Arbeitsverhalten und Bedingungsfaktoren in Arbeitsteams an beruflichen Schulen zu erlangen. Daraus resultieren folgende Forschungsfragen, welche mit Hilfe einer Querschnittstudie, und einer Längsschnittstudie, beantwortet werden sollen: Welche Zusammenhänge und Einflüsse bestehen zwischen (a) Bedingungsfaktoren und Teamlernprozessen (b) Teamlernprozessen und innovativen Arbeitsverhalten?

Teamlernen ist definiert als Ansammlung verschiedener Teamprozesse, welche unter verschiedenen Einflüssen stehen und zu Veränderung oder Verbesserung für das Team, seine Mitglieder oder die Organisation führen. Decuyper et al. (2010) postulieren in ihrem Teamlernmodell Verbindungen zwischen Teamlernprozessen, Teamlernprodukten und Bedingungsfaktoren.

In der Studie liegt der Fokus auf den Teamlernprozessen der Wissensteilung (Staples & Webster, 2008), Teamreflexion (Schippers, et al., 2007), Wissensspeicherung und -abrufung (van Offenbeek, 2001) und Boundary Spanning (Hirst & Mann, 2004), und dem verhaltensbezogenen Teamlernprodukt innovatives Arbeitsverhalten (Messmann, & Mulder, 2012).

Innovatives Arbeitsverhalten ist definiert als die Summe aller Aufgaben und Arbeitsaktivitäten, die zur Entwicklung von innovativen Produkten und Prozessen notwendig sind. Es wird als mehrdimensionales, dynamisches und kontextgebundenes Konstrukt gesehen. Demnach setzt sich innovatives Arbeitsverhalten aus den vier Dimensionen opportunity exploration, idea generation, idea promotion und idea realisation zusammen.

Folgende Bedingungsfaktoren können als Input mit den Teamlernprozessen in Bezug gesetzt werden: Teamstruktur (Bresman & Zellmer-Bruhn, 2013), Aufgabeninterdependenz (Ruppert, et al., 2011), und Gruppenwirksamkeit (Gibson et al., 2000).

#### Methode

Durch die querschnittliche Fragebogenstudie werden Zusammenhänge zwischen den Konstrukten analysiert. Die Stichprobe umfasst N=117 interdisziplinäre Arbeitsteams im Kontext der beruflichen Schule, bestehend aus n=595 Teammitgliedern.

Die längsschnittliche Fragebogenstudie (N=108 Teams, n=315 Teammitglieder) umfasst drei Messzeitpunkte, wobei der Querschnitt als erster Messzeitpunkt dient. Es sollen die Auswirkungen der Konstrukte auf innovatives Arbeitsverhalten als multidimensionales Konstrukt untersucht werden. Alle Variablen wurden mit validierten Skalen gemessen.

#### Auswertung

Neben Korrelationsanalysen, wurde ein Messmodell für die Teamlernprozesse spezifiziert. Zum Überprüfen des aufgestellten Forschungsmodells, wurde ein Strukturgleichungsmodell auf zwei Ebenen berechnet.

Neben Korrelationsanalysen erfolgt die Analyse im nächsten Schritt mittels eines autoregressiven Modells um reziproke und kausale Zusammenhänge zu analysieren. Die Analysen werden mit dem Programm Mplus 7 durchgeführt.

#### Ergebnisse

Die Ergebnisse der querschnittlichen Feldstudie weisen auf einen positiven Zusammenhang zwischen den Konstrukten hin. Besonders Teamreflexion ( $\beta=.54$ ,  $p<.05$ ) und Boundary Spanning ( $\beta=.61$ ,  $p<.05$ ) scheinen wichtige Teamlernprozesse für innovatives Arbeitsverhalten zu sein. Für die Teamlernprozesse zeigt sich vor allem Teamstruktur (von  $\beta=.73$  bis  $\beta=.29$ ,  $p<.05$ ) als wichtige Voraussetzung.

Die vorläufigen Ergebnisse der längsschnittlichen Feldstudie postulieren zu allen drei Messzeitpunkten einen positiven Zusammenhang zwischen allen Teamlernprozessen und innovativen Arbeitsverhalten (T1: von  $r=.27$  bis  $.51$ ,  $p<.01$ ; T2: von  $r=.27$  bis  $.56$ ,  $p<.01$ ; T3: von  $r=.29$  bis  $.55$ ,  $p<.01$ ) sowie zwischen den Bedingungsfaktoren und Teamlernprozessen (von  $r=.27$  bis  $.77$ ,  $p<.01$ ; T2:  $r=.20$  bis  $.72$ ,  $p<.01$ ; T3:  $r=.23$  bis  $.83$ ,  $p<.01$ ).

Die Ergebnisse postulieren für die Praxis, dass vor allem Teamreflexion und Boundary Spanning in Hinblick auf innovatives Arbeitsverhalten gefördert werden sollte, um so gesellschaftliche Herausforderungen, mit welchen berufliche Schulen konfrontiert sind, zu meistern. Auf der Tagung werden beide Studien mit Fokus auf die Längsschnittstudie (inklusive vollständiger Analysen) vorgestellt.

### Teamlernverhalten und affektives Verhalten von interdisziplinären Lehrenden-Teams an beruflichen Schulen

**Verena Watzek, Regina H. Mulder**

Universität Regensburg, Deutschland; [Verena.Watzek@ur.de](mailto:Verena.Watzek@ur.de)

Professionelles Handeln in der Bildungswissenschaft spielt besonders in der Domäne der beruflichen Bildung eine große Rolle. Insbesondere, die Teamarbeit als Bestandteil professionellen Handelns ist an beruflichen Schulen von großer Bedeutung, da Lehrende an komplexen Herausforderungen, wie der Integration von Flüchtlingen oder dem Qualitätsmanagement gemeinsam arbeiten müssen, um diese meistern zu können (Truijen et al., 2013). Das professionelle Handeln von Lehrenden-Teams in der Domäne der beruflichen Bildung wurde bislang anhand von vereinzelt Studien untersucht (z.B. Bouwmans et al., 2017; Truijen et al., 2013; Zoethout et al., 2017). Jedoch wurden bisher Untersuchungen vernachlässigt, inwiefern das

Teamlernverhalten in Zusammenhang mit dem affektiven Verhalten der Lehrenden-Teams steht und inwiefern sich Veränderungen der beiden Verhaltensweisen über einen längeren Zeitraum feststellen lassen. Affekte stellen eine bedeutende Komponente der Teamarbeit dar, die in Wechselwirkung mit dem Lernverhalten der Teams stehen (Cahour 2013; Van Kleef & Fischer, 2016), da das professionelle Handeln der Lehrenden von Affekten beeinflusst wird. Ziel unserer Studien ist es herauszufinden, inwiefern sich die Lehrenden-Teams in ihrem Lernverhalten unterscheiden sowie, ob sie sich über einen längeren Zeitraum hinweg in ihrem Lernverhalten verändern und ob diese Unterschiede bzw. Veränderungen im Lernverhalten anhand des affektiven Verhaltens der Teams erklärt werden können.

Die beiden Studien, die im Rahmen eines von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekts durchgeführt wurden, beinhalteten ein 'fully integrated mixed design' (Schoonenboom & Johnson, 2017). Die erste Studie bestand aus einer Umfrage von 117 interdisziplinären Lehrenden-Teams und einer vertiefenden Beobachtungsstudie. Um Unterschiede im Lernverhalten der Teams zu untersuchen wurden Clusteranalysen berechnet. Welche Rolle dabei das affektive Verhalten der Teams spielte, wurde sowohl anhand von affektiven Skalen der Fragebogenstudie als auch anhand der Daten der Beobachtungsstudie untersucht. Hierbei wurden sechs der befragten Teams, zwei pro Cluster, ausgewählt und in fünf aufeinanderfolgende Teamsitzungen gefilmt. Das Datenmaterial wurde mit dem Verfahren der qualitativen Inhaltsanalyse (Flick, 2014) ausgewertet im Hinblick auf die Identifikation des affektiven Verhaltens. Die resultierenden drei Cluster unterschieden sich signifikant im Lernverhalten der Teams (z.B. für die Variable Wissensteilung:  $p < .01$ ;  $\eta^2 = .63$ ) und wurden als Lerntypen von Teams festgesetzt: Sehr aktive ‚all-round‘ Teams; Aktive ‚all-round‘ Teams; Aktive Teams fokussiert auf Wissensteilung. Die Unterschiede im Lernverhalten der Teams konnten anhand von (signifikanten) Unterschieden im affektiven Verhalten der Teams erklärt werden. Die Teams, die am häufigsten gemeinsam lernten, erfuhren in ihrer Teamarbeit häufig positive Affekte und selten negative Affekte. Dies zeigte sich sowohl in den Ergebnissen der Umfrage als auch in den Ergebnissen der Beobachtungsstudie. Die zweite, nachfolgende Studie, beinhaltete eine längsschnittliche Fragebogenerhebung über drei Messzeitpunkte sowie weitere qualitative Inhaltsanalysen (Flick, 2014) des Datenmaterials der Beobachtungsstudie mit dem Ziel das Lernverhalten der Teams zu untersuchen, um einen tieferen Einblick in die Veränderungen der Verhaltensweisen geben zu können (Raes et al., 2015; 2016). Es wird untersucht, ob sich Veränderungen des Teamlernverhaltens der Teams identifizieren lassen und ob die Veränderungen anhand des affektiven Verhaltens der Teams erklärt werden können. Vorläufige Ergebnisse zeigen zu allen drei Messzeitpunkten signifikante Korrelationen zwischen den affektiven Verhaltensweisen und den Teamlernverhaltensweisen (z.B. für positive Affekte im Team und Teamlernverhaltensweisen: (T1) von  $r = .21$  bis  $.69$ ,  $p < .01$ ,  $p < .05$ ; (T2) von  $r = .27$  bis  $.74$   $p < .01$ ; (T3) von  $r = .22$  bis  $.67$   $p < .01$ ,  $p < .05$ ). Basierend auf diesen Ergebnissen wird in einem nächsten Schritt anhand eines autoregressiven Modells der wechselseitige Zusammenhang der beiden Variablen sowie deren Veränderung analysiert.

In der Präsentation werden die Ergebnisse der zweiten Studien im Detail erläutert. Die beiden durchgeführten Studien leisten einen Beitrag zum Diskurs des professionellen Handelns als Herausforderung für die Bildungsforschung. Es wurden empirisch fundierte Untersuchungen von Lehrenden-Teams an beruflichen Schulen durchgeführt, welche das professionelle Handeln von Lehrenden sowie deren Entwicklung im Detail behandeln.

## **Berufsfachliche Kompetenzentwicklung sowie erklärungsrelevante Prädiktoren zum Ende des ersten Ausbildungsjahres bei MalerInnen/LackiererInnen**

### **Didem Atik**

Universität Stuttgart Institut für Erziehungswissenschaft (IfE), Deutschland; [atik@bwt.uni-stuttgart.de](mailto:atik@bwt.uni-stuttgart.de)

Im Mittelpunkt des Beitrags steht die Frage, wie sich berufsfachliche Kompetenzen von MalerInnen/LackiererInnen im ersten Ausbildungsjahr entwickeln und welche Prädiktoren dafür bedeutsam sind. Die Forschungslage, in Bezug auf leistungsschwächere Jugendliche im beruflichen Bildungssystem und MalerInnen/LackiererInnen im Besonderen ist bisher unbefriedigend. Studien im Baubereich (Fliesenleger, Stuckateure, Maurer) machen drauf aufmerksam, dass im leistungsschwachen Segment durch ungünstige Voraussetzungen der Jugendlichen mit erheblichen Leistungsproblemen zu rechnen ist (Norwig/ Petsch/ Nickolaus, 2010). Diese Berufsgruppen weisen neben enormen Leistungsdefiziten auch überdurchschnittlich hohe Abbruchquoten auf und bedürfen gezielter Förderung. Eine Herausforderung stellt dabei die große Heterogenität der Gruppe dar. Im Anschluss an vorliegende Studien, die wenig konsistente Ergebnisse zur Entwicklung unterschiedlicher Leistungsgruppen in verschiedenen institutionellen Kontexten und didaktischen Settings erbrachten, soll der Frage nachgegangen werden, wie bei dieser hoch heterogenen Gruppe mit hohen Anteilen Leistungsschwächerer Kompetenzentwicklungen vollziehen. Berichtet werden in den Studien bisher einerseits Matthäuseffekte (Nickolaus/Norwig/Petsch, 2014; Behrendt/Nickolaus/Seeber, 2017), als auch Kompensationseffekte zugunsten der Leistungsschwächeren (Atik/Nickolaus, 2016). Letzteres scheint vor allem in beruflichen Domänen beobachtbar, in welchen Leistungsschwächere hohe Anteile stellen. Bezogen auf die erklärungsrelevanten Prädiktoren zur Kompetenzentwicklung in der gewerblich-technischen Ausbildung erweisen sich domänenübergreifend die kognitiven Eingangsvoraussetzungen als besonders bedeutsam (Nickolaus/Abele/Albus, 2015). Das Berufssegment der MalerInnen/LackiererInnen scheint insoweit von herausragendem Interesse, als das es zu den 20 am stärksten besetzten Ausbildungsberufen gehört, in den primär kognitiv schwächere Jugendliche einmünden. Zudem weist diese Gruppe innerhalb des Handwerks eine überdurchschnittliche Auszubildendenlösungsquote (bundesweit 41,7%, BW 36,5%) auf (BIBB 2015). Hilfreich für die Gewinnung von Ansatzpunkten für pädagogische Handlungsprogramme wären zum einen die längsschnittliche Modellierung der Kompetenzentwicklung sowie die Generierung elaborierter Erklärungsmodelle, die bisher für diesen Bereich noch ausstehen. So sollen im Rahmen der Untersuchung die zentralen Fragestellungen beantwortet werden: (1) Wie verläuft die fachliche Kompetenzentwicklung im ersten Ausbildungsjahr im Berufssegment der MalerInnen und LackiererInnen und lassen sich am Ende der Grundstufe substantielle Kompetenzzuwächse aufzeigen? (2) Lassen sich in diesem Beruf unterschiedliche Leistungsgruppen identifizieren und können im Gegensatz zu Berufen, in welche überwiegend Leistungsstarke einmünden und Matthäuseffekte beobachtet werden, Kompensationseffekte zugunsten der Leistungsschwächeren beobachtet werden. (3) Welche Prädiktoren haben einen Einfluss auf die berufliche Kompetenzentwicklung im Berufsfeld der MalerInnen und LackiererInnen im 1. Ausbildungsjahr?

Die Untersuchung ist längsschnittlich angelegt, erstreckt sich über ein Schuljahr und bezieht ca. 300 Jugendliche der einjährigen Berufsfachschule und des dualen Systems in der Farbtechnik ein. Ergänzend zu den berufsfachlichen Kompetenzen wurden die kognitiven Grundfähigkeit sowie relevante soziokulturelle und motivationale Merkmale erhoben. Für die Analysen der Kompetenzentwicklung wird eine IRT-Rasch Skalierung (Partial-Credit-Modell) vorgestellt. Die zwei Instrumente zur Erfassung der Fachkompetenz werden auf Grundlage des Anderson-Modells mit Ankeritems fixiert und liefern so Informationen über die Entwicklung der Probanden in der Grundstufe. Es wird für die Gesamtgruppe eine Entwicklung mit signifikantem mittleren Effekt ( $d \sim 0.6$ ) bestätigt. Darüber hinaus können bei Betrachtung der Leistungsgruppen deutliche Entwicklungsdynamiken zu Gunsten des unteren Leistungssegments ( $d \sim 1.17$ ) gezeigt werden, während für die

Leistungsstarken eher schwache Effekte ( $d \sim 0.32$ ) dokumentiert werden. Unterschiede zwischen den zwei Leistungsgruppen zu den einzelnen Messzeitpunkten deuten auf eine Homogenisierung der Leistungen in den Gesamtgruppen zum zweiten Messzeitpunkt hin. Im Anschluss an den Erkenntnisstand kann die hohe prädiktive Kraft des Vorwissens bestätigt werden. Weiterhin werden für das berufsfachliche Wissen am Ende des ersten Ausbildungsjahres die kognitive Grundfähigkeit, die Organisationsform, berufliche Aspirationen und die Klassenwiederholung bedeutsam. Ebenfalls Beiträge zur Varianzaufklärung der berufsfachlichen Kompetenz erbringen die Deutschnote und der Migrationshintergrund der Jugendlichen. Sprachliche Defizite in der Schülerschaft im unteren Leistungssegment scheinen für die Kompetenzentwicklung bedeutsam.

## C02

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus S104

### Welche alternativen Wege führen zur Hochschulreife? Eine längsschnittliche Analyse mit den Daten des Nationalen Bildungspanels

**Sandra Buchholz**

Leibniz Institut für Bildungsverläufe, Deutschland; [sandra.buchholz@uni-bamberg.de](mailto:sandra.buchholz@uni-bamberg.de)

Das deutsche Bildungssystem wird in der Regel als sehr rigide bezeichnet. Im Vergleich zu den Bildungssystemen anderer Länder werden Kinder in Deutschland vergleichsweise früh, nämlich nach vier oder sechs gemeinsamen Grundschuljahren, auf verschiedene Schullaufbahnen verteilt. Diese frühe Stratifizierung von Bildungskarrieren wird häufig dafür verantwortlich gemacht, dass Bildungsungleichheiten in Deutschland nach wie vor sehr stark ausgeprägt sind (siehe z.B. Baumert and Schümer, 2001; OECD, 2001, 2004).

Die Mehrheit der Studien zu Ungleichheiten in der deutschen Sekundarbildung konzentriert sich jedoch auf die Untersuchung traditioneller Bildungskarrieren: Entweder wird der frühe Übergang von der Grund- auf die weiterführende Schule in den Mittelpunkt gestellt oder es wird auf Basis von Querschnittsdaten, wie den PISA-Daten, die unterschiedliche Verteilung von Kindern und Jugendlichen auf verschiedene Schulformen in den Blick genommen. Unbestritten ist die erste bzw. frühe Allokation zum Sekundarschulsystem ein wichtiger Meilenstein. Das Problem ist jedoch, dass Bildungskarrieren nicht immer geradlinig verlaufen. Jüngere Arbeiten mit den längsschnittlichen Daten des Nationalen Bildungspanels zeigen, dass das als so rigide geltende deutsche Bildungssystem deutlich offener ist als in der Regel unterstellt: Fast jede dritte Hochschulreife wird in Deutschland erst über einen alternativen Bildungsweg, sozusagen „im zweiten Anlauf“, erreicht (Buchholz und Pratter, 2017). Darüber, wie genau alternative Wege zu höherer Schulbildung aussehen, ist bisher wenig bis gar nichts bekannt.

Deshalb setzt sich dieser Beitrag, meines Wissens erstmals für Deutschland, mit der Frage auseinander, wie alternative Wege zu höherer Sekundarschulbildung typischerweise aussehen, ob sich diesbezüglich soziale Disparitäten abzeichnen und wie sich diese Disparitäten theoretisch erklären lassen. Ich konzentriere mich dabei auf alternative Wege zur Hochschulreife. Grundlage der empirischen Analysen sind die längsschnittlichen Individualdaten des Nationalen Bildungspanels. Um verschiedene Typen alternativer Wege zur Hochschulreife zu identifizieren, werden sequenzanalytische Verfahren angewandt. Auf Basis der in den sequenzanalytischen Modellen identifizierten Cluster wird dann durch Nutzung logistischer Regressionsverfahren und Schätzung durchschnittlicher Marginaleffekte untersucht, ob bzw. inwiefern die Nutzung verschiedener Wege zur Hochschulreife durch individuelle Merkmale (soziale Herkunft, Geschlecht, Noten und den ursprünglichen Bildungsabschluss) beeinflusst wird.

Die Analysen zeigen, dass die Entscheidung, ob Jugendliche nach Abschluss der Haupt- oder Realschule die Hochschulreife nachholen, maßgeblich von ihrer sozialen Herkunft abhängt. Die klaren sozialen Herkunftseffekte lassen sich jedoch, anders als der frühe Übergang von der Grundschule aufs Gymnasium, kaum durch soziale Leistungsdisparitäten erklären. Auch zeigen sich deutliche Geschlechterunterschiede, jedoch anders als beim frühen Übergang auf das Gymnasium zum Nachteil von Mädchen bzw. weiblichen Jugendlichen (insbes. bei schwacher sozialer Herkunft). Dies ist ein Hinweis auf die Wirksamkeit geschlechtsspezifischer Stereotype.

Mit Blick auf die Frage, wie alternative Wege zur Hochschulreife typischerweise aussehen, lassen sich in den sequenzanalytischen Modellen zwei unterschiedliche Cluster identifizieren: Das erste Cluster (mit einer Bedeutung von knapp 67 Prozent) ist gekennzeichnet durch eine nahtlose Fortsetzung der schulischen Bildung; das zweite Cluster (mit einer Verbreitung von knapp 33 Prozent) ist dadurch gekennzeichnet, dass Haupt- und Realschulabsolventen zunächst eine berufliche Ausbildung absolvieren und anschließend die Hochschulreife nachholen. Es zeigen sich deutliche soziale Disparitäten: Vor allem Kinder aus sozial schwachen Familien wählen den „sicheren Umweg“ über eine berufliche Ausbildung; weibliche Jugendliche nutzen, wenn sie sich überhaupt zum Nachholen der Hochschulreife entschließen, durch eine nahtlose Fortsetzung der schulischen Laufbahn den zeitlich weniger intensiven Pfad zur Hochschulreife. Schulische Vorleistungen spielen dagegen keine Rolle. Zusammengefasst deutet dies darauf hin, dass sich soziale Disparitäten in alternativen Wegen zur Hochschulreife vor allem durch entscheidungstheoretische Modelle erklären lassen.

### Der Zusammenhang von Lebenszielen, Zielorientierungen und bildungsbezogenem Erziehungsverhalten – ein personenzentrierter Ansatz

**Sarah Becker<sup>1</sup>, PD Dr. Maximilian Pfost<sup>2</sup>, Prof. Dr. Cordula Artelt<sup>3</sup>**

<sup>1</sup>Universität Bamberg, Bamberg Graduate School of Social Sciences; <sup>2</sup>Universität Bamberg; <sup>3</sup>Universität Bamberg, Leibniz-Institut für Bildungsverläufe, Bamberg; [sarah.becker@uni-bamberg.de](mailto:sarah.becker@uni-bamberg.de)

Insbesondere für Schülerinnen und Schüler kurz vor ihrem schulischen Abschluss, können Ziele als "Navigationswerkzeuge" für einen erfolgreichen Übergang von der Schule zur tertiären Bildung und zur beruflichen Karriere gesehen werden (Nurmi, 2001). Die Zielinhalte können dabei hinsichtlich ihrer Situationsspezifität unterschieden werden. Lebensziele werden als eher situationsübergreifend betrachtet, wohingegen sich Zielorientierungen (i.S.v. Lern- und Leistungszielen) in einem spezifischen Kontext verorten lassen (z.B. Schule). Neben der Frage nach Zielinhalten, beschäftigt sich die Lebenszielforschung mit den Gründen, aufgrund derer ein Mensch ein Ziel erreichen möchte. Deci und Ryan (1985) sprechen in diesem Zusammenhang von Zielregulation und beschreiben den Grad der Autonomie der Gründe, der von extrinsisch (Arzt werden, weil Eltern es verlangen) bis intrinsisch (Arzt werden, weil Arbeit mit Menschen Spaß macht) variieren kann. Die Zielorientierungstheorie beschreibt ebenfalls Gründe, warum Individuen in Lernsituationen agieren. Leistungszielorientierte Menschen handeln, um eigene Fähigkeiten zu demonstrieren oder deren Mangel zu verbergen. Lernzielorientierte Menschen fokussieren auf den Zuwachs der eigenen Kompetenz. Auch hier lassen sich also eher extrinsische und intrinsische Gründe unterscheiden (Gaudreau, 2012; Michou, Vansteenkiste, Mouratidis, & Lens, 2014). Sowohl die Theorie der Lebenszielregulation als auch die der Zielorientierungen zeigen, dass es intrinsische und extrinsische Gründe dafür gibt, warum Menschen in mehr oder weniger spezifischen Situationen handeln. Die Beziehung zwischen beiden Konstrukten und der damit einhergehenden Frage, ob es eine Art übergeordneten Zielregulationstypen gibt, wurden hingegen wenig erforscht. Im vorliegenden Beitrag soll in Fragestellung<sup>1</sup> überprüft werden, ob sich Zusammenhänge zwischen eher intrinsischen bzw. extrinsischen Lebenszielen und Zielorientierungen finden lassen.

Weitere Studien zu Zielorientierungen zeigen, dass elterliches Erziehungsverhalten einen Einfluss auf Ziele der Kinder haben kann (Gonzalez, Doan Holbein, & Quilter, 2002; Wentzel, 1998). Dabei scheint insbesondere die zugestandene Autonomie

Lern- und Leistungsziele zu beeinflussen. Ob bildungsbezogenes Erziehungsverhalten einen Einfluss auf die in Fragestellung1 gefundenen Lebenszielregulationstypen hat, soll im vorliegenden Beitrag in Fragestellung2 untersucht werden.

Als Grundlage dient die Stichprobe der längsschnittlichen BiKS-Studien. Analysiert wurden die Daten von 663 Schülerinnen und Schülern der 12. Jahrgangsstufe und Auszubildenden im 2. Lehrjahr. Die Zielorientierungen wurden mit drei Skalen (Lern-, Annäherungsleistungs- & Vermeidungsleistungsziele) zur Erfassung der Lern- und Leistungsmotivation „SELLMO“ (Spinath, Stiensmeier-Pelster, Schöne, & Dickhäuser, 2002) erfasst. Die Regulation der Lebensziele wurde mit einem idiographisch-nomothetischen Verfahren erhoben. Dafür mussten je drei frei gewählte Lebensziele anschließend auf vier Items (intrinsisch, identifiziert, introjiert, extrinsisch) bewertet werden. Das bildungsbezogene Erziehungsverhalten, gemessen zu mehreren Messzeitpunkten durch Elterninterviews während der Sekundarstufe 1, wurde mit zwei Skalen (zugestandene Autonomie & Leistungserwartungen) in die Analysen integriert.

In Fragestellung1 lassen sich durch latente Klassenanalysen drei Typen von Lebenszielregulationstypen abbilden. Klasse1 (N=47) zeigt insgesamt niedrige Werte auf allen vier Regulationsitems, Klasse2 (N=480) weist besonders hohe Werte bei intrinsischer und identifizierter Regulation auf, Klasse3 (N=137) lässt sich durch hohe Werte bei introjiert und extrinsischer Regulation beschreiben. Um den Zusammenhang mit Zielorientierungen zu zeigen wurden anschließend ANOVAS berechnet. Es konnten keine Zusammenhänge mit Lernzielen gefunden werden. Für die beiden Leistungszielskalen konnte jedoch gezeigt werden, dass Studienteilnehmer mit hohen Werten auch eher in latenten Klassen mit extrinsischer Regulation zu finden sind.

Durch latente Multi-State-Single-Trait-Analysen wurde in Fragestellung2 zunächst gezeigt, dass das Erziehungsverhalten der Eltern über die betrachteten Messzeitpunkte während der Sekundarstufe 1 als stabil betrachtet werden kann. Anhand logistischer Regressionen konnte schließlich gezeigt werden, dass die Wahrscheinlichkeit eine eher intrinsische Lebenszielregulation (Klasse2) zu verfolgen, für die Kinder signifikant höher war, deren Eltern ihnen hohe Autonomie zugestanden haben und weniger stark auf Leistung fokussierten. Geringe zugestandene Autonomie und hohe Leistungserwartungen konnten hingegen eher extrinsische Lebenszielregulation (Klasse3) vorhersagen.

Unsere Analysen bestätigen den erwarteten Zusammenhang zwischen Lebenszielregulation und Zielorientierungen und können einen ersten Hinweis darauf geben, dass übergeordnete Zielregulationstypen definiert werden können. Weiterhin konnte gezeigt werden, dass elterliches Erziehungsverhalten die Lebenszielregulation der Kinder vorhersagen kann.

### **Was sagt mir die Note? Oberstufenreformen und die Einordnung von Fachnoten aus Kursen mit versus ohne Niveaudifferenzierung**

**Wolfgang Wagner<sup>1</sup>, Hübner Nicolas<sup>1</sup>, Hochweber Jan<sup>2</sup>, Neumann Marko<sup>3</sup>, Nagengast Benjamin<sup>1</sup>**

<sup>1</sup>Universität Tübingen, Deutschland; <sup>2</sup>Pädagogische Hochschule St. Gallen (PHSG), Schweiz; <sup>3</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland; [wolfgang.wagner@uni-tuebingen.de](mailto:wolfgang.wagner@uni-tuebingen.de)

Theoretischer Hintergrund. Seit Beginn dieses Jahrhunderts erfolgte in den meisten Bundesländern der Bundesrepublik Deutschland eine Reform der Gymnasialen Oberstufe, die der Möglichkeit individueller Profilierung im Rahmen eines Grund- und Leistungskurssystems durch Einschränkungen der Wahlmöglichkeiten entgegenwirkte, z.B. durch Einführung verbindlicher Kernfächer (Neumann & Trautwein, 2014; Trautwein & Neumann, 2008). Ziel war eine Stärkung der Fächer Deutsch und Mathematik sowie der Bereiche Fremdsprachen und Naturwissenschaften durch obligatorische Belegungen von Kursen mit leistungskursäquivalenten Curricula. Durch die Einführung von obligatorischen Kursbelegungen im gemeinsamen Klassenverband ist allerdings u.a. auch eine Veränderung der Schülerkomposition in den einzelnen Kursen zu erwarten, da sich die typischen Leistungskurschüler von den jeweiligen Grundkurstschülern hinsichtlich ihrer Lernvoraussetzungen unterscheiden (Hübner, Wille et al., 2017; Nagy, Neumann, Trautwein & Lüdtke, 2010). So sind beispielsweise höhere Anteile an leistungsschwächeren Schülerinnen und Schülern im Kernfach verglichen mit dem Leistungskurs zu erwarten. Damit verändert sich die jeweilige Referenzgruppe, die bei der Notenvergabe auf Basis sozialer Bezugsnormen („grading on the curve“) eine wichtige Rolle spielt (Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Auch in der gymnasialen Oberstufe konnten negative Kompositionseffekte der auf Kursebene aggregierten Fachleistung (erfasst anhand standardisierter Leistungstests) auf die Fachnote nachgewiesen werden, die als Referenzgruppeneffekte interpretiert werden können (Neumann, Trautwein, & Nagy, 2011).

Fragestellung. In der vorliegenden Studie wurde für die Fächer Mathematik und Englisch untersucht, ob reformbedingte Veränderungen bezüglich der Kurswahl und der damit einhergehenden Veränderungen der Schülerkomposition zu Unterschieden hinsichtlich der zu erwartenden Fachleistungen bei jeweils identischen Notenstufen vor und nach der Reform führen. Es wurde angenommen, dass Schülerinnen und Schüler bei vergleichbaren Fachleistungen die günstigsten Noten im Grundkurs (mit vergleichsweise niedriger Leistungskomposition) und die am wenigsten günstigen Noten im Leistungskurs (hohe Leistungskomposition) erhalten, während die Kernfachnoten im reformierten System sich zwischen den beiden Kursniveaus vor Systemumstellung einordnen sollten. Entsprechend wurden bei identischen Noten je nach Kursniveau unterschiedliche Ergebnisse in standardisierten Leistungstests erwartet. Außerdem wurden aufgrund der Veränderungen der Kurskomposition unterschiedlich starke Zusammenhänge zwischen der Note und der standardisierten Leistung erwartet.

Methode. Anhand eines Kohorten-Kontroll-Designs (Shadish, Cook & Campbell, 2002) basierend auf zwei repräsentativen Stichproben von Abiturientinnen und Abiturienten aus zwei Bundesländern (Baden-Württemberg: N = 5,574; Thüringen: N = 2,202; NEPS, 2011; Wagner et al., 2011) vor und nach der jeweils weitgehend vergleichbaren organisationalen Oberstufenreform wurden standardisierte Fachleistungen – modelliert als latente Variablen in einem Item Response Theory-Modell – auf Basis von kursniveauspezifischen Schulnoten vorhergesagt. Dabei wurden zur adäquaten Standardfehlerschätzung in Mplus (Muthén & Muthén, 1998-2012) sowohl die Cluster-Struktur der Daten als auch die erforderlichen Surveygewichte berücksichtigt.

Ergebnisse. Es zeigten sich bedeutsame Leistungsunterschiede zwischen den kursspezifischen Leistungen der Schülerinnen und Schüler vor und nach der Reform für identische Schulnoten in beiden Bundesländern. So wiesen beispielsweise Schülerinnen und Schüler im Leistungskurs Mathematik mit der Schulnote „gut“ im Mittel 58,3 Punkte im standardisierten Leistungstest auf, während für Schülerinnen und Schüler im Grundkurs im Mittel nur eine Leistung von 46,4 Punkten vorhergesagt wurde bei einer Standardabweichung von 10 Punkten ( $p < .001$ ). Nach der Reform (im Kernfach) lag die standardisierte Leistung von Schülerinnen und Schülern mit der Schulnote „gut“ hingegen bei 55,3 Punkten und damit statistisch signifikant unter der vorhergesagten Leistung des Leistungskurses ( $p < .001$ ) und statistisch signifikant über der Leistung im Grundkurs ( $p < .001$ ) vor der Reform. Dieses Ergebnismuster zeigte sich in Mathematik in beiden Bundesländern und in Englisch in Baden-Württemberg und unterstreicht die besondere Bedeutung der Referenzgruppe für die Notenvergabe auf Basis sozialer Bezugsnormen. Die Ergebnisse und deren Implikationen für die Bedeutung von Noten am Ende der Sekundarstufe II werden kritisch diskutiert.

## **Bildungsverläufe in Gymnasien: Eine Prädiktionsmodellierung unter Einbezug von Migrations-, Leistungs- und Motivationsmassen**

**Albert Düggeli, Dominique Oesch**

Pädagogische Hochschule FHNW, Basel, Schweiz; [albert.dueggeli@fhnw.ch](mailto:albert.dueggeli@fhnw.ch), [dominique.oesch@fhnw.ch](mailto:dominique.oesch@fhnw.ch)

In der Schweiz sind gymnasiale Ausbildungswege strukturell und curricular über die interkantonale Anerkennung von Maturitätsausweisen (MAR) koordiniert (EDK, 1995). Auch wenn somit eine Art nationale Steuerung vorliegt, und auch Verlaufsdaten auf nationaler Ebene vorliegen (BFS, 2015), erfordert die Analyse gymnasialer Bildungsverläufe den Blick auf kantonale Gegebenheiten, weil beispielsweise Eintritts- bzw. Abschlussquoten, oder die insgesamt an Gymnasien vorgesehenen Ausbildungsdauern kantonal erheblich variieren (vgl. z. B. SKBF, 2014). Besonders in Kantonen, in denen eine vergleichsweise hohe Aufnahmequote mit einem erhöhten Anteil jener verbunden ist, die das Gymnasium vorzeitig und damit ohne Matura verlassen (SKBF, 2014), stellt sich vordringlich die Frage, weshalb die gymnasialen Bildungswege nicht zu Ende geführt werden. Stamm (2010) verweist diesbezüglich darauf, dass neben erzielten Noten (Mathematik) auch motivationale Aspekte mit Repetitionen oder Abbrüchen assoziiert sind, und dass zwischenmenschliche Probleme in Form von Konflikten mit Lehrpersonen oder Mitslernenden diesbezüglich ebenfalls einen Einfluss haben. Neuere Befunde zur Erklärung nicht-linearer oder so genannter Abbruchverläufe fehlen jedoch. So ist beispielsweise das Wissen zu bildungsbiografischen Verläufen von Lernenden, die in der Schweiz geboren sind, und solchen, die erst nach ihrer Geburt in die Schweiz eingewandert sind, noch wenig ausdifferenziert. Ähnlich sind motivationale Aspekte unter fachlicher Perspektive nur wenig geschärft, beispielsweise mit Blick auf fachspezifische Selbstkonzepte bzw. fachbezogene Anstrengungsbereitschaften. Und ganz besonders fehlt es an Prädiktionsmodellen, die bereits zu einem frühen Zeitpunkt gymnasialer Laufbahnen einen möglichen Einfluss dieser genannten Merkmale aufzeigen. Hier setzt die Fragestellung dieses Beitrags an. Untersucht wird, inwiefern sich gymnasiale Repetitions- oder Drop-Out-Verläufe im Vergleich zu ungebrochenen Biografien durch die Noten, die drei Jahre vor dem Abschluss erzielt wurden (Leistungsaspekt), durch den Migrationsstatus (Aspekte der familiären Herkunft) oder die Leistungsbereitschaft (motivationaler Aspekt) erklären lassen. Den Analysen liegen Daten aus der TIDES-Durchlässigkeitsstudie zugrunde, die mittlerweile längsschnittliche Untersuchungen zu gymnasialen Bildungsverläufen in Basel-Stadt ermöglicht (N=335). Erste Befunde weisen darauf hin, dass sich Austritte durch geringere Noten in den Fächern Deutsch, Mathematik und Englisch erklären lassen. Bei Repetitionen jedoch fällt besonders das Fach Mathematik ins Gewicht. Motivationale Merkmale scheinen sich insgesamt nur marginal auszuwirken, während sich der Migrationsstatus für Lernende der ersten Einwanderergeneration als Bedeutsam erweist. Für sie, nicht aber für Lernende der zweiten Generation, ist die Wahrscheinlichkeit eines Abbruchs oder einer Repetition im Vergleich zu jenen, die einen Normverlauf aufweisen, systematisch erhöht. Diese Befunde werden im Kontext chancengerechter Ausbildung, bildungssystemischer Steuerung und individueller Förderung diskutiert, und zwar mit besonderem Blick auf das in Basel auf der Basis der vorliegenden Ergebnisse neu gestartete Projekt ChaBäle. Es ist eine Förderintervention für jene Jugendlichen, die bei der Gestaltung ihrer gymnasialen Bildungskarriere vor allem migrationsbedingte Herausforderungen zu meistern haben.

## C03

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus S105

### **Gesundheitsförderung durch Supervision – Eine Studie zur Evaluation von Einzelsupervision für VolksschullehrerInnen**

**Christina Mogg**

Universität Wien, Österreich; [christina.mogg@univie.ac.at](mailto:christina.mogg@univie.ac.at)

Der Lehrberuf geht mit hohen physischen und psychischen Anforderungen einher, die die Gesundheit der Lehrpersonen potentiell beeinträchtigen können. In einer vorhergehenden qualitativ-empirischen Studie der Autorin über Maßnahmen zur Gesundheitsförderung und Prävention im Lehrberuf stellte sich die Einzelsupervision als besonders wichtige Maßnahme heraus. Auch in der Literatur gilt Supervision als bedeutsame Maßnahme, bei der Arbeitsprozesse, Beziehungskonstellationen und Probleme des schulischen Alltags reflektiert, Lösungsstrategien erarbeitet und dadurch langfristig Veränderungen im Arbeitsumfeld erreicht werden können (Günther, 2014; Krause, 2013; Thiel, 2013). Empirische Ergebnisse zur Erforschung der Wirksamkeit von Einzelsupervision für LehrerInnen liegen im deutschsprachigen Raum zum aktuellen Stand nicht vor.

In dieser Interventionsstudie wurde die Wirkung einer einsemestrigen Einzelsupervision für VolksschullehrerInnen in Wien evaluiert, um die Effektivität dieser Maßnahme der Gesundheitsförderung zu untersuchen und Empfehlungen von ExpertInnen über die optimale Organisation von Supervision zusammenzufassen. Dabei sollte die Forschungsfrage beantwortet werden, inwiefern sich der Gesundheitszustand, die Lebensqualität und Arbeitsfähigkeit von LehrerInnen durch die Inanspruchnahme von Supervision verändern, und wie diese gestaltet sein sollte. Es wurde ein Methodenmix aus quantitativen und qualitativen Methoden eingesetzt. Mittels quantitativer Längsschnitterhebung zu zwei Messzeitpunkten jeweils vor und nach der Intervention wurde analysiert, ob und inwieweit sich der Zustand der supervidierten LehrerInnen (n=25) im Vergleich zur Kontrollgruppe (n=35) verändert. Der eingesetzte Fragebogen enthält Messinstrumente zur Erhebung von Gesundheitszustand, Lebensqualität, Arbeitsfähigkeit, Arbeitsbelastung, Erschöpfungsempfinden, Distanzierungsfähigkeit und Selbstwirksamkeit. Nach der Intervention fand eine Fokusgruppendifkussion mit fünf SupervisorInnen statt, um den Nutzen der Supervision aus ExpertInnensicht einschätzen sowie die inhaltlichen Schwerpunkte und optimale Gestaltung der Einheiten feststellen zu können.

Die statistische Datenanalyse der Veränderung relevanter Merkmale im Zeitverlauf zeigt, dass sich die Variablen zur Messung von Arbeitsfähigkeit und Arbeitsbelastung, physischem und psychischem Gesundheitszustand, Gesundheitszufriedenheit, Distanzierungsfähigkeit, Selbstwirksamkeit und Emotionale Erschöpfung bei der supervidierten Gruppe im Vergleich zur Kontrollgruppe positiver verändern. Mit Ausnahme der beiden Variablen zur Erfassung von Lebenszufriedenheit und Glücksempfinden haben sich die Mittelwerte der Versuchsgruppe im Vergleich zu jenen der Kontrollgruppe durchwegs positiver verändert, auch wenn die Unterschiede nicht signifikant ausfallen ( $p < 0,05$ ). In der Fokusgruppendifkussion wurden ausschließlich positive Aspekte hervorgebracht und diskutiert. Die SupervisorInnen konnten keine Nachteile finden, die sich dadurch ergeben könnten, sofern der Grundsatz der absoluten Vertraulichkeit der besprochenen Inhalte durch die SupervisorInnen eingehalten wird. Die empirischen Ergebnisse legen den Schluss nahe, dass Einzelsupervision als Maßnahme der Gesundheitsförderung den einzelnen LehrerInnen sehr hilft und daher für alle LehrerInnen angeboten werden sollte.

### **Berufliche Selbstregulation von Lehrkräften: Überprüfung eines Kurzinstrumentes**

**Claudia Menge, Hildegard Schaeper**

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (DZHW), Deutschland; [menge@dzhw.eu](mailto:menge@dzhw.eu), [schaeper@dzhw.eu](mailto:schaeper@dzhw.eu)

#### Theoretischer Hintergrund

Berufliche Selbstregulation beschreibt die Fähigkeit, im beruflichen Kontext verantwortungsvoll mit den persönlichen Ressourcen zu haushalten (vgl. Klusmann 2011). Diese Fähigkeit spielt als Element professioneller Handlungskompetenz eine zentrale Rolle bei der Bewältigung beruflicher Anforderungen (vgl. Klusmann et al. 2006; Roloff Henoch et al. 2015; Oesterreich 2015). Sie wird im Modell professioneller Kompetenz (angehender) Lehrkräfte von Baumert und Kunter (2006) als einer von vier zentralen Kompetenzaspekten unterschieden.

Einschlägige Studien zum Umgang mit beruflichen Anforderungen haben aufgezeigt, dass Lehrkräfte im Vergleich psychosozial beanspruchter Berufsgruppen – darunter z. B. Polizisten, Pflegekräfte und Beschäftigte bei der Feuerwehr – besonders häufig Erlebens- und Verhaltensstile aufweisen, die für ihre psychische und physische Gesundheit riskant sind (vgl. Schaarschmidt und Kieschke 2007a; Schaarschmidt und Kieschke 2007b). Mögliche Auswirkungen der risikobehafteten Bewältigungsstile sind beispielsweise eine erhöhte emotionale Erschöpfung, gesteigerte Berufsunzufriedenheit, geringere berufliche Produktivität und Effektivität, eine höhere Anzahl an (krankheitsbedingten) Fehltagen und dauerhafte Dienstunfähigkeit (vgl. Klusmann et al. 2006; Neugebauer und Wilbert 2010).

#### Fragestellung

In diesem Beitrag soll eine Kurzskaala des Nationalen Bildungspanels (NEPS) zur beruflichen Selbstregulation von Lehrkräften überprüft werden. Grundlage der Kurzskaala ist das AVEM-Inventar, welches von Schaarschmidt und Fischer (2001) im Rahmen der Potsdamer Lehrerstudie entwickelt wurde und seither regelmäßig in Untersuchungen zum Lehrberuf, insbesondere zur Identifikation und Analyse von Burnout-Risiken, Verwendung findet. Dieses Instrument erfasst anhand von elf Dimensionen zu jeweils sechs (Vollversion) bzw. vier Items (Kurzversion) arbeitsbezogene Verhaltens- und Erlebensmuster auf Basis von Merkmalen des beruflichen Engagements (Ressourceninvestition), der Widerstandsfähigkeit (Ressourcenerhaltung) sowie der berufsbezogenen Emotionen. Im Rahmen des NEPS werden wegen der begrenzten Befragungszeit nur jeweils zwei Dimensionen des beruflichen Engagements und der Widerstandsfähigkeit mit drei bis vier Items (insgesamt 13 Items) erfasst.

Die besondere Bedeutung des AVEM-Inventars ergibt sich aus der Möglichkeit, aus den Antwortmustern clusteranalytisch vier Typen beruflicher Selbstregulation abzuleiten, die prädiktiv für die Gesundheit, den Verbleib im Beruf sowie die Qualität der Berufsausübung von Lehrkräften sind (vgl. Baumert und Kunter 2006). Ob sich diese Typologie mithilfe der 13 im NEPS eingesetzten Items reproduzieren lässt, oder sich ähnlich prädiktive Muster unterscheiden lassen, steht neben der Überprüfung der faktoriellen Struktur und diskriminanten Validität der Kurzskaala im Zentrum des vorliegenden Beitrags.

#### Daten und Methoden

Die AVEM-Kurzskaala wurde im Nationalen Bildungspanel (NEPS) in der Lehramtsstichprobe der Studierendenkohorte eingesetzt. Die Analysen stützen sich auf die Angaben von 1.059 Personen, die sich zum Erhebungszeitpunkt Ende 2016 in

einem Referendariat befanden oder als Lehrkraft erwerbstätig waren. Die faktorielle Struktur und die diskriminante Validität wurden mittels konfirmatorischer Faktorenanalysen überprüft. Für die Identifizierung von Selbstregulationstypen wurden latente Profilanalysen durchgeführt. Weitere Analysen der gefundenen Profillösung erfolgten mit multinomialen Logitmodellen.

#### Ergebnisse

Die eingesetzten Subskalen verfügen mit  $\alpha \geq 0,77$  über eine zufriedenstellende interne Konsistenz, die mit anderen Studien zum Belastungserleben von Lehrkräften vergleichbar ist (z. B. Schröder und Kieschke 2006; Klusmann et al. 2008; Beer und Benischek 2014; Roloff Henoch et al. 2015; Oesterreich 2015). Die Faktorenstruktur wurde in konfirmatorischen Faktorenanalysen bestätigt, wobei auch ein Modell mit zwei Faktoren höherer Ordnung geschätzt werden konnte (RMSEA = 0,066; CFI = 0,944; TLI = 0,929). Zusammenhänge mit Konstrukten wie dem globalen Selbstwertgefühl und den Big Five-Persönlichkeitsmerkmalen sprechen für die diskriminante Validität des Instruments. Mit dem Vorliegen von Daten aus nachfolgenden Wellen wird darüber hinaus die prädiktive Validität zu überprüfen sein.

Durch latente Profilanalysen wurde die Unterscheidung von vier Mustern beruflicher Selbstregulation nahegelegt. Diese lassen sich inhaltlich analog zur Typologie nach Schaarschmidt interpretieren. Analysen mit ausgewählten Prädiktoren der Regulationsmuster haben weitreichende Übereinstimmungen mit den Befunden anderer Untersuchungen ergeben.

Bewährt sich die Kurzskaala auch in den ausstehenden Analysen, so liegt ein valides und reliables Kurzinstrument vor, das bei begrenzter Befragungszeit gut für Analysen zum Lehrerberuf geeignet ist. Diese umrissenen Ergebnisse werden präsentiert und ihre Implikationen diskutiert.

### **Anforderungsspezifische Unterschiede in der Resilienz von Lehrkräften**

**Udo Käser, Elisabeth Müller, Julia Maisenbacher, Janine Uklei**

Rheinische Friedrich-Wilhelms-Universität Bonn, Deutschland; [ukaeser@uni-bonn.de](mailto:ukaeser@uni-bonn.de)

Eine genaue Bestimmung des Begriffs „Resilienz“ ist schwierig. Einheitlich liegt ein Verständnis von Resilienz als Widerstandsfähigkeit vor, die daran deutlich wird, dass ein Organismus dazu in der Lage ist bei erlebter Belastung in den unbelasteten Normalzustand zurückzufedern (Turner, 2002). Darüber hinaus liegen unterschiedliche Konzepte von Resilienz vor. Während Block und Kremen (1996) Resilienz ursprünglich als relativ stabiles Persönlichkeitsmerkmal auffassen, wird Resilienz heute eher als Fähigkeit verstanden, die sich im Kontext von Person-Umwelt-Interaktionen situativ entwickelt. Jessor (1993) bezeichnet hiermit ein relationales Konstrukt, das sich erst in spezifischen Konstellationen beeinträchtigter Ressourcen zeigt. Kumpfer (1999) fasst Resilienz als Prozess oder Ergebnis biopsychosozialer Anpassung an widrige Lebensumstände auf. Insofern stehen unterschiedliche Auffassungen einander gegenüber, bei denen Resilienz entweder als Persönlichkeitsmerkmal ins Individuelle verwiesen wird, es einen Adaptionsprozess zwischen Individuum und seinem mikrosozialen Umfeld bezeichnet oder das Ergebnis eines solchen Prozesses ist (Maasten & Reed, 2002). Ein dynamisches im Unterschied zu einem statischen Verständnis von Resilienz führt hierbei dazu, dass Resilienz nicht zeitstabil ist, sondern Organismen zu einem Zeitpunkt resilient sein können und zu einem anderen Zeitpunkt nicht. Resilienz wird dann auch nicht als globale Größe, sondern bereichsspezifisch verstanden. Wustmann (2004) spricht daher von Resilienz als Widerstandsfähigkeit im Sinne eines variablen und situationspezifischen Konstrukts, welches sich dynamisch entwickelt.

Folgt man diesem Verständnis von Resilienz, so ergibt sich mit Bezug auf das Belastungserleben von Lehrkräften die Frage, inwieweit nicht auch hier eine Differenzierung nach spezifischen Anforderungen des Lehrerberufs erforderlich ist. Diese Anforderungen können mit Blick auf den schulgesetzlichen Rahmen der Tätigkeit von Lehrkräften leicht konkretisiert werden: Lehrkräfte haben die Aufgabe, Schülerinnen und Schüler zu unterrichten, zu erziehen, zu beraten, zu beurteilen, zu betreuen und umfassend zu fördern, bei der Gestaltung des Schullebens und der Fortentwicklung der Qualität schulischer Arbeit mitzuwirken sowie sich fortzubilden (Scharenberg & Käser, 2015). Daher wird hier der Frage nachgegangen, ob bei der Resilienz von Lehrkräften von einem Globalkonstrukt ausgegangen werden kann oder ob es sinnvoll ist, anforderungsspezifisch von Bereichen zu sprechen, in denen die Resilienz einer Lehrkraft unterschiedlich ausfällt. Außerdem wird untersucht, inwieweit sich für die unterschiedlichen Anforderungsbereiche Belastungserleben und Resilienz unterscheiden.

Hierzu wurde ein online-Fragebogen entwickelt, in dem in 58 Items mögliche Belastungssituationen des Alltags von Lehrkräften geschildert werden, für die eingeschätzt werden muss, wie häufig solche Situationen auftreten und wie nachhaltig eine mögliche Beanspruchung wirkt. Der Fragebogen wurde an einer Stichprobe von 331 Lehrkräften realisiert, 211 Lehrerinnen (64%) und 120 Lehrer (36%). Ihr Alter variierte zwischen 25 und 68 Jahren ( $M = 47$  Jahre,  $SD = 10$  Jahre). Die Berufserfahrung lag zwischen einem und 42 Jahren ( $M = 17$  Jahre,  $SD = 11$  Jahre). Die Lehrkräfte unterrichteten an Gymnasien (36%), Berufskollegs (31%), Gesamtschulen (19%), Realschulen (10%) oder Hauptschulen (5%). Es bestätigten sich neun Faktoren: Unterrichten, Umgang mit Störungen / schwierigen Schülerinnen und Schülern, Unterrichtskonzeption, Beurteilen und Beraten, Diagnose und Förderung, Aufsichtspflichten, Elternarbeit, Verwaltung und Organisation sowie fehlende Wertschätzung. Für die Reliabilitätskoeffizienten ergaben sich akzeptable bis gute Werte zwischen ungefähr 70 und 90 Prozent. Zusätzlich wurde neben Merkmalen der Person (Alter, Geschlecht, Berufserfahrung) und der beruflichen Tätigkeit (Schulform, Standort der Schule) noch die berufliche Selbstwirksamkeit als psychisches Merkmal nach Käser und Wasch (2009) erfasst (16 Items mit Cronbachs Alpha von 83%).

Dabei zeigen sich sowohl für das Belastungserleben als auch für die Nachhaltigkeit der Belastung mehrheitlich hohe Interkorrelationen zwischen den jeweils neun Faktoren der verschiedenen Anforderungsbereiche. Die Häufigkeit, mit der Belastungen in den verschiedenen Bereichen erlebt werden, und die Nachhaltigkeit, die der Beanspruchung zukommt, unterscheiden sich aber je nach Anforderungsbereich deutlich. Auch liegen je nach Anforderungsbereich deutliche unterschiedliche Zusammenhänge zu den Personen- und Schulmerkmalen vor. Dies spricht dafür, dass eine anforderungsspezifische Differenzierung von Resilienz gerade hinsichtlich möglicher Interventionsmaßnahmen sinnvoll sein kann.

### **‘Erfolg’ und ‘Scheitern’ in (berufs-)biographischen Selbstkonstruktionen von Lehrpersonen**

**Susanne Keller**

ZQF, Universität Zürich, Schweiz; [susanne.keller0@gmail.com](mailto:susanne.keller0@gmail.com)

Das Projekt untersuchte (berufs-)biographisch-narrative Interviews mit Lehrpersonen, die zum Interviewzeitpunkt erschöpfungsbedingt krankgeschrieben waren wie auch mit Lehrpersonen, die im aktiven Schuldienst standen und sich als zufrieden in ihrem Beruf ansahen. In den Interviews mit den krankgeschriebenen Lehrpersonen wurde deutlich, dass ein Burnout für Lehrpersonen ein stigmatisierendes lebensgeschichtliches Ereignis (Goffman 1975) darstellt – sie laufen Gefahr,



als gescheitert oder nicht geeignet für den Lehrberuf angesehen zu werden -, das ihre professionelle Identität massiv irritiert und aufwändig erklärt und in die Lebens- und/oder Berufsgeschichte integriert werden muss (Mead 1973, Strauss 1974). Entsprechend lassen sich in den Interviews mit den krankgeschriebenen Lehrpersonen intensive und ausgedehnte, theoretisierende Aktivitäten ausmachen, die darauf gerichtet sind, sich selbst und dem/der Interviewenden das Entstehen ihrer beruflichen Krise plausibel zu machen. Doch nicht nur die krankgeschriebenen Lehrpersonen leisten intensiv biographische Arbeit (Corbin/Strauss 2004), auch in den Interviews mit den berufszufriedenen Lehrpersonen finden sich umfangreiche und teilweise sehr elaborierte Erklärungen für das Ausbleiben beruflicher Schwierigkeiten oder eines «Burnout». Vor diesem Hintergrund entwickelte sich die Fragestellung, wie Lehrpersonen ihren Berufsverlauf erleben und theoretisch verarbeiten und wie sie ein lebensgeschichtlich stigmatisierendes Ereignis, im Falle der krankgeschriebenen Lehrpersonen ein «Burnout», in ihre (Berufs-)Biographie integrieren. Für die Beantwortung der Fragestellung interessieren vor allem die theoretischen Anstrengungen der Lehrpersonen, die ihren Berufsverlauf erklären sollen, um jedoch zu verstehen, wie eine Lehrperson ihren Berufsverlauf deutet und interpretiert, ist es notwendig zuvor die Lebensereignisse, die sich faktisch ereignet haben, und die biographische Aufschichtung der Erfahrung zu rekonstruieren (Schütze 1984). Die Interviews wurden deshalb mittels des narrationsanalytischen Verfahrens (Schütze 2008a, 2008b) analysiert, um die biographische Erfahrungsaufschichtung und die darin zum Ausdruck kommenden biographischen Prozesse (ebd.) zu rekonstruieren. In einem anschließenden Analyseschritt wurden die theoretisierenden Aktivitäten der Lehrpersonen argumentationsanalytisch untersucht (Schütze 1978, 1987; Riemann 1986).

Ein Ergebnis der Untersuchung sind Rekonstruktionen der berufsbiographischen Erzählungen von sechs Lehrpersonen, die jeweils sehr unterschiedliche Verläufe aufweisen. Für den Tagungsbeitrag werden daraus zwei Fallbeispiele ausgewählt und zusammengefasst kontrastiert: 1) Die Sekundarlehrerin Iris Alder durchleidet einen Enthausungsprozess, als sie nach einer langanhaltenden Ehekrise aus dem gemeinsamen Haus ausziehen und ihre Kinder beim Ehemann lassen muss. Infolge Disziplinschwierigkeiten und professioneller Fehler gerät sie in einen seitens der Eltern initiierten Mobbingprozess. Zum Zeitpunkt des Interviews gelingt es Frau Alder noch nicht, ihre beruflichen Schwierigkeiten aus sozialwissenschaftlicher Perspektive zu betrachten und zu verstehen, dass ihre Reformschulorientierung in einer Standardschule mit ihrer Schulkultur wie auch Schüler- und Elternkultur zu einer Quelle immer neuer Schwierigkeiten werden kann. Stattdessen werden ihre Schwierigkeiten heruntergebrochen auf psychologische Erklärungen («ADS», «Krise»). Ihre Verwendung quasi-professioneller Prädikate verweist auf Diskurse, in denen es selbstverständlich oder zum Diktum geworden ist, den Blick auf sich selbst zu richten und berufliche Schwierigkeiten mit psychologischen Defiziten zu erklären. 2) Der Generalist Herr Schnegg: Nach Beendigung des Lehrerseminars entdeckt Herr Schnegg im jahrgangsübergreifenden Unterricht, in der Kombination von verschiedenen Altersstufen und Disziplinen eine wichtige Ressource zur Herausbildung einer beruflichen Identität und Sinnquelle für seinen Beruf. Infolge schrumpfender Schülerzahlen kommt es an Herrn Schneggs Schule zu Umstrukturierungsmassnahmen, die eine Zusammenlegung von Klassen zur Folge haben. Mit der Klassenzusammenlegung beginnen die beruflichen Schwierigkeiten

Herrn Schneggs, für die ihm von Eltern und Kolleg/innen die Verantwortung zugeschrieben werde. Er beobachtet an sich Somatisierungsphänomene und beschliesst, aus dem Lehrberuf auszusteigen. In seiner Argumentation bezieht sich Herr Schnegg auf gesellschaftliche Wandlungsprozesse, die sich, von ihm unbemerkt, ereignet und seine Energieressourcen erschöpft hätten. Er begibt sich damit auf die Ebene eines Diskurses über gesellschaftliche Verfallserscheinungen im Zuge von Individualisierungs- und Modernisierungsprozessen, der zur Erklärung seines Erschöpfungssyndroms und einer Entfremdung vom Beruf herangezogen wird.

Im Tagungsbeitrag werden, als ein weiteres Ergebnis der Untersuchung, die Theorien der Lehrpersonen über den Misserfolg ihrer Tätigkeiten theoretisch integriert.

## C04

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Rosshofgasse, S02

### **Studienabbruch im Lehramtsstudium: Bedeutung von sozialer Integration für die Studienabbruchintention von Lehramtsstudierenden mit Migrationshintergrund**

**Kristin Wolf, Christina Maurer, Mareike Kunter**

Goethe-Universität Frankfurt, Deutschland; [wolf@paed.psych.uni-frankfurt.de](mailto:wolf@paed.psych.uni-frankfurt.de)

#### Theoretischer Hintergrund

Beim Studienabbruch handelt es sich um einen komplexen Entscheidungsprozess, der von vielen inneren und äußeren Faktoren beeinflusst wird (Heublein et al., 2017). Als mögliche Ursachen beschreiben bereits Tinto (1975) und Spady (1970) mangelnde soziale und akademische Integration und auch neuere theoretische Modelle berücksichtigen soziale Integration als relevanten Einflussfaktor (Heublein et al., 2017). Empirische Forschungsergebnisse bestätigen dies und zeigen signifikante Zusammenhänge zwischen sozialer Integration und Studienabbruch bzw. Verbleib im Studium (Ebert & Heublein, 2017; Heublein et al., 2017; Robbins et al., 2004). Besonders im Lehramtsstudium kann soziale Integration aufgrund der Rahmenbedingungen erschwert werden: Lehramtsstudierende studieren meist mehrere Fächer und besuchen zusätzlich zu den Fachwissenschaften fachdidaktische und bildungswissenschaftliche Veranstaltungen. Sie müssen sich daher in wechselnde Lerngruppen integrieren und studieren nicht gemeinsam als Kohorte. Vor diesem Hintergrund stellen Lehramtsstudierende eine besonders interessante Stichprobe dar.

Unabhängig vom Studiengang zeigen sich für Studierende mit Migrationshintergrund vergleichsweise hohe Studienabbruchraten (Bundesamt für Migration und Flüchtlinge, 2011; Ebert & Heublein, 2017; Heublein et al., 2017). Forschungsergebnisse, die Zusammenhänge zwischen sozialer Integration und Migrationshintergrund untersuchen, sind widersprüchlich: während Studierende mit Migrationshintergrund in qualitativen Interviewstudien häufig geringeres Zugehörigkeitsgefühl berichten (Meinhardt & Zittlau, 2009; Naumann, 2011), können quantitativen Untersuchungen dies bisher nicht bestätigen (Hagerty et al., 1996; Hausmann et al., 2007).

#### Fragestellung

Der vorliegende Beitrag setzt hier an und hat das Ziel die Bedeutung wahrgenommener sozialer Integration für die Studienabbruchintention von Lehramtsstudierenden mit Migrationshintergrund zu untersuchen. Es wird angenommen, dass Lehramtsstudierenden mit Migrationshintergrund eine geringere soziale Integration aufweisen und diese als Mediator des Zusammenhangs von Migrationshintergrund und Studienabbruchintention fungiert.

#### Methode

Es wurden 796 Lehramtsstudierende der Universität Duisburg-Essen (n=243), der Goethe-Universität Frankfurt (n=386), der Technischen Universität München (n=122) und der Eberhard Karls Universität Tübingen (n=45) in einer Online-Studie befragt. Die Studierenden (71% weiblich) waren durchschnittlich 22.54 Jahre alt (SD=3.54, Range: 17-46).

Zur Bildung der unabhängigen Variable Migrationshintergrund wurden Angaben zum Geburtsland der Studierenden sowie deren Eltern berücksichtigt. Als Studierende mit Migrationshintergrund werden solche Studierenden verstanden, die selbst oder von denen mindestens ein Elternteil im Ausland geboren wurden. Als abhängige Variable wurde die Studienabbruchintention („Denken Sie momentan daran Ihr Lehramtsstudium abzubrechen?“) mit einem zweistufigen Antwortformat (1=ja, 2=nein) erfasst. Die wahrgenommene soziale Integration der Studierenden wurde mit einer angepassten, deutschen Version einer Subskala der Student Engagement Scale aus PISA 2000 (OECD, 2002) erfasst, jeweils für die universitären Veranstaltungen in den Bildungswissenschaften ( $\omega=.83$ ) und Fachwissenschaften ( $\omega=.86$ ) getrennt. Um die Fragestellung zu beantworten, wurde ein Strukturgleichungsmodell analysiert, wobei fehlende Werte (max. 4% pro Variable) mithilfe des Full Information Maximum Likelihood-Verfahrens (Muthén & Muthén, 1998-2010) berücksichtigt wurden. Als Kovariaten wurden Muttersprache, Bildungshintergrund, Abiturnote, Fachkombination, Alter, Geschlecht und Standort berücksichtigt.

#### Ergebnisse

Die Ergebnisse zeigen erwartungsgemäß direkte Effekte des Migrationshintergrundes auf die wahrgenommene soziale Integration sowohl in den Bildungswissenschaften ( $b=-.26$ ,  $p<.05$ ) als auch in den Fachwissenschaften ( $b=-.29$ ,  $p<.05$ ). Studierende mit Migrationshintergrund fühlen sich demnach sowohl in bildungswissenschaftlichen als auch in fachwissenschaftlichen Veranstaltungen weniger sozial integriert als Studierende ohne Migrationshintergrund. Zur Vorhersage der Studienabbruchintention erweisen sich die wahrgenommene Integration in den Fachwissenschaften ( $b=.45$ ,  $p<.05$ ) sowie das Alter der Studierenden ( $b=.13$ ,  $p<.05$ ) als signifikante direkte Prädiktoren. Studierende, die sich in den fachwissenschaftlichen Veranstaltungen integrierter fühlen, denken weniger an Studienabbruch. Dies gilt ebenso für ältere Studierende. Für soziale Integration in den Bildungswissenschaften zeigt sich weder ein direkter Effekt auf Studienabbruchintention noch indirekte Effekte. Wohingegen sich die wahrgenommene Integration in den Fachwissenschaften erwartungsgemäß als signifikanter Mediator des Zusammenhangs zwischen Migrationshintergrund und Studienabbruchintention erweist ( $b=-.13$ ,  $p<.05$ ). Die Ergebnisse zeigen, dass die Studienabbruchintention von Lehramtsstudierenden mit Migrationshintergrund durch geringe soziale Integration in den Fachwissenschaften erklärt werden kann, dies zeigt sich nicht für soziale Integration in bildungswissenschaftlichen Veranstaltungen.

### **Wissen über Gründe für problematische Studienverläufe – eine Grundlage für den professionalisierten Umgang mit Studienabbruchneigung an der TU Dresden?!**

**Franziska Schulze-Stocker, Robert Pelz**

TU Dresden, Deutschland; [franziska.schulze-stocker@tu-dresden.de](mailto:franziska.schulze-stocker@tu-dresden.de)

Administrative Vorgaben wie bspw. die Bologna-Reform, die Exzellenzinitiative oder der Hochschulpakt haben zu Neuerungen im deutschen Bildungssystem geführt, die auch die Gestaltung universitärer Bildungsprozesse betreffen. Die hohen Abbruchzahlen werden in diesem Zusammenhang umfassend diskutiert (für Bachelorstudiengänge: Heublein, 2014). Im Mittelpunkt stehen verschiedene den Studienabbruch bedingende Faktoren, die aus unterschiedlichen theoretischen Ansätzen (Interaktionismus, kulturelles Kapital, rationale Wahl, psychologische Ansätze) hergeleitet werden können (Ahles et al., 2016). Zudem werden verschiedene Präventions- und Interventionsmaßnahmen für den universitären Bereich diskutiert (z.B. Ahles et

al. 2016; Blüthmann et al., 2008), u.a. persönliche Studienberatung, Orientierungsangebote (z.B. Einführungskurse), Teilzeitstudium bzw. Flexibilisierung des Studiums oder Etablierung von Frühwarnsystemen.

Für die Konzeption und Durchführung passgenauer und professioneller Beratungsangebote für abbruchgefährdete Studierende fehlen Informationen zu kritischen Studienverläufen. Mittels sog. Frühwarnsysteme, welche derzeit an deutschen Hochschulen erprobt werden, lassen sich diese anhand vorab bestimmter Merkmale identifizieren. Die Gründe für Probleme im Studium bleiben dabei allerdings ungeklärt.

Der Beitrag stellt dar, wie die Erforschung der Gründe problematischer Studienverläufe mithilfe einer Methodentriangulation (quantitative Befragung mit offenen und geschlossenen Fragen sowie qualitative Interviews) in Handlungsempfehlungen für Studienberatung und Frühwarnsysteme münden kann. Zentrale Frage hierbei ist, welche Gründe hinter problematischen Studienverläufen stehen.

In einer Auftakterhebung wurden 463 Studierende ausgewählter Studiengänge der TU Dresden zu ihrem Studium befragt (51,9% Frauen, Alter: M = 22 Jahre). Insgesamt liegen ca. 1.400 selbstformulierte Angaben bezüglich der Frage nach Gründen kritischer Studienverläufe vor, die über die unzureichende Anzahl an Leistungspunkten sowie eine bevorstehende zweite Wiederholungsprüfung operationalisiert wurden. Diese Angaben wurden anhand eines umfassenden Kategoriensystems qualitativ-inhaltsanalytisch ausgewertet (Mayring, 2010; Cohen's  $\kappa = .822$ ). Dieses spiegelt das Spektrum der Gründe für problematische Studienverläufe aus Studierendensicht wider. Erste Ergebnisse belegen einen Zusammenhang kritischer Studienverläufe mit Gedanken an Studienabbruch. Vertiefende Analysen zu unterschiedlichen Studierendengruppen, die sich mit verschiedenen Problemszenarien konfrontiert sehen, sind geplant. Mithilfe der vorliegenden Ergebnisse sollen ein Abgleich von Bedarfen und angebotenen Beratungs- und Unterstützungsformaten vorgenommen und Handlungsempfehlungen formuliert werden, die einen professionellen Umgang mit diesen Phänomenen ermöglichen.

## **Gekommen, um zu bleiben? Die Promotionsphase als Wegbereiter wissenschaftlicher Karrieren**

**Gesche Brandt, Kolja Briedis, Steffen Jaksztat, Susanne de Vogel**

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Deutschland; [devogel@dzhw.eu](mailto:devogel@dzhw.eu)

Eine abgeschlossene Promotion gilt als Voraussetzung, um eine Karriere in der Wissenschaft einzuschlagen. Doch längst nicht jede(r) Promovierende strebt mit der Promotion eine wissenschaftliche Karriere an (Fabian et al. 2013), u. a. weil eine Promotion häufig auch die Berufschancen auf dem außerakademischen Arbeitsmarkt verbessert (Enders & Bormann 2001; Engelage & Schubert 2009; Heineck & Matthes 2012; Falk & Küpper 2013).

Je nach formaler Promotionsform sind die Promovierenden während ihrer Promotionsphase bereits mehr oder weniger stark in das Wissenschaftssystem eingebunden (Berning & Falk 2006; Enders & Kottmann 2009; Jaksztat et al. 2012). Hieraus ergeben sich möglicherweise kontextspezifische Unterstützungsstrukturen, die wiederum unterschiedliche Voraussetzungen für eine akademische Laufbahn schaffen können. Empirische Studien, die den Zusammenhang zwischen bestimmten Promotionsbedingungen und den Werdegängen nach der Promotion untersuchen, sind jedoch selten (z.B. Berweger 2008). Die vorliegenden Analysen setzen hier an und überprüfen, inwiefern sich die Lern- und Entwicklungsbedingungen, die Promovierte während ihrer Promotionsphase in unterschiedlichen formalen Promotionsformen vorgefunden haben, auf die wissenschaftliche Karriereintention und auf Verbleib in der Wissenschaft auswirken.

Welche formalen Promotionsformen und Unterstützungsbedingungen während der Promotionsphase befördern eine wissenschaftliche Karriereintention Promovierter und ihren tatsächlichen Verbleib in der Wissenschaft?

Aus der Perspektive der sozial-kognitiven Laufbahnthorie (Lent, Brown & Hackett 1994) wird angenommen, dass die Unterstützung beim Ausbau wissenschaftlicher Netzwerke, eine umfassende fachliche und persönliche Unterstützung sowie Bedingungen, die die eigene wissenschaftliche Arbeit der Promovierenden fördern, das Interesse an einer wissenschaftlichen Karriere begünstigen. Entsprechend sollten positive Lernerfahrungen und –ergebnisse sollten wiederum dazu führen, dass Promovierte ein höheres Maß an forschungsbezogenen Selbstwirksamkeitserwartungen und Interessen ausbilden und deshalb häufiger berufliche Ziele innerhalb der Wissenschaft anstreben. Aus arbeitsmarktökonomischer Perspektive dürften die Outcomes dieser Unterstützungsbedingungen – wie etwa Publikationen, Vorträge oder Mobilitätserfahrungen – wiederum wichtige Signale darstellen, welche die tatsächlichen Verbleibchancen in der Wissenschaft wesentlich bestimmen (Spence 1973). Aus signal- und netzwerktheoretischer Sicht sollte zudem ein großes Karrierenetzwerk maßgeblich sein (Granovetter 1973; Lin 2002).

Um diese Annahmen zu prüfen, werden die Daten des repräsentativen, bundesweiten DZHW Promoviertenpanels genutzt. Insgesamt beteiligten sich rund 5.400 Promovierte des Prüfungsjahres 2014 aller Fachrichtungen und Promotionsformen an der Befragung. Durch die Fokussierung auf einen aktuellen Promoviertenjahrgang ist sichergestellt, dass verschiedene Formen der Individualpromotion und strukturierte Promotionsprogramme miteinander verglichen werden können. Mit den ersten zwei Erhebungswellen liegen Informationen für die ersten beiden Jahre nach dem Abschluss der Promotion vor.

Mit logistischen Regressionsmodellen wird im ersten Schritt untersucht, inwieweit die unterschiedlichen formalen Promotionsformen und die Unterstützungsbedingungen in der Promotionsphase das Interesse an einer wissenschaftlichen Karriere beeinflussen. Im zweiten Schritt wird geprüft, wie die Promotionsformen und die Outcomes der Unterstützungsbedingungen die tatsächlichen Verbleibchancen in der Wissenschaft begünstigen. In den Analysen werden dabei sieben verschiedene Förderdimensionen differenziert: Die Qualität des Promotionsergebnisses, wissenschaftliche Publikationen, wissenschaftliche Tagungen, forschungsbezogene Auslandsmobilität, Hochschullehre, der Praxisbezug sowie der Ausbau wissenschaftlicher Netzwerke.

Den Analysen zufolge stehen die Unterstützungsbedingungen während der Promotionsphase mit der wissenschaftlichen Karriereintention Promovierter in Zusammenhang. Insbesondere jene Promovierten, die während der Promotionsphase Unterstützung bei der Einbindung in die Scientific Community erhalten haben, wollen nach Abschluss ihrer Promotion mit höherer Wahrscheinlichkeit weiterhin in der Wissenschaft tätig sein. Erwartungskonform erhöhen eine gute Promotionsnote, Tagungsbesuche in der Promotionsphase, eine gute Netzwerkeinbindung und – mit Einschränkungen – Forschungsaufenthalte im Ausland, die Wahrscheinlichkeit, nach dem Abschluss der Promotion weiterhin in der Wissenschaft tätig zu sein.

Hinsichtlich der Promotionsform bieten die Unterstützungsbedingungen in strukturierten Programmen und auf wissenschaftlichen Mitarbeiterstellen ähnlich gute Voraussetzungen für eine wissenschaftliche Karriere. Insbesondere Promovierte aus Drittmittelprojekten sind in den ersten zwei Jahren nach der Promotion mit höherer Wahrscheinlichkeit weiterhin in der Wissenschaft beschäftigt. Freie Promotionen machen eine Karriere im Wissenschaftssystem hingegen deutlich unwahrscheinlicher.

## Studienverläufe in Lehramtsstudiengängen - Befunde zur Revision und Stabilität von Studienentscheidungen

**Andreas Ortenburger<sup>1</sup>, Hilde Schaeper<sup>1</sup>, Sebastian Franz<sup>2</sup>**

<sup>1</sup>DZHW - Deutsches Zentrum für Hochschul- und Wissenschaftsforschung; <sup>2</sup>LifBi – Leibniz-Institut für Bildungsverläufe e.V.;  
[ortenburger@dzhw.eu](mailto:ortenburger@dzhw.eu), [sebastian.franz@lifbi.de](mailto:sebastian.franz@lifbi.de)

Die universitäre Lehrerbildung ist die erste Phase eines längeren Professionalisierungsprozesses, der in erfolgreiches Unterrichtshandeln münden soll (Cochran-Smith 2001; Terhart 2004) und mit der Entscheidung für ein Lehramtsstudium beginnt. Diese Entscheidung wird in manchen Fällen nach Umwegen getroffen, in anderen Fällen revidiert, z. B. durch Abbruch des Studiums. Während zum Thema Studienabbruch insgesamt eine Vielzahl an Untersuchungen vorliegt, ist über Studienabbruch von Lehramtsstudierenden im Besonderen nur wenig bekannt. Bisherige Studien konzentrieren sich auf Unterschiede zwischen Studienabbrecher(inne)n mit und ohne Lehramtsorientierung (Heublein et al. 2012; Blüthmann und Thiel 2011; Heublein et al. 2003), sind auf eine Fachrichtung beschränkt (z. B. Blömeke 2009) oder betrachten ausschließlich und regional begrenzt Lehramtsstudienabbrecher(innen) (Herfter et al. 2015; Gesk 1997). Dabei wird häufig nicht unterschieden, ob das Hochschulsystem verlassen oder in ein anderes Studium gewechselt wird. Unberücksichtigt bleibt meistens ebenfalls, wer seine Entscheidung im Hinblick auf das angestrebte Lehramt zwar revidiert, dem Berufsziel an sich aber treu bleibt (Ausnahme, allerdings auf eine Universität beschränkt: Gausch und van Buer 2011). Selten thematisiert wird schließlich auch die Gruppe derer, die aus einem anderen Studiengang in ein Lehramtsstudium gewechselt sind (Ausnahme z. B.: Rothland 2011). Von punktuellen Einblicken abgesehen fehlt es demnach an einer systematischen Untersuchung unterschiedlicher Verlaufsmuster in einem Lehramtsstudium.

Dieses Desiderat aufgreifend sollen im Beitrag in einem ersten Schritt verschiedene Gruppen von Lehramtsstudierenden hinsichtlich ihrer Studienverlaufsentscheidungen differenziert und quantifiziert werden. In einem zweiten Schritt wird der Frage nachgegangen, welche soziostrukturellen und psychologisch-individuellen Merkmale eine Korrektur der ursprünglichen Studienentscheidung begünstigen. Dabei wird im Anschluss an das psychologische Erwartungs-Wert-Modell (Wigfield und Eccles 1992) zum einen untersucht, inwieweit niedrige Leistungen, ein geringes akademisches Selbstkonzept und falsche Erwartungen bzw. eine mangelnde Passung von Studium und Interessen zu einem Abbruch oder Wechsel des Studiums führen. Den theoretischen Überlegungen von Spady (1970) und Tinto (1975) folgend wird zum anderen der Frage nachgegangen, welche Rolle die soziale und akademische Integration für die Revision der Studienentscheidung von Lehramtsstudierenden spielt.

Die Analysen erfolgen mit den Daten der Startkohorte 5 des Nationalen Bildungspanels (NEPS; doi:10.5157/NEPS:SC5:9.0.0). In dieser NEPS-Teilstudie wird eine Kohorte von Studienanfänger(inne)n des Wintersemesters 2010/2011 an deutschen Hochschulen im Längsschnitt untersucht und jede Veränderung im Studienverlauf erfasst.

Betrachtet werden 4.646 Personen, die ein Erststudium mit dem Abschlussziel Lehramt aufgenommen haben und bis zur Befragung 2015 (aktuell verfügbare Daten) im Panel verblieben sind. Die größte Gruppe machen Studierende aus, bei denen keine Veränderung festgestellt werden konnte, d.h. der angestrebte bzw. erreichte Abschluss entspricht dem Studienziel zum Zeitpunkt der Studienaufnahme. Ihr Anteil liegt mit 78 Prozent so hoch wie bei Gausch und van Buer (2011). Weitere sechs Prozent studieren im NEPS weiter auf Lehramt, haben jedoch die Art des angestrebten Lehramts gewechselt. Rund neun Prozent haben das Lehramtsstudium angebrochen, um etwas anderes zu studieren. Ein kompletter Abbruch des Studiums ist für lediglich drei Prozent zu verzeichnen. Weitere drei Prozent des Analysesamples sind in ein Lehramtsstudium hineingewechselt.

Für die Untersuchung der Faktoren, die die Revision von lehramtsbezogenen Studienentscheidungen beeinflussen, werden multinomiale Logitmodelle geschätzt. Die Analysen, die derzeit noch nicht abgeschlossen sind, versprechen weitere Aufschlüsse über die Studienentscheidungen von Lehramtsstudierenden. Perspektivisch kann die Unterscheidung verschiedener Gruppen von Absolvent(inn)en eines Lehramtsstudiums hinsichtlich ihres Studienverlaufs auch für die Untersuchung des Outcomes (professionelle Kompetenz) der Lehrerbildung fruchtbar gemacht werden.

## C05

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Rosshofgasse, S01

### Von der informellen zur formellen Diagnose – ein experimenteller Ansatz zur Untersuchung des diagnostischen Prozesses bei Lehramtsstudierenden

**Jacqueline Wißmann, Annette Kinder**

Freie Universität Berlin, Deutschland; [j.wissmann@fu-berlin.de](mailto:j.wissmann@fu-berlin.de)

#### Hintergrund und Fragestellung

Die Beurteilung von Schüler\*innen ist anfällig für Fehler und Verzerrungen (vgl. Südkamp et al., 2012). Für eine Diagnose hoher Güte ist eine hypothesengeleitete, systematische Herangehensweise unter Verwendung evidenzbasierter Verfahren unerlässlich (formelle Diagnose; vgl. Schrader, 2010). Die Förderung von Beurteilungskompetenzen ist daher bereits im Rahmen der Lehramtsausbildung von entscheidender Bedeutung (vgl. KMK, 2004). Zur Evaluation, ob diese tatsächlich im Rahmen der Ausbildung erworben wurden, mangelt es derzeit noch an geeigneten Messinstrumenten. Eine Schwierigkeit besteht darin, dass für die Förderung wie für die Erfassung von Beurteilungskompetenzen – operationalisiert als Urteilsgenauigkeit - in der Regel die Erhebung von realen Schülerleistungen notwendig ist (vgl. Schrader & Helmke, 1987; Helmke, 2009). Über die Simulation virtueller Schüler\*innen konnte jedoch die Erfassung von Beurteilungskompetenzen bereits ohne Realdaten ermöglicht werden (vgl. Schülerinventar; Kaiser et al., 2015).

Im Rahmen der Studie wurde ein Instrument erprobt, mit Hilfe dessen prozedurales Wissen von Lehramtsstudierenden im Bereich der Diagnostik von Lernbesonderheiten (z.B. Leseschwäche, Hochbegabung) erfasst und trainiert werden können. Dabei sollten auch Prozessvariablen, die korrekte diagnostische Entscheidungen begünstigen, identifiziert werden.

#### Methoden

An der Studie nahmen 241 Studierende teil, die sich entweder im ersten oder im dritten Lehramts-Master-Semester befanden. Ihnen wurde ein Fallinventar mit acht realistischen Fallvignetten vorgelegt, in dem Schüler\*innen mit und ohne Lernbesonderheiten vorgestellt wurden. Die Vignetten enthielten Informationen über drei schulisch relevante Fähigkeitsbereiche (Lesekompetenz, mathematische Kompetenzen, kognitive Fähigkeiten), die sich hinsichtlich ihrer diagnostischen Güte unterschieden: Neben Ergebnissen evidenzbasierter Testverfahren (formelle Information) wurden bspw. auch mehr oder weniger systematische Beobachtungen von Lehrkräften oder Eltern beschrieben (semiformelle bzw. informelle Information). Die Studierenden sollten für jede Fallvignette eine Verdachtsdiagnose und eine pädagogische Empfehlung aus je sechs Optionen auswählen. Anschließend gaben die Studierenden an, welche Information sie jeweils für die Diagnose- und die Empfehlungsentscheidung genutzt hatten (formell, semiformell oder informell). Ferner waren sie zur Angabe ihrer subjektiven Sicherheit beim Treffen der Entscheidungen gebeten. Die Erhebung erfolgte einmalig über eine Online-Plattform und zusätzliche Vorlage von Fallheften. Außerdem wurde erfasst, ob die Studierenden die Übung sinnvoll zur Vorbereitung auf die Klausur „Pädagogische Diagnostik“ sowie zur Vorbereitung auf die spätere Berufspraxis fanden. Die Musterlösungen wurden durch Expertenbefragung abgesichert.

#### Ergebnisse

Die Itemschwierigkeiten variierten für Diagnosen und Empfehlungen in Abhängigkeit von den (inhaltlich heterogenen) Fällen im Bereich  $p=.25-.91$ . Die Lösungsraten für die Diagnosen waren dabei mit über 70% signifikant höher als bei Empfehlungen, die in rund 57% der Entscheidungen korrekt gewählt wurden. Bei beiden Entscheidungen gaben die Studierenden zumeist an, die formellen Informationen genutzt zu haben. Für Diagnosen war die Bevorzugung formeller Diagnoseinformationen jedoch deutlicher. Wenn formelle Informationen genutzt wurden, waren außerdem die Lösungen im Mittel häufiger korrekt als bei der Nutzung semiformeller oder informeller Information. Die subjektive Sicherheit unterschied sich allerdings nicht zwischen Diagnosen und Empfehlungen, bei beiden Entscheidungen waren sich die Studierenden im Mittel eigenen Angaben zufolge „eher sicher“ (1=gar nicht sicher, 6=sehr sicher).

Die Übung wurde insgesamt im Mittel als „eher hilfreich“ (1=gar nicht hilfreich, 6=sehr hilfreich) für die Prüfung im Modul Pädagogische Diagnostik und „hilfreich“ für die spätere Berufspraxis eingeschätzt.

#### Diskussion

Das Fallinventar ist eine praxisnahe Übungsmöglichkeit von Beurteilungs-kompetenzen, die von Studierenden als hilfreich empfunden wird. Die relativen Schwierigkeiten beim Ableiten pädagogischer Empfehlungen geben Hinweis auf einen Übungsbedarf der Studierenden. In Lerngelegenheiten zur gezielten Einübung des datengestützten diagnostischen Prozesses (z.B. Jäger, 2006; Mandinach & Gummer, 2016) könnte diesem Defizit ggf. begegnet werden. Zur Evaluation solcher Lerngelegenheiten könnte das Instrument als kriterialer Test zur Ermittlung des Lernstandes von Studierenden zum Einsatz kommen.

#### Ausblick

Aktuell werden die erhobenen Daten auch in Hinsicht auf Zusammenhänge mit relevanten Kovariaten (Einstellung zu Diagnostik, selbsteingeschätzte diagnostische Kompetenz, Intelligenz, u.a.) und auf Gruppenunterschiede (Master-Semester) analysiert. Weiterhin wird experimentellen Fragestellungen zum diagnostischen Prozess (Stereotypeneffekt Fallgeschlecht, Reihenfolge Diagnose – Empfehlung) nachgegangen.

### Formativ-diagnostisches Handeln von Lehrpersonen im Unterricht: Überzeugungen von Lehrpersonen und Wahrnehmung durch die Schülerinnen und Schüler

**Alois Buholzer<sup>1</sup>, Matthias Baer<sup>2</sup>, Hanni Loetscher<sup>1</sup>, Sandra Zulliger<sup>1</sup>, Merle Ruelmann<sup>1</sup>, Loredana Torchetti<sup>1</sup>**

<sup>1</sup>Pädagogische Hochschule Luzern, Schweiz; <sup>2</sup>Pädagogische Hochschule Zürich, Schweiz; [sandra.zulliger@phlu.ch](mailto:sandra.zulliger@phlu.ch), [merle.ruelmann@phlu.ch](mailto:merle.ruelmann@phlu.ch)

Theoretischer Hintergrund: Diagnostische Informationen, die während des laufenden Unterrichts („on the fly“) von der Lehrperson gewonnen werden, haben eine zentrale Funktion für die Steuerung des Lehren und Lernens. Auf ihrer Grundlage können Lehrpersonen lernrelevante Entscheidungen informiert treffen und begründen (Praetorius & Südkamp 2017). Diese unterrichtsbegleitende Diagnose und Förderung wird auch als „formativ-diagnostisches Handeln“ (Maier, 2014) oder im englischsprachigen Raum als „formative assessment“ (Filsecker & Kerres, 2012) bezeichnet.

Neueren Forschungsarbeiten zufolge (Gotwals, Philhower, Cisterna & Bennett, 2015; Ruiz-Primo & Furtak, 2006) zeigt sich formativ-diagnostisches Handeln von hoher Qualität dann, wenn die Lehrperson: 1) Lernziele und Erfolgskriterien mit den Schülerinnen und Schülern bespricht, 2) anregende Fragen zum Lernprozess stellt, 3) Self-Assessments und 4) Peer-Assessments anleitet sowie 5) lernförderliches Feedback gibt. Die Effekte dieser Strategien, insbesondere auf die schulische Leistung, werden durch eine reichhaltige Forschung gestützt (für eine Übersicht McMillan, Venable & Varier, 2013). Welche Faktoren den Einsatz der Strategien im Unterricht begünstigen, ist jedoch bisher weitgehend ungeklärt.

Oser und Blömeke (2012) zufolge haben Überzeugungen (beliefs) einen entscheidenden Einfluss auf das professionelle Handeln von Lehrpersonen. Es ist deshalb auch davon auszugehen, dass spezifisch für das formativ-diagnostische Handeln Überzeugungen zu Lehr- und Lernprozessen eine handlungsleitende Rolle spielen. Hinsichtlich dieses Zusammenhangs liegen gemäß unseren Recherchen kaum robuste Daten vor.

Im Rahmen der Unterrichtsqualitätsforschung konnte des Weiteren gezeigt werden, dass Lernende, Lehrpersonen und externe Beobachter das Unterrichtsgeschehen z.T. unterschiedlich wahrnehmen (u.a. Maulana & Helms-Lorenz, 2016). Lehrpersonen tendieren beispielsweise im Vergleich zu ihren Schülerinnen und Schülern dazu, ihr Lehrverhalten qualitativ besser einzuschätzen (u.a. Fraser & O'Brien, 1985). Es kann angenommen werden, dass es auch im Hinblick auf das formativ-diagnostische Handeln Unterschiede in der Wahrnehmung gibt. Eine hohe Relevanz der Betrachtung der verschiedenen Wahrnehmungen auf das Lehrpersonenhandeln besteht v.a. durch dessen Zusammenhang mit der Lernentwicklung der Schülerinnen und Schüler (Fauth et al., 2014). Seidel und Shavelson (2007) konnten diesbezüglich zeigen, dass die gemessene Effektstärke auf die Leistung in Abhängigkeit von der Datengrundlage (Lehrperson, Schülerinnen und Schüler oder Beobachter) variiert.

Die Fragestellungen, die an der Tagung präsentiert und diskutiert werden, fokussieren darauf, wie Lehrpersonen ihr formativ-diagnostisches Handeln im Unterricht einschätzen und wie ihre Überzeugungen zum Lehr- und Lernprozess damit zusammenhängen. Außerdem werden die Sichtweisen von Lehrperson und Lernenden auf das formativ-diagnostische Handeln vergleichend untersucht.

Forschungsfragen:

(1) Wie schätzen Lehrpersonen ihr formativ-diagnostisches Handeln im Unterricht ein?

(2) Wie hängen berufsbezogene Überzeugungen der Lehrpersonen (zur formativen Beurteilung, konstruktivistische vs. transmissive Lehrüberzeugungen und Bezugsnormorientierungen) mit dem selbstberichteten formativ-diagnostischen Handeln zusammen?

(3) Inwieweit hängen das durch die Lehrpersonen berichtete mit dem durch die Schülerinnen und Schüler wahrgenommene formativ-diagnostische Handeln zusammen?

Methode:

Die an der Tagung präsentierte Teiluntersuchung zur Wahrnehmung des formativ-diagnostischen Handelns ist an ein umfassenderes Forschungsprojekt angegliedert, das durch den Schweizerischen Nationalfonds (SNF) gefördert wird (Projekt-Nr. 100019 169771). Die Hauptuntersuchung ist eine Videostudie und erfasst die formativ-diagnostische Performanz von Lehrpersonen, mögliche Prädiktoren des Handelns und die Auswirkungen auf den Lernprozess der Schülerinnen und Schüler.

An dem Forschungsprojekt nahmen 52 Lehrpersonen der Zentralschweiz und ihre Schülerinnen und Schüler (N = 621) teil. Neben 90-minütigen Videoaufzeichnungen des Mathematikunterrichts wurden in der Studie umfangreiche Fragebogendaten erhoben, welche Grundlage der präsentierten Analyse sind.

Zur Untersuchung der oben genannten Fragestellungen beantworteten Lehrpersonen sowie Schülerinnen und Schüler parallelisierte Fragebögen zur formativ-diagnostischen Performanz der Lehrperson in fünf Bereichen (Einsatz von Lernzielen, Fragen, Self- und Peer-Assessment sowie Feedback). Darüber hinaus gaben die Lehrpersonen mittels Fragebögen über ihre professionellen Überzeugungen Auskunft.

Ergebnisse: Die Datenerhebung der Hauptstudie wurde im Juli 2017 abgeschlossen, zurzeit werden die Daten aufbereitet und ausgewertet. Die statistische Auswertung wird mittels Regressions- und Multilevel-Analysen erfolgen. Die Ergebnisse werden am Kongress vorgestellt.

### **Professionelle Kompetenzen pädagogischer Fachkräfte in Bezug auf Diagnostik und Förderung im Kindergarten – Analysen schriftlicher Antworten zu einer Fallvignette**

**Lea Höltge<sup>1,2</sup>, Janine Buchholz<sup>1</sup>, Ulrike Hartmann<sup>2,3</sup>, Jan-Henning Ehm<sup>1,2</sup>, Marcus Hasselhorn<sup>1,2</sup>**

<sup>1</sup>DIPF, Deutschland; <sup>2</sup>IDEA-Zentrum, Deutschland; <sup>3</sup>Bergische Universität Wuppertal, Deutschland; [hoeltge@dipf.de](mailto:hoeltge@dipf.de)

In den letzten zwei Jahrzehnten hat der Bildungsauftrag des Kindergartens zunehmend Aufmerksamkeit erfahren, weswegen insbesondere die Unterstützung kindlicher Lernprozesse als Kernaufgabe frühpädagogischer Fachkräfte angesehen werden kann. So sieht die Kultusministerkonferenz (KMK, 2004) einen Schwerpunkt des Bildungsauftrags darin, Kinder in ihren individuellen Potentialen zu stärken und sie im Lernen des Lernens zu unterstützen. Des Weiteren macht sie deutlich, dass auch die spezifische Förderung von Kindern mit Entwicklungsrisiken zu den Aufgaben frühpädagogischer Fachkräfte gehört. Voraussetzung für die Erfüllung dieses Auftrags ist, dass Entwicklungsrisiken und Förderbedarfe erkannt werden. So verwundert es nicht, dass der verstärkte Fokus auf den Bildungsauftrag auch eine rege Debatte über die Professionalisierung frühpädagogischer Fachkräfte entfacht hat. Bislang liegt jedoch nur wenig Forschung zu ihren professionellen Kompetenzen und deren Bedeutung für ihr praktisches Handeln vor (Roßbach, 2013). Zu den Facetten professioneller Handlungskompetenz frühpädagogischer Fachkräfte gehören neben professionellem Wissen und Fähigkeiten auch Einstellungen, Motivation, emotionale Aspekte und Selbstregulation; Facetten, die sich über die berufliche Karriere fortlaufend entwickeln (Anders, 2012). Zwar schätzen die Fachkräfte ihre Fähigkeiten selbst zumeist hoch ein (z.B. Mischo et al., 2013; Höltge et al., 2017), doch scheinen im Hinblick auf die Unterstützung kindlicher Lernprozesse deutsche Fachkräfte im internationalen Vergleich weniger gut abzuschneiden (vgl. Wadepohl, 2015) und besonders im sprachlichen und mathematischen Bereich eine geringe Prozessqualität vorzuliegen (Smidt, 2012).

Vor diesem Hintergrund stellt sich die Frage, welche Maßnahmen der Diagnostik und Förderung pädagogische Fachkräfte für zielführend halten und in ihr professionelles Handlungsrepertoire aufgenommen haben und welche Aspekte sie dabei berücksichtigen. Weiterhin interessiert uns, welche personenspezifischen Merkmale hiermit zusammenhängen.

Methode

Zur Erfassung des Handlungsrepertoires wurde die Fallbeschreibung eines potentiell förderbedürftigen Kindes eingesetzt. Die Fachkräfte sollten beschreiben, was sie zur Überprüfung eines Förderbedarfs und zur Förderung empfehlen würden. Antworten liegen von N=313 frühpädagogischen Fachkräften aus 114 baden-württembergischen Kindertageseinrichtungen vor. Die schriftlichen Antworten wurden inhaltsanalytisch ausgewertet. Die Kategorien beinhalten sowohl diagnostische als auch Fördermaßnahmen (z.B. Durchführung von Beobachtungen, Einsatz von Förderprogrammen) und weitere Aspekte, die von den Fachkräften beachtet werden (z.B. die individuellen Merkmale des Kindes). Pro genannte Kategorie wurden bis zu zwei Punkte vergeben. Ein hoher Punktwert spiegelt somit eine hohe Anzahl genannter Kategorien wider. Die Beurteilungsüberstimmung betrug für die einzelnen Kategorien durchschnittlich  $\kappa=,703$ . Anhand latenter Klassenanalysen (LCA) wurden Gruppen ähnlich antwortender Fachkräfte gebildet, die hinsichtlich weiterer demographischer und kompetenzbezogener individueller Merkmale (Alter, Berufserfahrung, Selbstwirksamkeit, Einstellungen, Berufsbelastung) über die Berechnung einfaktorierlicher Varianzanalysen (ANOVA) miteinander verglichen wurden.

#### Ergebnisse

Viele Antworten der Fachkräfte fielen sehr knapp aus. Im Schnitt erhielten die Fachkräfte 16% der möglichen Punkte. Im diagnostischen Bereich waren die am häufigsten genannten Kategorien die Durchführung von Beobachtungen, Austausch über Beobachtungen, sprachliche und mathematische Angebote, anhand derer sich die Fähigkeiten beobachten lassen, sowie die Berücksichtigung individueller Merkmale des Kindes. In Bezug auf die Förderung wurden ebenfalls Angebote im sprachlichen und mathematischen Bereich genannt. Auch hier wurde betont, dass individuelle Merkmale berücksichtigt werden sollen. Fast die Hälfte der Teilnehmenden empfahl den Einbezug weiterer Personen in die Förderung.

Die Ergebnisse der LCAs weisen auf fünf Gruppen und anschließende ANOVAs auf Gruppenunterschiede hinsichtlich des Alters ( $F(4,300)=3.183, p<.05$ ) und der Berufsbelastung ( $F(4,309)=2.950, p<.05$ ) hin. So berichteten die Personen, die in erster Linie andere Fachkräfte in die Diagnostik und Förderung einbeziehen (35%) die geringste Berufsbelastung und bilden gleichzeitig mit durchschnittlich 38 Jahren die jüngste Gruppe. Mit durchschnittlich 44 Jahren sind Personen, die Ideen zu Beobachtungstätigkeiten im sprachlichen und mathematischen Bereich haben und in Bezug auf Förderung das externe Umfeld einbinden (12%) die älteste Gruppe. Die höchste Berufsbelastung wird von der Gruppe berichtet, die sowohl Ideen zur Beobachtung als auch zur Förderung hat (14%). 24% der Teilnehmenden nannten keine der berücksichtigten Kategorien und 15% äußerten vor allem Ideen zur Förderung im sprachlichen und mathematischen Bereich.

### **Ist Empathie förderlich für diagnostische Kompetenz? Verarbeitung diagnostischer Informationen aus textbasierten Fällen in der ersten Phase der Lehrkräftebildung**

**Alexander Wedel<sup>1</sup>, Franziska Greiner<sup>2</sup>, Christin Müller<sup>1</sup>**

<sup>1</sup>Technische Universität Berlin, Deutschland; <sup>2</sup>Friedrich-Schiller-Universität Jena, Deutschland; [alexander.wedel@tu-berlin.de](mailto:alexander.wedel@tu-berlin.de), [franziska.greiner@uni-jena.de](mailto:franziska.greiner@uni-jena.de), [christin.rosemarie.mueller@gmail.com](mailto:christin.rosemarie.mueller@gmail.com)

#### Relevanz der Studie und theoretischer Hintergrund

Die Professionalisierung der Lehrkräftebildung konzentriert sich aktuell u.a. auf die Entwicklung evidenzbasierter Lehr-Lern-Arrangements zur gezielten Förderung pädagogischer Kompetenzen. Für den Aufbau diagnostischer Kompetenz haben sich fallbasierte Lehr-Lern-Arrangements als wirksam erwiesen (Yadav et al., 2011; Stürmer et al., 2013). Als Lernmaterial wurden bislang insbesondere videobasierte Fälle untersucht, dagegen fehlen Erkenntnisse über textbasierte Fälle und deren Effekt auf die Kompetenzentwicklung (Syring et al., 2015).

Dem prozessorientierten Modell diagnostischer Kompetenz von Klug et al. (2013, 2017) zufolge ist die systematische Gewinnung und Verarbeitung diagnostischer Informationen die Grundlage weiterer Entscheidungsprozesse im Verlauf der Diagnostik. Ein zentrales Ziel der Entwicklung diagnostischer Kompetenz ist daher die Fähigkeit, personenbezogene Informationen in ein kohärentes Modell zu integrieren und daraus überprüfbare Hypothesen abzuleiten. Querschnittsvergleiche von Expert\_innen und Noviz\_innen identifizieren dahingehend einige Entwicklungsbereiche in der Verarbeitung diagnostischer Informationen. So können Expert\_innen generell besser diagnostisch relevante von irrelevanten Informationen unterscheiden (Klug, Bruder, & Schmitz, 2016), sie je nach Zielsetzung differenzierter auswählen (Krolak-Schwerdt, Böhmer, & Gräsel, 2009) und akkuratere Inferenzen über die Entwicklung von Unterricht und Schüler\_innenverhalten bilden (Santagata & Yeh, 2016).

Weitere Erkenntnisse zur diagnostischen Informationsverarbeitung kommen aus der Empathieforschung. Qualitative Studien weisen darauf hin, dass erfahrene Lehrkräfte insbesondere kognitive Empathie, d. h. Perspektivübernahme nutzen, um Informationen über Schüler\_innen zu gewinnen (Warren, 2017). Quantitative Studien zeigen, dass kognitive Empathie einen positiven Effekt auf die Urteilsgenauigkeit für Leistungsmerkmale von Schüler\_innen hat (Arghode, 2013; Hellmann, 2015), insbesondere auf das Speichern personenbezogener Informationen im Arbeitsgedächtnis (Bailey & Henry, 2008; Gao et al., 2016). Affektive Empathie hat dagegen einen negativen Effekt auf die erinnerte Informationsmenge (Ze et al., 2014).

Diese Effekte wurden auch in der Forschung zum Textverstehen repliziert. In mehreren Studien konnte gezeigt werden, dass Texte im psycho-narrativen Stil, d. h. mit Beschreibungen der Gedanken und Gefühle der Protagonisten zu einem Anstieg kognitiver State-Empathie wie auch zur Bildung realitätsnaher kognitiver Repräsentationen der im Text beschriebenen Situationen und Personen führen (Gillioz et al. 2012; Salem 2017). Diese sogenannten ‚situational models‘ erleichtern die Bildung von Inferenzen (van Dijk & Kintsch 1983; van der Schoot et al., 2012) und sind daher für pädagogische Diagnostik besonders wertvoll. Inwiefern diese beschriebenen Effekte auch in der Verarbeitung diagnostischer Informationen durch angehende Lehrkräfte wirken, ist jedoch bislang nicht untersucht.

#### Fragestellung und Hypothesen

Die vorliegende Studie untersucht daher die Effekte von Empathie und textstrukturellen Merkmalen diagnostischer Fälle auf die Qualität der kognitiven Repräsentationen diagnostisch relevanter Situationen im Schulkontext. Dazu werden folgende Hypothesen geprüft:

- Psycho-narrative diagnostische Textfälle führen eher zur Bildung von ‚situational models‘ als external fokalisierte Textfälle.
- Kognitive Trait- und State-Empathie begünstigen die Speicherung diagnostischer Informationen im Arbeitsgedächtnis.
- Affektive State-Empathie beschränkt die Speicherung diagnostischer Informationen im Arbeitsgedächtnis.

#### Methoden

Die Erhebung findet im Wintersemester 2017/18 in einer Stichprobe von Lehramtsstudierenden (N = 250) mit einem Experimentaldesign statt. Nach der Erhebung von Selbsteinschätzungen zu kognitiver und affektiver Trait-Empathie (SPF: Paulus, 2009) lesen die Proband\_innen einen psycho-narrativen oder external fokalierten diagnostischen Fall. Anschließend

beantworten sie Fragebögen zu den Kontrollvariablen subjektive Textverständlichkeit, Immersion der Texte, Ausmaß der kognitiven Elaboration, Ähnlichkeit zu Protagonisten und kognitiver und affektiver State-Empathie. Zur Erhebung der kognitiven Repräsentationen im Arbeitsgedächtnis wird ein Rekognitionstest (Schiefele, 1991) eingesetzt.

#### Ergebnisse

Die Hypothesen werden mit einer multiplen hierarchischen Regression geprüft. Im ersten Schritt werden die Kontrollvariablen ins Modell aufgenommen, im zweiten Schritt kognitive und affektive Empathie, im dritten Schritt die Eigenschaften der Textfälle und im vierten Schritt die Interaktion kognitiver und affektiver Empathie mit den Eigenschaften der Textfälle. Die Interaktionen werden anschließend mit einer Simple-Slope-Analyse untersucht. Die Ergebnisse der Studie geben einen Einblick in den Zusammenhang von diagnostischer Kompetenz mit Empathie und der Verarbeitung diagnostischer Informationen aus textbasierten Fällen.



## C06

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus S212

### **Erweiterung des Professionswissens durch multiprofessionelle Kooperation im Lehramtsstudium**

**Philipp Krämer<sup>1</sup>, Roswitha Ritter<sup>2</sup>, Antje Wehner<sup>3</sup>, Gertrud Lohaus<sup>2</sup>**

<sup>1</sup>Arbeitsbereich Rehabilitationswissenschaften, Institut für Bildungsforschung, Bergische Universität Wuppertal, Deutschland;

<sup>2</sup>Fachgruppe Biologie, Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Deutschland;

<sup>3</sup>School of Education, Bergische Universität Wuppertal, Deutschland; [ritter@uni-wuppertal.de](mailto:ritter@uni-wuppertal.de)

#### Theoretischer Hintergrund

Die multiprofessionelle Kooperation von Lehrkräften gilt als eine zentrale Gelingensbedingung der schulischen Inklusion (Schwager 2011, Friend et al. 2010). Die Kooperation soll dabei einen Expertisentransfer ermöglichen, so dass sich das Professionswissen der Lehrkräfte weiterentwickelt (Scruggs, Mastropieri & McDuffie 2007). Das Professionswissen wiederum ist ein zentraler Bestandteil der professionellen Kompetenz, welche den Unterrichtserfolg maßgeblich beeinflusst (Voss et al. 2015).

Formale Lerngelegenheiten bieten eine Möglichkeit, die professionelle Kompetenz weiterzuentwickeln und die Kooperation zu initiieren (Kunter et al. 2011). Daraus lässt sich die Forderung ableiten, die multiprofessionelle Kooperation als festen Bestandteil im Studium zu integrieren (Hochschulrektorenkonferenz 2015). Ob die multiprofessionelle Kooperation im Studium aber tatsächlich zu einem Expertisentransfer und damit zu einer Weiterentwicklung der professionellen Kompetenz bezüglich schulischer Inklusion führt, ist bislang nicht erwiesen.

#### Fragestellung

Das Forschungsprojekt verfolgt die Fragestellung, inwieweit die multiprofessionelle Kooperation gegenüber der monoprofessionellen Kooperation während des Studiums das Professionswissen von Lehramtsstudierenden bezüglich schulischer Inklusion beeinflusst.

#### Methodik

**Seminar & Probanden:** Das Forschungsprojekt ist eingebettet in ein disziplinübergreifendes Seminar für Lehramtsstudierende für Regelschulen und Lehramtsstudierende für sonderpädagogische Förderung an der Bergischen Universität Wuppertal. Das Seminar umfasst einen universitären Teil in Form einer Vorbesprechung und eines Blocktages sowie einen schulischen Teil in Form einer wöchentlichen Schulhospitation. Ziel der universitären Phase ist es, den Studierenden grundlegendes Wissen zu inklusivem Unterricht aus fachdidaktischer und sonderpädagogischer Perspektive zu vermitteln sowie die Studierenden auf das Co-Teaching vorzubereiten. Innerhalb der universitären Phase werden Teams gebildet, die entweder aus einem Lehramtsstudierenden für Regelschulen und einem Lehramtsstudierenden für sonderpädagogische Förderung bestehen (Interventionsgruppe) oder aber aus zwei Lehramtsstudierenden einer Disziplin (Kontrollgruppe). Ziel der schulischen Phase ist es, dass die Teams semesterbegleitend gemeinsam einen Tag in der Woche in einer inklusiven Klasse verbringen. Dabei wird der Unterricht für die inklusive Klasse durch die Teams nach dem Prinzip des Co-Teachings geplant, durchgeführt und analysiert. An dem Seminar haben insgesamt 65 Lehramtsstudierende aus drei konsekutiven Semestern teilgenommen (31 für Regelschule aus dem Master und 34 für Sonderpädagogik aus dem Bachelor), davon 41 in der Interventionsgruppe und 24 in der Kontrollgruppe.

**Datenerhebung & -auswertung:** Das Professionswissen der Studierenden wird mit Hilfe von Concept Maps zu drei verschiedenen Zeitpunkten erhoben: t1 = vor dem Seminar, t2 = nach der universitären Phase und t3 = nach der schulischen Phase. Concept Maps visualisieren Wissen in einem strukturierten, hierarchischen System bestehend aus Konzepten, welche mit logischen Verknüpfungen miteinander verbunden sind (Novak & Cañas 2008). Die beschreibende Analyse der Concept Maps wird unter Graph-theoretischen Aspekten durchgeführt (Stracke 2004). Die Auswertung erfolgt mittels induktiver, zusammenfassender qualitativer Inhaltsanalyse (Mayring 2008) der enthaltenen Propositionen. Um die Daten vor dem Hintergrund möglicher Störvariablen relativieren zu können, wird das Seminar qualitativ und quantitativ evaluiert, werden die Lehrkräfte der inklusiven Klassen vorab interviewet und füllen die Studierenden an jedem Hospitationstag ein Lerntagebuch aus.

#### Ergebnisse

Derzeit erfolgt die qualitative Analyse der erhobenen Daten.

Durch die induktive und zusammenfassende Inhaltsanalyse aller Concept Maps (t1, t2 und t3 gemeinsam) wird ein Kategoriensystem erstellt. Die Concept Maps der drei Messzeitpunkte (t1, t2 und t3 getrennt) werden dann jeweils dem Kategoriensystem zugeordnet, so dass sich die Entwicklung des Professionswissens abbilden lässt. Anschließend wird die Entwicklung der Interventionsgruppe mit der Entwicklung der Kontrollgruppe verglichen. Die Gütekriterien der qualitativen Forschung (Mayring 2010) sowie insbesondere die Berechnungen der Interraterreliabilitäten (Cohen 1960, Landis & Koch 1977) werden dabei berücksichtigt. Die Schlussfolgerungen sollen auf Grundlage der Ergebnisse die Frage beantworten, ob die multiprofessionelle Kooperation gegenüber der monoprofessionellen Kooperation tatsächlich zu einer stärkeren Weiterentwicklung der professionellen Kompetenz führt.

Die Ergebnisse werden voraussichtlich bis Ende des Jahres vorliegen.

### **Unterstützung professioneller Entwicklung angehender Lehrkräfte durch Lehreinheiten zu inklusiver Fachdidaktik in der Lehramtsausbildung**

**Bianka Troll, Michael Besser, Simone Abels, Steffen Greve, Jessica Süßenbach**

Leuphana Universität Lüneburg, Deutschland; [troll@leuphana.de](mailto:troll@leuphana.de)

#### THEORETISCHER HINTERGRUND

Der Umgang mit Diversität in inklusiver Schule ist als zentrale Herausforderung des deutschen Bildungssystems zu verstehen. Da vor allem professionelle Handlungskompetenz von Lehrkräften Unterrichtsqualität und schulisches Lernen bedingt (Baumert u. a., 2010; Kleickmann u. a., 2017), ist die Unterstützung professioneller Entwicklung angehender Lehrkräfte in der universitären Ausbildungsphase zur Vorbereitung auf inklusive Lehr-Lern-Settings von besonderer Bedeutung (Wolfswinkler, Fritz-Stratmann & Scherer, 2014). Neben bildungswissenschaftlichen Modulen zum Umgang mit Diversität ist dabei insbesondere auch der Aufbau inklusionsorientierter Ausbildungsinhalte in den universitären Fachdidaktiken erforderlich

(Amrhein & Dziak-Mahler, 2014). Hier stehen die Hochschulen jedoch noch am Anfang (Greiten, 2014). Um die Diskussion inklusiver Fachdidaktiken an der Hochschule evidenzbasiert voranzubringen, wurden im Forschungsprojekt IBall (Projektleitung: Besser) im interdisziplinären Verbund aus Bildungswissenschaft und Fachdidaktiken verschiedener Fächer (Naturwissenschaften, Mathematik, Musik, Sport) an der Leuphana Universität Lüneburg theorie- und empiriebasierte Lehreinheiten, die gezielt eine Unterstützung der Entwicklung professioneller Handlungskompetenz (Baumert & Kunter, 2006) angehender Lehrkräfte in inklusiver Schule fokussieren, entwickelt und in bestehende Lehrveranstaltungen implementiert. Eine fächerübergreifende wissenschaftliche Begleitung dieser Lehreinheiten untersucht empirisch deren Wirkung auf die Entwicklung spezifischer Facetten professioneller Handlungskompetenzen (Einstellung, Selbstwirksamkeitsüberzeugungen und selbsteingeschätzte fachdidaktische Fähigkeit zu inklusiver Schule).

## FRAGESTELLUNG

In der vorliegenden Studie sind folgenden Fragestellungen handlungsleitend:

1. Status Quo: Welche Einstellung, Selbstwirksamkeitsüberzeugungen und selbsteingeschätzte fachdidaktische Fähigkeiten zu inklusiver Schule haben angehende Lehrkräfte vor der Teilnahme an einer fachspezifischen Lehreinheit zu Inklusion im Fachunterricht? Wie unterscheiden sich die Ergebnisse zwischen den Fächern?
2. Entwicklung: Wie verändern sich Einstellung, Selbstwirksamkeitsüberzeugungen und selbsteingeschätzte fachdidaktische Fähigkeiten zu inklusiver Bildung angehender Lehrkräfte durch die Teilnahme an einer fachspezifischen Lehreinheit zu Inklusion im Fachunterricht? Können Unterschiede in den Ergebnissen zwischen den Fächern festgestellt werden?

## METHODE

Um den eher explorativ angelegten Fragestellungen zur Wirkung dieser Lehreinheiten nachgehen zu können, wurden Fragebögen in einem Prä-Post-Design eingesetzt, welche die Einstellung zur fachspezifischen Inklusion (6 Items;  $\alpha = .78$  (Prä)/  $.74$  (Post), adaptiert von Seifried und Heyl (2016)), die Selbstwirksamkeit gegenüber inklusivem Fachunterricht (4 Items;  $\alpha = .68$  (Prä)/  $.71$  (Post), adaptiert von Bosse und Spörer (2014)) und die selbstwahrgenommene fachdidaktische Fähigkeit im inklusiven Fachunterricht (4 Items;  $\alpha = .72$  (Prä)/  $.72$  (Post); selbst entwickelt) erfragen. Ebenso wurde nach dem persönlichen Nutzen (4 Items;  $\alpha = .88$  (Post), adaptiert von Gläßer u.a. (2002)) der Lehreinheit gefragt. Die Likert-Skalen reichen von 0 = keine Zustimmung bis 3 = volle Zustimmung. Insgesamt haben  $N = 56$  Studierende (Mathematik: 11; Musik: 9; Naturwissenschaft: 16; Sport: 20) zu beiden Messzeitpunkten der Befragung teilgenommen. Zur Auswertung der Fragestellungen wurden Post-hoc Tests sowie Varianz- und Regressionsanalysen herangezogen.

## ERGEBNISSE UND DISKUSSION

Vor Durchführung der spezifischen Lehreinheiten zeigen sich im Mittel eher durchschnittliche Ausprägung von Einstellung ( $M=1.56$ ,  $SD=.53$ ), Selbstwirksamkeit ( $M=1.66$ ,  $SD=.50$ ) und selbstberichteter fachdidaktischer Fähigkeiten ( $M=1.54$ ;  $SD=.52$ ) der Studierenden. Multivariate Gruppenvergleiche zeigen im Prätest kaum Unterschiede zwischen den Fächern, allein Studierende mit Fach Musik verfügen über vergleichsweise geringere Selbstwirksamkeit (im Vergleich zu Naturwissenschaften;  $p=.044$ ;  $d=1.22$ ) sowie vergleichsweise geringere selbstberichtete fachdidaktische Fähigkeiten (im Vergleich zu Sport;  $p=.036$ ;  $d=1.25$ ) (Forschungsfrage 1). Die Ergebnisse der Längsschnittanalysen zeigen, dass hauptsächlich positive Entwicklungen der selbsteingeschätzten fachdidaktischen Fähigkeiten der Studierenden in den Lehreinheiten Mathematik ( $t(10)=-2.45$ ,  $p=.034$ ), Musik ( $t(8)=-4.24$ ,  $p=.003$ ) und Naturwissenschaften ( $t(15)=-3.85$ ,  $p=.002$ ) vorliegen. Varianzanalysen verdeutlichen, dass diese Entwicklung für das Fach Naturwissenschaften signifikant stärker ausfällt, als in allen anderen Fächern ( $F(4,51)=7.15$ ,  $p<.001$ ). Erklärt werden kann dies durch einen direkten Effekt des wahrgenommenen persönlichen Nutzens (Regressionsanalysen mit  $F(2,53)=8.50$ ,  $p=.001$ ) (Forschungsfrage 2).

Die Ergebnisse machen deutlich: Lehramtsstudierende verfügen trotz offensichtlicher Notwendigkeit nur über bedingt ausgebildete, fachspezifische professionelle Kompetenzen zu inklusiver Schule. Selbst kurze fachdidaktische Lehreinheiten zu inklusiver Schule können professionelle Entwicklung jedoch erfolgreich anstoßen, dieser Effekt hängt aber vom wahrgenommenen Nutzen der Lehreinheit ab.

### **Wie unterscheiden sich Lehrpersonen gleicher Wissensprofile hinsichtlich ihrer Überzeugungen, Selbstwirksamkeit und Ausbildungsmerkmale?**

**Doreen Holtsch<sup>1</sup>, Johannes Hartig<sup>2</sup>, Fabio Sticca<sup>1</sup>, Franz Eberle<sup>1</sup>**

<sup>1</sup>Universität Zürich, Deutschland; <sup>2</sup>DIPF, Frankfurt; [doreen.holtsch@uzh.ch](mailto:doreen.holtsch@uzh.ch)

\*Theoretischer Hintergrund\*

Die Forschung zur Professionalität von Lehrpersonen hat in den letzten Jahren einerseits auf die direkte Erfassung von Wissen der Lehrpersonen und andererseits auf die Erfassung der formalen Lerngelegenheiten fokussiert. Da Wissen als Kern professioneller Kompetenz betrachtet wird (Baumert & Kunter, 2011), wurde vor allem der Zusammenhang zwischen dem fachlichen sowie fachdidaktischen Wissen von angehenden Lehrpersonen und ihren Lerngelegenheiten im Studium untersucht (z. B. Blömeke, Suhl, Kaiser, Felbrich, & Schmotz, 2010). Neben dem Wissen der Lehrpersonen spielen ihre instruktionalen und konstruktivistischen Überzeugungen zum Lehren und Lernen (Voss, Kleickmann, Kunter, & Hachfeld, 2011) und ihre Selbstwirksamkeit als motivationale Orientierung (Kunter, 2011) für ihre professionelle Kompetenz sowie für die Unterrichtsgestaltung ebenfalls eine wichtige Rolle.

Analysen in der Deutschschweiz im Projekt XXX (blinded for review) zeigten sehr heterogene Ausbildungsmerkmale von Lehrpersonen für Wirtschaft und Gesellschaft (W&G) an Berufsfachschulen (XXX, blinded for review), so dass bei Zusammenhangsanalysen Lehrpersonen in einer sehr heterogenen Gruppe betrachtet werden würden. Dies wirft die weiterführende Frage auf, inwieweit sich Lehrpersonen mit ähnlichem Wissensprofil hinsichtlich ihrer Überzeugungen und Selbstwirksamkeit sowie formaler Lerngelegenheiten (Ausbildungsmerkmale) unterscheiden.

\*Fragestellung\*

1. Können die Lehrpersonen in Klassen unterteilt werden, die sich durch die Kombination der Ausprägungen ihres fachlichen und fachdidaktischen Wissens (d.h. Wissensprofile) unterscheiden lassen?
2. Wie unterscheiden sich die latenten Klassen hinsichtlich ihrer Überzeugungen und Selbstwirksamkeit sowie ihrer Ausbildungsmerkmale?

\*Methode\*

155 Lehrpersonen (71.6% männlich, Alter  $M=46.5$ ,  $SD=9.7$ ) für W&G an kaufmännischen Berufsfachschulen in der Deutschschweiz lösten einen fachlichen (FW) und fachdidaktischen (FDW) Wissenstest und beantworteten Fragen zu Überzeugungen zum Lehren und Lernen und zur Selbstwirksamkeit. Hinsichtlich der Ausbildungsmerkmale hatten 36.0% der Lehrpersonen eine kaufmännische Berufsausbildung absolviert, 78.5% hatten einen universitären Abschluss und 84.3% hatten ein Lehrdiplom für W&G abgeschlossen.

Für die Erfassung des fachlichen Wissens wurden die Testteile für Mikroökonomie (Lösungsquote 60.9%) sowie Makroökonomie (Lösungsquote 73.3%) aus dem von WiwiKom adaptierten TUCE (Zlatkin-Troitschanskaia et al., 2013; 2015) verwendet. Das fachdidaktische Wissen wurde mit einem neu entwickelten Test erfasst (Lösungsquote 52.1%). Als psychologische Merkmale wurden instruktionale und konstruktivistische Überzeugungen (Seifried, 2009) sowie die Selbstwirksamkeit (Schwarzer & Schmitz, 1999) einbezogen.

Auf der Basis des mikroökonomischen, makroökonomischen und fachdidaktischen Wissens wurden durch eine latente Klassenanalyse zunächst Klassen von Lehrpersonen gebildet. Die passendste Klassenlösung wurde nach statistischen sowie inhaltlichen Kriterien ermittelt (vgl. Kunter & Klusmann, 2010). Anschliessend wurde untersucht, ob sich die Klassen hinsichtlich der Überzeugungen und der Ausbildungsmerkmale unterschieden.

\*Ergebnisse\*

Die Ergebnisse der latenten Klassenanalyse suggerierten, dass die Lösung mit vier Klassen die Daten am besten abbildete. Die erste Klasse (18% der Stichprobe) zeichnete sich durch hohe Ausprägungen in allen Wissensfacetten aus („hohes Wissen“: Lösungsquote  $FW_{\text{mikro}}=79.6\%$ ,  $FW_{\text{makro}}=82.5\%$ ,  $FDW=62.7\%$ ). Die zweite Klasse (30%) hatte eine hohe Ausprägung des FDW, während das FW eine mittlere Ausprägung hatte („FDW dominant“:  $FW_{\text{mikro}}=51.9\%$ ,  $FW_{\text{makro}}=71.4\%$ ,  $FDW=59.7\%$ ). Die dritte Klasse (20%) profilierte sich durch ein hohes FW und ein geringes FDW („FW dominant“:  $FW_{\text{mikro}}=67.6\%$ ,  $FW_{\text{makro}}=77.8\%$ ,  $FDW=37.4\%$ ). Die vierte Klasse (32%) wies hingegen tiefe Ausprägungen in allen Wissensfacetten auf („niedriges Wissen“:  $FW_{\text{mikro}}=36.2\%$ ,  $FW_{\text{makro}}=57.0\%$ ,  $FDW=36.7\%$ ). Zusammengefasst konnten sehr heterogene Wissensprofile gebildet werden, die sich durch allgemein hohes Wissen, dominant fachdidaktisches Wissen, dominant fachliches Wissen und allgemein tiefes Wissen charakterisieren liessen.

Der Vergleich dieser Klassen zeigte, dass es keine signifikanten Unterschiede in den instruktionalen und konstruktivistischen Überzeugungen zum Lehren und Lernen gab, jedoch hinsichtlich ihrer Selbstwirksamkeit, die bei Lehrpersonen der Klasse „hohes Wissen“ niedriger ausgeprägt war als in den anderen Klassen. Hinsichtlich der Unterschiede in den Ausbildungsmerkmalen zeigten sich für die Berufsausbildung keine signifikanten Unterschiede. Allerdings hatten in den Klassen „hohes Wissen“ und „tiefes Wissen“ signifikant mehr Lehrpersonen einen Universitätsabschluss als in den anderen Klassen. In der Klasse „tiefes Wissen“ verfügten ausserdem signifikant weniger Lehrpersonen über ein Lehrdiplom als in den anderen Klassen.

## Kompetenzorientiertes Fachwissen von Bewerber\*innen auf ein Mathematik-Lehramtsstudium

**Michael Besser, Denise Depping, Timo Ehmke, Dominik Leiss**

Leuphana Universität Lüneburg, Deutschland; [besser@leuphana.de](mailto:besser@leuphana.de)

Fachwissen von Lehrkräften beeinflusst schulische Lehr-Lern-Prozesse und gilt als notwendige Bedingung für die Entwicklung fachdidaktischen Wissens (Baumert u. a., 2010; Krauss u. a., 2017). Die Herausbildung von profundem, vernetztem, flexibel abrufbarem Fachwissen stellt entsprechend ein zentrales, elementares und wichtiges Ziel professioneller Entwicklung (angehender) Lehrkräfte in allen Phasen der Lehrerbildung dar (Baumert & Kunter, 2013). Verschiedene Meta-Studien der letzten Jahre haben daher analysiert, welche strukturellen Rahmenbedingungen professionelle Entwicklung in der Lehrerbildung positiv beeinflussen (Darling-Hammond, Hyster, & Gardner, 2017; Desimone, 2009; Kennedy, 2016). Weitestgehend unklar ist jedoch, über welches fachliche Vorwissen – das wiederum die Wissensentwicklung und somit die erwünschte professionelle Entwicklung bedingt (Depaepe, Verschaffel, & Kelchtermans, 2013; Kleickmann u. a., 2017) – angehende Lehrkräfte zu Studienbeginn verfügen. Oder etwas plakativ: Welches Fachwissen bringen junge Menschen mit, die ein Lehramtsstudium beginnen wollen?

Das vorliegende Forschungsprojekt greift dieses Desideratum auf und geht am Beispiel des Unterrichtsfachs Mathematik folgender Forschungsfrage nach: Über welches mathematische Fachwissen verfügen Bewerber\*innen auf ein Mathematik-Lehramtsstudium?

Zur empirischen Auseinandersetzung mit dieser Forschungsfrage ist in den Jahren 2015 bis 2017 an der Leuphana Universität Lüneburg ein mathematischer Fachwissenstest als Paper-Pencil-Test entwickelt, an Lehramtsstudierenden im Bachelor Mathematik ( $N=274$ ; zweites bis sechstes Semester) pilotiert und validiert worden. Insgesamt liegen 49 voneinander unabhängige Aufgaben (Items) im Multiple-Choice-Format oder im Complex-Multiple-Choice-Format vor, die sich inhaltlich an den für das Fach Mathematik verbindlichen Bildungsstandards für den Mittleren Schulabschluss (Kultusministerkonferenz, 2003) orientieren. Die Aufgaben erfassen somit explizit kompetenzorientiertes mathematisches Fachwissen der Sekundarstufe I, über welches aus normativer Sicht alle Studienanfänger\*innen eines Mathematik-Lehramtsstudiums verfügen sollten. Empirisch wurden die Aufgaben auf der Bildungsstandardskala für den Mittleren Schulabschluss im Fach Mathematik verankert. Die EAP-Reliabilität dieses Gesamttests über alle 49 Aufgaben beträgt .69, der Gesamttest korreliert signifikant mit der Anzahl an absolvierten Fachsemestern Mathematik und ist somit sensitiv bzgl. der Wissensentwicklung über das Lehramtsstudium ( $r=.24$ ,  $p<.01$ ), es liegen jedoch keine Korrelationen mit Geschlecht, Alter, der Anzahl absolvierter Hochschulsemester oder der Hochschulzugangsberechtigungsnote (Abiturnote) vor. Im Sommer 2017 wurden 16 dieser 49 Aufgaben als kompetenzorientierter Fachwissenstest als verpflichtender Teil des Auswahlverfahrens der Leuphana Universität Lüneburg bei allen Bewerber\*innen für ein Mathematik-Lehramtsstudium administriert ( $N=137$ ). Jede einzelne Aufgabe wird dichotom kodiert, maximal können somit 16 Punkte erreicht werden. Die Testauswertung erfolgt sowohl mittels klassischer (Interne Konsistenz = .62) als auch mittels probabilistischer (WLE-Reliabilität = .62) Testtheorie.

Testauswertungen zeigen folgende Ergebnisse auf: (Auswertungen mittels klassischer Testtheorie) Im Mittel erzielen die Bewerber\*innen  $M=8.2$  ( $SD=2.8$ ) von maximal möglichen 16 Punkten, der Median liegt bei 8 Punkten, der Modalwert bei 7 Punkten. Die empirische Spanne erstreckt sich dabei von  $\text{min}=3$  bis  $\text{max}=15$  korrekt bearbeiteter Aufgaben. Keine Bewerberin bzw. kein Bewerber hat somit alle Aufgaben korrekt gelöst. Die Lösungshäufigkeit der verschiedenen Aufgaben reicht von 10% bis 83%. (Auswertungen mittels probabilistischer Testtheorie) Eine Rasch-Skalierung liefert auf Itemebene Trennschärfen zwischen 0.17 und 0.58, der Weighted Meansquare der Items reicht von 0.91 bis 1.10. Die Fähigkeitswerte der Bewerber\*innen liegen auf der Logit-Skala zwischen -1.68 (entspricht 3 Punkten) und 2.81 (entspricht 15 Punkten). Die Itemschwierigkeiten schwanken zwischen -1.77 und 2.39. Im Mittel können 36% der Bewerber\*innen diejenigen Items, die empirisch der

Kompetenzstufe 4 und 5 (also den beiden höchsten Anforderungsstufen) der Bildungsstandardskala für den Mittleren Schulabschluss zuzuordnen sind, nicht erfolgreich bearbeiten.

Mit Blick auf die aufgezeigte, zentrale Forschungsfrage ist festzuhalten: Mathematisches Fachwissen zu Inhalten der Sekundarstufe I ist bei vielen Bewerber\*Innen für ein Mathematik-Lehramtsstudium nur bedingt ausgeprägt. Mehr als die Hälfte der am Auswahlverfahren teilnehmenden Bewerber\*innen auf ein Mathematik-Lehramtsstudium bearbeitet nicht einmal die Hälfte der administrierten Aufgaben korrekt. Mit Blick auf die besondere Bedeutung ausgeprägten Fachwissens für die Qualität von Schule sind diese Ergebnisse angesichts der aktuellen Diskussionen über einen befürchteten Qualitätsverlust universitärer Ausbildung im Fach Mathematik (Konferenz der Fachbereichstage, 2017; Kritische Stellungnahme, 2017) kritisch zu reflektieren.

## C07

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Bernoullianum 32, H120

### Längsschnittliche Zusammenhänge zwischen Erwartungs-Wert-Komponenten in drei Fächern

**Katrin Arens<sup>1</sup>, Isabelle Schmidt<sup>2</sup>, Preckel Franzis<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland; <sup>2</sup>Universität Trier, Deutschland; [arens@dipf.de](mailto:arens@dipf.de)

Nach der Erwartungs-Wert-Theorie (Eccles & Wigfield, 1995) umfasst die schulische Motivation eine Erwartungskomponente und eine Wertkomponente. Die Erwartungskomponente bildet Selbstwahrnehmungen eigener Fähigkeiten ab und wird oft als akademisches Selbstkonzept operationalisiert. Die Wertkomponente lässt sich in verschiedene Teilfacetten unterteilen, wobei zwischen intrinsischem Wert (Freude, Interesse), subjektiver Relevanz/Wichtigkeit, Nützlichkeit und Kosten/Aufwand unterschieden wird. Die Erwartungs-Wert-Theorie nimmt an, dass die Facetten positive Zusammenhänge untereinander aufweisen, spezifiziert diese Zusammenhänge jedoch nicht näher. In der Tat zeigen querschnittliche Studien Zusammenhänge zwischen der Erwartungskomponente und der Wertkomponente; dieser Zusammenhang fällt besonders hoch zwischen der Erwartungskomponente und intrinsischem Wert aus. Weiter zeigen sich positive Zusammenhänge innerhalb der Wertkomponente selbst und hier insbesondere zwischen intrinsischem Wert und Relevanz. In längsschnittlichen Studien finden sich jedoch häufig keine oder nur schwache Zusammenhänge zwischen der Erwartungskomponente und der Wertkomponente und innerhalb der Wertkomponente. Einige Studien verweisen zwar auf einen Einfluss der Erwartungskomponente auf nachfolgenden intrinsischen Wert (Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005). Dieser Zusammenhang konnte aber in anderen Studien nicht belegt werden (Nurmi & Aunola, 2005; Skaalvik & Valas, 1999; Spinath & Spinath, 2005; Spinath & Steinmayr, 2008).

Die vorliegende Studie untersucht längsschnittliche Zusammenhänge zwischen der Erwartungskomponente und der Wertkomponente und innerhalb der Wertkomponente. Sie erweitert vorhandene Studien, indem sie die Zusammenhänge in drei Fächern (Deutsch, Englisch, Mathematik) betrachtet. Damit können fachspezifische Unterschiede entdeckt werden. Weiter werden zwei Facetten der Wertkomponente, nämlich intrinsischer Wert und Relevanz, berücksichtigt. Zudem deckt die Studie einen langen Zeitraum mit fünf Messzeitpunkten ab.

Es wurden die Daten der Studie „Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter“ (BiKS 8-14; Artelt, Blossfeld, Faust, Roßbach & Weinert, 2013) sekundäranalytisch ausgewertet. Betrachtet werden die letzten fünf Messzeitpunkte, die einmal jährlich von Klassenstufen 5 bis 9 stattfanden. Die hier betrachtete Stichprobe umfasste N=726 (44.9% männlich; 55.1% weiblich) Schüler mit mindestens einem gültigen Wert zu jedem der fünf Messzeitpunkte auf den untersuchten Skalen. Die Erwartungskomponente wurde anhand des akademischen Selbstkonzepts in Deutsch, Englisch und Mathematik erfasst. Der intrinsische Wert und die subjektive Relevanz in den drei Fächern wurden durch je zwei Items erfasst. Die Reliabilitäten der Skalen lagen in allen Fächern zu allen fünf Messzeitpunkten in einem guten Bereich.

Für jedes Fach (Deutsch, Englisch, Mathematik) wurden zuerst konfirmatorische Faktorenanalysen zur Prüfung der Messmodelle durchgeführt. Hier wurden separate Faktoren für jede motivationale Facette (Erwartung/Selbstkonzept, intrinsischer Wert, Relevanz) zu jedem Messzeitpunkt angenommen. Anschließend wurde geprüft, ob die Faktorladungen invariant über die Zeit sind. Dies ist eine notwendige Voraussetzung zur Prüfung von Zusammenhängen zwischen Konstrukten über die Zeit. „Cross-lagged panel“ Modelle (Curran & Bollen, 2001) wurden berechnet, die bei Kontrolle der Stabilitäten der Konstrukte die längsschnittlichen Zusammenhänge zwischen den Konstrukten schätzen. Aufgrund der fünf Messzeitpunkte wurde jeder zeitlicher Zusammenhang zwischen zwei Konstrukten durch vier Pfade (t1-t2; t2-t3; t3-t4; t4-t5) abgebildet.

Die Ergebnisse zeigten eine hohe Stabilität der Konstrukte in allen drei Fächern, wobei das Selbstkonzept die höchsten Stabilitäten aufwies. Die meisten längsschnittlichen Zusammenhänge zwischen den Konstrukten fanden sich in Mathematik: Hier wurden drei Pfade vom vorherigen Selbstkonzept auf nachfolgenden intrinsischen Wert signifikant. Für den Effekt des vorherigen Selbstkonzepts auf die nachfolgende Relevanz wurden zwei Pfade signifikant. Zudem zeigte der intrinsische Wert einen Effekt auf die Relevanz. Effekte des Selbstkonzepts auf intrinsischen Wert und Relevanz fanden sich auch in Englisch, kamen hier aber seltener vor als in Mathematik. In Deutsch zeigte sich nur ein signifikanter Effekt der Relevanz auf den intrinsischen Wert.

Diese Studie erweitert vorhandene Befunde zu längsschnittlichen Zusammenhängen der Komponenten des Erwartungs-Wert-Modells. Die Befunde verweisen auf fachspezifische Zusammenhangsmuster. Insgesamt fallen die längsschnittlichen Zusammenhänge schwach aus; wenn es zu Zusammenhängen kommt, dann sind solche am häufigsten vom Selbstkonzept/der Erwartungskomponente auf den intrinsischen Wert zu finden.

### Große Fische im großen Teich? Die Veränderung des akademischen Selbstkonzepts bei Studienanfängern der Psychologie

**Tom Rosman<sup>1</sup>, Anne-Kathrin Mayer<sup>1</sup>, Günter Krampen<sup>1,2</sup>**

<sup>1</sup>Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID), Deutschland; <sup>2</sup>Universität Trier, Deutschland; [rosman@zpid.de](mailto:rosman@zpid.de)

Theoretischer Hintergrund und Fragestellung

Der Übergang zur Hochschule stellt für viele junge Menschen eine zentrale Entwicklungsaufgabe dar (Coertjens, Brahm, Trautwein, & Lindblom-Ylänne, 2017; Gale & Parker, 2014). Während das selbstregulierte Lernen im Rahmen des Hochschulstudiums zunehmend wichtiger wird (Briggs, Clark, and Hall, 2012) steigen gleichzeitig die Leistungsanforderungen (Trautwein & Bosse, 2017), was häufig mit einem Absinken des Notenspiegels einhergeht (z.B. Abele-Brehm, 2017). Laut des Big Fish Little Pond-Effekts (BFLPE; Marsh, 1987) führt eine solche Transition von einem kleinen zu einem großen „Teich“ zu einem Absinken des akademischen Selbstkonzepts, da sich die Studierenden mit einer leistungsstärkeren Referenzgruppe vergleichen. Entsprechende Effekte sollten bei Psychologiestudierenden besonders ausgeprägt sein, da ihr Notendurchschnitt in der Sekundarstufe in der Regel deutlich über dem Klassendurchschnitt liegt, im Hochschulkontext jedoch nicht mehr (Abele-Brehm, 2017).

Während eine Vielzahl an Längsschnittstudien zur Entwicklung des akademischen Selbstkonzepts im Sekundarschulbereich vorliegen, gibt es wenig entsprechende Erkenntnisse im Hochschulkontext (Jackman, Wilson, Seaton, and Craven 2011; Jackson 2003; Tett, Cree, and Christie 2016). Lohnend erscheinen solche Studien aber allemal, da das akademische Selbstkonzept gerade im Hochschulkontext einen deutlichen Einfluss auf Studienpersistenz (House, 1992) und -leistung

besitzt. Aus diesem Grund stellt der vorliegende Beitrag eine Längsschnittstudie zur Entwicklung des akademischen Selbstkonzepts über die ersten drei Semester des Bachelor-Psychologiestudiums vor. Dabei wurden folgende Forschungsfragen bearbeitet: Gibt es einen BFLPE (also ein Absinken des akademischen Selbstkonzepts) bei Studienanfängern der Psychologie? Moderieren Studiennoten und das Geschlecht die Entwicklung des akademischen Selbstkonzepts?

#### Methode

Zur Prüfung der Forschungsfragen (und den daraus abgeleiteten Hypothesen) wurden Daten einer Längsschnittstudie zur Wissensentwicklung bei BSc-Psychologiestudierenden herangezogen. Das akademische Selbstkonzept wurde dabei an insgesamt vier Messzeitpunkten erfasst: Eine Baseline-Messung fand direkt nach Studieneintritt statt, gefolgt von weiteren Messungen zu Beginn des zweiten, dritten und vierten Semesters. Am ersten Messzeitpunkt nahmen 137 Studierende teil (82 % Frauen;  $M = 20.43$  [ $SD = 2.53$ ] Jahre alt); am vierten Messzeitpunkt verblieben noch 115 Personen in der Stichprobe. Im Rahmen von Dropout-Analysen (Vergleiche zwischen Abbrechern und Personen, die in der Stichprobe verblieben) konnte keine Evidenz für einen systematisch bedingten Dropout gefunden werden. Das akademische Selbstkonzept wurde mit einer Skala zur Erfassung des domänenspezifischen akademischen Selbstkonzepts von Dickhäuser, Schöne, Spinath und Stiensmeier-Pelster (2002) erfasst (5 Items; 7-stufige Likert-Skala; Beispielitem: „Meine studienbezogenen Fähigkeiten sind ... niedrig / hoch“); die empirische Reliabilität (Cronbach's Alpha) ist dabei je nach Messzeitpunkt mit Werten zwischen  $\alpha = .84$  und  $\alpha = .88$  als „gut“ zu bewerten. Studiennoten wurden zum vierten Messzeitpunkt erfasst; dazu wurden die Studierenden gebeten, eine Kopie ihrer Studienverlaufsbescheinigung einzureichen. Im Anschluss wurde der Notendurchschnitt aller schriftlichen Klausuren berechnet und so umkodiert, dass höhere Werte bessere Noten indizieren.

#### Ergebnisse und Diskussion

Zur Datenanalyse wurde auf Mehrebenenanalysen im Rahmen des SPSS MIXED-Befehls zurückgegriffen. Dabei zeigte sich keine signifikante Veränderung im akademischen Selbstkonzept über die Zeit ( $\beta = .013$ ;  $SE = .025$ ;  $p = ns$ ). Nach Aufnahme der Studiennoten (inkl. deren Interaktion mit dem Zeitfaktor) in das Modell zeigte sich hingegen ein signifikanter Interaktionseffekt NOTEN\*ZEIT ( $\beta = -.069$ ;  $SE = .025$ ;  $p < .01$ ): Je besser die Studiennoten, umso eher kommt es zu einem Anstieg des akademischen Selbstkonzepts. Zudem zeigte sich ein signifikanter Interaktionseffekt GESCHLECHT\*ZEIT ( $\beta = .184$ ;  $SE = .061$ ;  $p < .01$ ): Im Vergleich zu Frauen kam es bei männlichen Probanden zu einem deutlich stärkeren Anstieg des akademischen Selbstkonzepts.

Insgesamt legt die vorliegende Studie eine relative Stabilität des akademischen Selbstkonzepts im Hochschulkontext nahe, was gegen einen BFLPE spricht. Der signifikante Moderatoreffekt von Studiennoten spricht jedoch für das internal/external frame of reference-Modell (Parker, Marsh, Lüdtke, & Trautwein, 2013), welches unter anderem einen Vergleich der eigenen Leistung mit externen Standards als ursächlich für Veränderungen im Selbstkonzept ansieht. Der Moderatoreffekt der Variable Geschlecht bestätigt schließlich entsprechende Forschungsarbeiten aus der Gender-Forschung (z.B. Roberts & Nolen-Hoeksema, 1989).

## **Die Rolle von Lehrerverantwortung im Attributionsprozess: Zusammenhänge zwischen Lehrerverantwortung, Schuldgefühlen und der Bereitschaft, Lernende im Kontext von schlechter Leistung zu unterstützen**

**Michael Evers, Fani Lauer mann**

Rheinische Friedrich-Wilhelms-Universität, Bonn, Deutschland; [gmevers@uni-bonn.de](mailto:gmevers@uni-bonn.de)

#### Theorie

Es gehört allgemein hin zu den professionellen Anforderungen an Lehrkräfte, dass sie Verantwortung für ihre Schülerinnen und Schüler übernehmen. Was genau die Implikationen von Verantwortungsübernahme im Lehrberuf sind, ist empirisch kaum belegt (Lauer mann & Karabenick, 2011). Dennoch zeichnet sich in der Literatur eine für Lehrkräfte motivierende Tendenz ab, wenn sie sich verantwortlich fühlen, z.B. in Form von hoher Zielbindung (Bierhoff et al., 2005), der Nutzung von evidenzbasierten Unterrichtsmethoden im Hinblick auf konkrete Ziele (Daniels, Radil, & Wagner, 2016) oder der Bereitschaft, Schülerinnen und Schüler zu unterstützen (Lauer mann, 2014).

Inbesondere im Umgang mit Lernenden, die schlechte Leistungen erbringen und daher besondere Unterstützung benötigen, könnten diese Aspekte eine wichtige Rolle spielen. Die Attributionstheorie von Weiner (1995) ist einer der zentralen Ansätze zur Erforschung solcher Verantwortungsurteile in Unterrichtssituationen. Allerdings stehen in der bisherigen Literatur die Urteile der Lehrkräfte über die Verantwortung der Lernenden an schlechter Leistung im Vordergrund (z.B. Georgiou, Christou, Stavrinides, & Panaoura, 2002; Medway, 1979; Reyna & Weiner, 2001); die Evidenz zu Selbsturteilen über die eigene Verantwortung von Lehrkräften ist dagegen nicht eindeutig (cf. Matteucci, 2007; Matteucci & Gosling, 2004).

Dementsprechend werden in der vorliegenden Studie Selbsturteile von Lehrkräften über die eigene Verantwortung an schlechter SchülerInnenleistung in Weiners Attributionstheorie integriert, um die entsprechenden Konsequenzen zu erforschen. Es wird angenommen, dass Lehrkräfte mit Verärgerung reagieren, wenn sie Lernende für verantwortlich an ihrer schlechten Leistung halten; mit Mitleid, wenn die wahrgenommene Schülerverantwortung gering ist; und mit Schuldgefühlen, wenn sie sich selbst verantwortlich einschätzen. Diese Emotionen können handlungsbestimmend für die Lehrkraft wirken: Bestrafung bei Verärgerung versus Unterstützung bei Mitleid und Schuldgefühlen. Weiterhin wird angenommen, dass Lehrkräfte über ihr Schuldgefühl hinaus aufgrund ihres eigenen Verantwortungsgefühls unterstützend handeln könnten.

Ziel der Studie ist es herauszufinden, (1) inwiefern Lehrerurteile über die eigene Verantwortung deren emotionale Reaktionen (Ärger, Mitleid und Schuldgefühl) vorhersagen können; sowie (2) inwiefern diese emotionalen Reaktionen und die eigene Verantwortung, die Bereitschaft, scheiternde SchülerInnen zu unterstützen, vorhersagen können. Kontrollüberzeugungen und demographische Daten werden als Kontrollvariablen berücksichtigt.

#### Methode

Ein Fragebogen und zwei Vignetten zu Lehrerurteilen über SchülerInnenverantwortung wurden aus Reyna and Weiner (2001) übernommen und um eine Skala zu LehrerInnenverantwortung ergänzt. Die Vignetten präsentieren zwei Situationen schlechter SchülerInnenleistung: in der einen Situation aufgrund von Faulheit, in der anderen Situation aufgrund von mangelnder Fähigkeit. Daraufhin wurden lehrer- und schülerbezogene Kontrollüberzeugungen erfasst, sowie das Empfinden von Ärger, Mitleid und Schuld, und entsprechende Handlungsintentionen (strafendes und unterstützendes Feedback, Bereitschaft für Unterstützung). Lehramtsstudierende ( $N = 224$ ) und erfahrene Lehrkräfte aus drei Schulen ( $N = 134$ ) beteiligten sich an der Studie.

## Ergebnisse

Da die Ergebnisse für beide Situationen ähnlich sind, wird nur die der faulen Schülerin / des faulen Schülers beschrieben. Zu (1) zeigten Regressionsanalysen, dass Urteile über die SchülerInnenverantwortung schülerorientierte Emotionen (Ärger und Mitleid) vorhersagen ( $\beta_{\text{Ärger}} = .181$   $p = .007$ ;  $\beta_{\text{Mitleid}} = -.163$   $p = .015$ ), wohingegen Urteile über die eigene Verantwortung Schuldgefühle vorhersagen ( $\beta = .490$   $p < .001$ ). Dabei wurde für Geschlecht und Lehrerstatus (StudentIn vs. Lehrkraft) kontrolliert. In Bezug auf (2) wurde die Bereitschaft, scheiternde SchülerInnen zu unterstützen, durch Schuldgefühl ( $\beta = .182$   $p = .009$ ) und Urteile über die eigene Verantwortung ( $\beta = .312$   $p < .001$ ) vorhergesagt. Als Kontrollvariablen wurden SchülerInnenverantwortung, weitere Emotionen, Geschlecht und Lehrerstatus berücksichtigt.

Die Ergebnisse unterstützen die Annahmen der Attributionstheorie (Weiner, 1995) und ergänzen diese um die Perspektive der Eigenverantwortung von Lehrkräften. Dabei hat sich gezeigt, dass Lehrerverantwortung eine zentrale Rolle einnimmt in Bezug auf die Bereitschaft von Lehrkräften, Lernende mit schwacher Leistung gesondert zu unterstützen. Dies geschieht nicht nur in Abhängigkeit von Schuldgefühlen, sondern Lehrerverantwortung stellt einen eigenständigen Prädiktor von Hilfsbereitschaft dar (vgl. Lauerermann & Karabenick, 2011).

## Genese und Wirkung aufgabenspezifischen Fähigkeitsselbstkonzepts

**Robert Grassinger**

Pädagogische Hochschule Weingarten, Deutschland; [robert.grassinger@phil.uni-augsburg.de](mailto:robert.grassinger@phil.uni-augsburg.de)

### Theoretischer Hintergrund

Das Fähigkeitsselbstkonzept beschreibt kognitive Repräsentationen eigener Fähigkeiten, die sich auf allgemeinere (z.B. Schule, Knobelaufgaben) oder spezifischere Anforderungsbereiche wie einzelne Schulfächer (z.B. Mathematik) oder Aufgaben (z.B. Bruchrechnen, Sudoku) beziehen und beispielsweise die Kurs- oder Aufgabenwahl (Köller, Trautwein, Lüdtke & Baumert, 2006; Marsh & Yeung, 1998; Meyer, 1983) oder die Persistenz und die Anstrengung bei der Bearbeitung einer Aufgabe (Gordon, 1997; Perels, Löb, Schmitz, & Haberstroh, 2006; Trautwein, Lüdtke, Schnyder, & Niggli, 2006) beeinflussen. Vor diesem Hintergrund waren vielfältige Arbeiten durch die Fragen inspiriert, wie Personen zu ihrem Fähigkeitsselbstkonzept gelangen und wie dieses strukturiert ist. Das Internal/External Frame-of-Reference-Modell (Marsh, 1986) postuliert, dass Lernende ihre Leistungen in einer Domäne (z.B. Mathematik) mit der Leistung anderer in dieser Domäne (external, sozialer Vergleich) sowie mit der eigenen Leistung in anderen Domänen wie Französisch (internal, dimensionaler Vergleich) vergleichen und so auf ihre eigenen Fähigkeiten schließen. Dickhäuser und Galfe (2004) argumentieren darüber hinaus, dass bisherige Erfahrungen in der Bearbeitung domänenspezifischer Aufgaben memoriert werden und temporale Leistungsvergleiche die Ausprägung des domänenspezifischen Fähigkeitsselbstkonzepts beeinflussen. Das Nested Marsh/Shavelson-Modell (Brunner et al., 2010) integriert bisherige theoretische Überlegungen sowie empirische Befunde zur Frage der Struktur des Fähigkeitsselbstkonzepts und postuliert eine hierarchische Struktur der kognitiven Repräsentationen eigener Fähigkeiten. Demnach beeinflusst ein allgemeineres Fähigkeitsselbstkonzept das Fähigkeitsselbstkonzept einer jeweiligen Subdomäne bzw. einer subdomänenspezifischen Aufgabe. Als vorwiegend handlungsrelevant wird dabei das spezifischere Fähigkeitsselbstkonzept angenommen (Faber, 1992; Marsh & Craven, 1997, Rost & Sparfeldt, 2002).

### Fragestellung

Weniger untersucht ist der gesamte Prozess von der Genese eines aufgabenspezifischen Fähigkeitsselbstkonzepts bei der Konfrontation mit einer Aufgabe hin zu dessen Wirkung auf Lern- und Leistungshandlungen bei dieser Aufgabe. Auch ist unklar, inwieweit ein allgemeineres Fähigkeitsselbstkonzept in Ergänzung zu sozialen, dimensional und temporalen Vergleichen die Ausprägung eines spezifischeren Fähigkeitsselbstkonzepts beeinflusst. Gemäß den Theorien zur Selbsterkenntnis (Brown, 1990; Sedikides, 1993; Sedikides & Strube, 1997; Trope, 1986) und zur Selbstverifikation (Swann, Rentfrow & Guinn, 2002) sind Menschen daran interessiert, ein möglichst genaues und zutreffendes Bild der eigenen Person zu gewinnen. So sollten möglichst vielfältige Informationen bei der Genese des Fähigkeitsselbstkonzepts Berücksichtigung finden. Entsprechend wird erwartet, dass soziale, dimensionale und temporale Vergleiche sowie ein allgemeineres Fähigkeitsselbstkonzept die Ausprägung eines aufgabenspezifischen Fähigkeitsselbstkonzepts beeinflussen und Letztgenanntes bedeutsam mit der Aufgabenwahl und der Persistenz assoziiert ist.

### Methodik

Zur Prüfung der Hypothesen wurden 167 Lehramtsstudierende mit der Aufgabe konfrontiert, zwischen einem leichten oder schweren Sudoku zu wählen und dieses zu lösen. In der Tat waren beide Sudokus identisch und unlösbar. Nach der Wahl und bevor den Studierenden das gewählte Sudoku ausgeteilt wurde, beantworteten diese einen Fragebogen. Nach zehn Minuten Bearbeitungszeit erhielten die Studierenden die Gelegenheit, ein neues Sudoku zu wählen. Das aufgabenspezifische Fähigkeitsselbstkonzept für Sudoku wurde in Analogie zu Brunner et al (2010) als correlated trait-correlated method minus one – Modell (CT-C(M\_1) spezifiziert. Hierbei luden sämtliche Items der Vergleiche und des allgemeineres Fähigkeitsselbstkonzepts auf das aufgabenspezifische Fähigkeitsselbstkonzept. Zudem wurden die Methodenfaktoren sozialer Vergleich, dimensionaler Vergleich und temporaler Vergleich modelliert. Für Items zum allgemeineren Fähigkeitsselbstkonzept wurde in Analogie zu Brunner et al. (2010) kein Methodenfaktor spezifiziert, diese fungierten als „method minus one“. Das so modellierte aufgabenspezifische Fähigkeitsselbstkonzept wurde zur empirischen Prüfung der Annahme logarithmisch auf die Aufgabenwahl (leicht, schwer) sowie die Persistenz (ja, nein) regressiert.

### Ergebnisse

Übereinstimmend zur Annahme nutzten die Studierenden zur Generierung ihres aufgabenspezifischen Fähigkeitsselbstkonzepts vielfältige Informationen, die aufgrund des allgemeineren Fähigkeitsselbstkonzepts sowie der sozialen, temporalen und zum Teil dimensional Vergleiche zur Verfügung standen. Dieses so generierte aufgabenspezifische Fähigkeitsselbstkonzept wiederum beeinflusste die Aufgabenwahl und die Persistenz. Die Befunde geben einen Hinweis darauf, dass für die Genese eines aufgabenspezifischen Fähigkeitsselbstkonzepts neben sozialen, temporalen und dimensional Vergleichen auch top-down-Vergleiche relevant sind.

## C08

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Bernoullianum 32, H103

### Vergleich der prädiktiven Validität von Lerntagebuch und Fragebogen zur Erfassung von Lernstrategien im Studium

**Natalie Enders, Stefan Venekamp, Christian Weinzierl**

Leibniz Universität Hannover, Deutschland; [enders@psychologie.uni-hannover.de](mailto:enders@psychologie.uni-hannover.de)

Selbstregulatorischen Fähigkeiten wird sowohl als Voraussetzung als auch als Resultat für erfolgreiches akademisches Lernen hohe Bedeutung beigemessen (Steuer et al., 2015). Ihre Förderung wurde in der Konsequenz von der KMK (2005) in einem Qualifikationsrahmen für deutsche Hochschulabschlüsse beschlossen.

Lernstrategien sind zentrale Bestandteile theoretischer Modelle selbstregulierten Lernens (Boekaerts, 1999; Schiefele & Pekrun, 1996; Zimmerman, 2000). Nach Weinstein und Mayer (1986) werden kognitive Strategien (Wiederholung, Elaboration, Organisation), metakognitive Strategien und Strategien des Ressourcenmanagements unterschieden. Obwohl ihnen als Determinanten erfolgreichen akademischen Lernens hohe Bedeutung beigemessen wird (Wild, 2000), zeigen sich lediglich niedrige bis mittlere Korrelationen mit Leistungsindikatoren (Streblo & Schiefele, 2006; Wild, 2000, 2005). Als Ursache werden unzulängliche Erfassungsmethoden diskutiert. An Lernstrategiefragebögen wird kritisiert, dass sie habituelles Lernverhalten erfragen und eine geringe Passung zu konkreten Lernhandlungen aufweisen (Artelt, 2000). Weiterhin ist fraglich, ob mit Fragebögen nicht lediglich Strategiewissen erfasst wird (Lompscher, 1996; Spörer & Brunstein, 2006). Eine alternative Erfassungsmethode von Lernstrategien stellen Lerntagebücher dar. Im Vortrag wird der Fragestellung nachgegangen, ob sich ein Lernstrategiefragebogen oder ein Lerntagebuch besser zur Vorhersage der Lernleistung Lehramtsstudierender eignet.

An drei aufeinanderfolgenden Semestern nahmen n=113 Studierende an einer E-Learning-Veranstaltung zu pädagogisch-psychologischen Grundlagen für Lehramtsstudierende teil. Zu Semesterbeginn (MZP1) wurden die Lernstrategien der Studierenden mittels LIST-Fragebogen (Wild & Schiefele, 1994) erfasst. Anschließend bearbeiteten die Studierenden innerhalb vierzehntägiger Lerneinheiten (MZP2–5) verschiedene Softwarekapitel und reichten zu jeder Einheit jeweils eine schriftliche Hausaufgabe ein. Zusätzlich dokumentierten sie ihre Lernaktivitäten beim E-Learning und bei der Vorbereitung auf die Abschlussklausur (MZP6) in einem standardisierten Lerntagebuch mit offenem Antwortformat.

Als Indikatoren für die Nutzungshäufigkeit der fünf verschiedenen Lernstrategietypen (Weinstein & Mayer, 1986) wurden entsprechende Subskalen des LIST sowie die Lerntagebucheinträge herangezogen. Letztere wurden von zwei unabhängigen Ratern segmentiert und hinsichtlich der Lernstrategieklassifikation von Weinstein und Mayer (ebenda) kodiert (Interrater-Reliabilitäten:  $.59 < \text{Kendalls Tau} < .97$ ).

Die Ergebnisse zeigen, dass die im Lerntagebuch selbstberichtete Lernzeit zu keinem Messzeitpunkt signifikante Korrelationen mit den Leistungsmaßen (Hausaufgaben bzw. Klausur) aufweist ( $.01 < r < .18$ ). Allerdings bestehen zwischen der Anzahl der mittels Lerntagebuch erfassten Lernstrategien signifikante positive Korrelationen mit den Leistungsmaßen ( $.02 < r < .46^{***}$ ), welche die Zusammenhänge der LIST-Skalen mit den Leistungsmaßen ( $-.14 < r < .24^*$ ) in Anzahl und Höhe deutlich übertreffen.

In einer hierarchischen Regressionsanalyse zur Vorhersage der erreichten Punkte in der Abschlussklausur wurden als Prädiktoren in drei Schritten (1) der Lernerfolg während des Semesters (Gesamtpunktzahl aller Hausaufgaben), (2) die zur Klausurvorbereitung investierte Lernzeit sowie (3) die anhand des Lerntagebuchs zur Klausurvorbereitung ermittelten absoluten Nutzungshäufigkeiten der fünf Lernstrategietypen eingefügt. Der bisherige Lernerfolg klärt im ersten Regressionsmodell 14% der Varianz der Lernleistung auf [ $R^2 = .14$ ,  $F(1,89) = 15.01$ ,  $p < .001$ ]. Die Lernzeit im zweiten Modell erbringt kein signifikantes Inkrement an aufgeklärter Varianz, durch Aufnahme der kognitiven Lernstrategien im dritten Modell ergibt sich jedoch ein deutlicher signifikanter Zuwachs der aufgeklärten Varianz auf insgesamt 32% [ $\Delta R^2 = .18$ ,  $F(7,83) = 5.57$ ,  $p < .001$ ], wobei sich Wiederholungs- und Elaborationsstrategien als signifikante Prädiktoren erweisen.

In einer analogen hierarchischen Regressionsanalyse, in welcher im dritten Schritt die LIST-Skalen als Indikatoren für die Lernstrategienutzung eingefügt wurden, können im letzten Schritt zwar 25% der Varianz in der Klausurleistung aufgeklärt werden [ $R^2 = .25$ ,  $F(7,82) = 3.85$ ,  $p < .001$ ]. Allerdings liefert in allen drei Modellen lediglich der bereits in Schritt 1 hinzugefügte bisherige Lernerfolg einen signifikanten Beitrag zur Vorhersage der Klausurleistung. Weder die Lernzeit (Schritt 2) noch die LIST-Skalen (Schritt 3) führen zu einer statistisch signifikanten Vorhersageverbesserung.

Insgesamt sprechen die Ergebnisse für eine bessere prädiktive Validität des Lerntagebuchs als des LIST-Fragebogens. Zusätzlich liefern die anhand der Lerntagebücher gewonnenen Daten interessante Informationen über den Zeitaufwand und den Lernstrategieeinsatz der Studierenden im Semesterverlauf. Dies erlaubt beispielsweise eine differenzierte Betrachtung des strategischen Lernens bei der Vorbereitung auf unterschiedliche Prüfungsanlässe (Hausaufgabenerstellung vs. Klausurvorbereitung). Zusammenfassend werden die Vor- und Nachteile der unterschiedlichen Erfassungsmethoden (z.B. Ökonomie, Gütekriterien, Reaktivität) gegeneinander abgewogen und das Potenzial einer multimethodalen Erfassung diskutiert.

### Einfluss motivationaler Selbstregulation auf die Leistung, Zufriedenheit und Abbruchneigung von Lehramtsstudierenden – welche Motivationsregulationsstrategien sind erfolgsversprechend?

**Olena Kryshko, Jens Fleischer, Detlev Leutner**

Universität Duisburg-Essen, Deutschland; [olena.kryshko@uni-due.de](mailto:olena.kryshko@uni-due.de)

Die Sicherstellung einer erfolgreichen akademischen Ausbildung zukünftiger Lehrkräfte ist besonders in Zeiten des Lehrermangels (Fokken, 2017) ein dringendes bildungspolitisches Anliegen. Da motivationale Defizite neben Leistungsproblemen zu den wichtigsten Studienabbruchgründen von Lehramtsstudierenden zählen (Heublein et al., 2009), können Erkenntnisse aus der Forschung zur motivationalen Selbstregulation gewinnbringende Ansätze zur Steigerung der Studienerfolgsquoten liefern.

Frühere Untersuchungen haben gezeigt, dass Studierende unterschiedliche Strategien nutzen, um ihre Motivation beim Lernen zu erhöhen bzw. aufrechtzuerhalten (Schwinger et al., 2007; Wolters, 1998; Wolters & Benzon, 2013). Es gibt außerdem erste Hinweise dafür, dass sich bestimmte motivationale Selbstregulationsstrategien günstig auf die Anstrengungsbereitschaft, das Wohlbefinden sowie die Leistungen im Studium auswirken (Grunschel et al., 2016; Schwinger & Otterpohl, 2017). Für eine motivationsregulatorische Lernstrategie konnten Leutner et al. (2001) zeigen, dass deren zielführende Anwendung erfolgreich trainiert werden kann. Ausgehend von diesen Befunden wird in der vorliegenden Studie die Relevanz von neun motivationalen



Selbstregulationsstrategien für ein erfolgreiches Lehramtsstudium untersucht, um effektive Strategien für die Konzeption eines Trainings motivationaler Selbstregulation abzuleiten.

Es wird von einem positiven Einfluss motivationaler Selbstregulation auf den Studienerfolg ausgegangen und überprüft, ob sich dieser auch bei Kontrolle der Abiturnote, die als stärkster Einzelprädiktor für Studienerfolg gilt, nachweisen lässt. Es wird also der Frage nach dem inkrementellen Beitrag motivationaler Selbstregulationsstrategien zum Studienerfolg nachgegangen, der in der bisherigen Forschung unzureichend berücksichtigt worden ist. Dabei werden die Studienleistung, die Abbruchneigung sowie die Studienzufriedenheit als Kriterien für Studienerfolg herangezogen.

Die erhobenen Daten stammen aus einer Querschnittsstudie, die im Rahmen einer Pflichtvorlesung für Lehramtsbachelorstudiengänge im Sommersemester 2017 an einer nordrhein-westfälischen Universität durchgeführt wurde. Die Teilnahme an der Studie war freiwillig und wurde nicht entlohnt. Befragt wurden 249 Bachelorstudierende mit unterschiedlichen Fächerkombinationen, die sich durchschnittlich im 3. Semester ( $SD=1.88$ ) befanden und das Lehramt an unterschiedlichen Schulformen anstrebten. Der Frauenanteil lag bei 66.3 %, und das Alter betrug im Mittel 21.6 Jahre ( $SD=3.72$ ). Die Befragten hatten ihre Hochschulzugangsberechtigung im Durchschnitt mit einer Abiturnote von 2.4 ( $SD=0.55$ ) erworben.

Die Operationalisierung motivationaler Selbstregulation erfolgte mit dem Fragebogen von Schwinger et al. (2007), der die Nutzungshäufigkeit von acht Motivationsregulationsstrategien anhand einer 5-stufigen Skala erfasst: Leistungszielbezogene Selbstinstruktion (5 Items,  $\alpha=.81$ ), Selbstbelohnung (4 Items,  $\alpha=.84$ ), Teilziele setzen (3 Items,  $\alpha=.74$ ), Umweltkontrolle (3 Items,  $\alpha=.62$ ), Steigerung der persönlichen Bedeutsamkeit (3 Items,  $\alpha=.76$ ), Lernzielbezogene Selbstinstruktion (4 Items,  $\alpha=.71$ ), Steigerung des situationalen Interesses (5 Items,  $\alpha=.80$ ) sowie Vermeidungs-leistungszielbezogene Selbstinstruktion (3 Items,  $\alpha=.75$ ). Zudem wurde eine weitere Motivationsregulationsstrategie (adaptiert aus McCann & Garcia, 1999) ergänzt: Selbstwirksamkeitsmanagement (5 Items,  $\alpha=.65$ ).

Die Studienleistung wurde anhand des selbstberichteten Notendurchschnitts operationalisiert. Die Erfassung der Studienabbruchneigung (6 Items,  $\alpha=.86$ ) erfolgte in Anlehnung an Dresel und Grassinger (2013) sowie Ströhlein (1983). Schließlich wurden zwei Komponenten der Studienzufriedenheit (Zufriedenheit mit Studieninhalten: 3 Items,  $\alpha=.83$ ; Zufriedenheit mit der Bewältigung der Studienbelastungen: 3 Items,  $\alpha=.81$ ) mit dem Fragebogen von Westermann et al. (1996) erfasst.

Die Ergebnisse der Regressionsanalysen belegen den Effekt der Nutzung motivationaler Selbstregulationsstrategien auf unterschiedliche Aspekte des Studienerfolgs. So lässt sich beispielsweise sowohl für den Strategiegesamtscore als auch für sieben von neun einzelnen Motivationsregulationsstrategien ein signifikanter positiver Einfluss auf den Notendurchschnitt nachweisen. Dieser Einfluss bleibt für den Gesamtscore ( $\beta=-.323$ ,  $p<.001$ ) sowie für folgende fünf Einzelstrategien auch nach Kontrolle der Abiturnote statistisch bedeutsam: Lernzielbezogene Selbstinstruktion ( $\beta=-.309$ ;  $p<.001$ ), Selbstwirksamkeitsmanagement ( $\beta=-.241$ ,  $p=.003$ ), Leistungszielbezogene Selbstinstruktion ( $\beta=-.196$ ;  $p=.016$ ), Steigerung der persönlichen Bedeutsamkeit ( $\beta=-.195$ ;  $p=.010$ ) und Steigerung des situationalen Interesses ( $\beta=-.163$ ,  $p=.044$ ). Zudem erweist sich der Zuwachs an aufgeklärter Varianz für diese Strategien und den Gesamtscore als statistisch signifikant. Die Befunde zu den weiteren untersuchten Studienerfolgskriterien fallen ebenfalls überwiegend erwartungskonform aus und werden im Vortrag präsentiert.

Somit liefern die erzielten Ergebnisse weitere Hinweise zur Bedeutsamkeit motivationaler Selbstregulation im Studium und werden im Hinblick auf die Auswahl effektiver Motivationsregulationsstrategien für die Entwicklung möglicher Interventions- bzw. Trainingsmaßnahmen für Studierende diskutiert.

## **Originalartikel versus Skript: Überzeugungen Studierender über den Nutzen unterschiedlicher Informationsquellen im Studium**

**Sebastian Schmid, Sabrina Viehauser, Klaus-Peter Wild**  
Universität Regensburg, Deutschland; [seb.schmid@ur.de](mailto:seb.schmid@ur.de)

### Theoretischer Hintergrund

Eine Besonderheit des universitären Lernens besteht darin, dass Studierende vergleichsweise selbständig darüber entscheiden können, welche Informationsquellen sie zum Lernen heranziehen. Das Spektrum potenziell relevanter Informationsquellen umfasst Lehrbücher ebenso wie wissenschaftliche Originalliteratur und populärwissenschaftliche Literatur (Goldman & Bisanz, 2002). Hinzu kommen informelle Informationsquellen wie Skripte und Mitschriften anderer Studierender. Umfassende Rahmenmodelle zum Umgang mit multiplen Dokumenten (z.B. MD-TRACE; Rouet & Britt, 2011) konzipieren die Auswahl von Informationsquellen und ihre kognitive Verarbeitung als zielgerichteten Prozess, bei dem mehrere Relevanzdimensionen berücksichtigt werden. Neben der thematischen Relevanz spielt eine Rolle, ob ein Dokument zu einem Genre gehört, von dem Lernende erwarten, dass es geeignet ist, ihre jeweiligen Ziele zu erreichen. Es ist wenig darüber bekannt, welche Instrumentalitätsüberzeugungen Studierende mit den Genres verbinden, die sie im Studium nutzen. Wir betrachten diese Überzeugungen als einen Bestandteil des Genrewissens Studierender, der sich im Laufe ihrer akademischen Sozialisation zunehmend ausdifferenziert (vgl. Rouet, Favart, Britt, & Perfetti, 1997). Ein Ziel der Untersuchung bestand darin herauszufinden, welche Instrumentalität Studierende unterschiedlichen Quellen für das Erreichen der Ziele "Inhalte wiedergeben können" (im Sinne von Oberflächenlernen) und "ein tieferes Verständnis gewinnen" (im Sinne von Tiefenlernen) zuschreiben. Ein weiteres Ziel bestand darin zu untersuchen, ob und wie diese Instrumentalitätsüberzeugungen mit der Dauer des bisherigen Studiums zusammenhängen.

### Methode

Untersuchungsteilnehmer(innen) waren 452 Studierende (63.1% weiblich, 34.7 männlich) naturwissenschaftlicher (51.5%), sozialwissenschaftlicher (24.1%) und geisteswissenschaftlicher (24.3%) Fächer in ihrem ersten (43.5), zweiten (30.6%) und dritten (25.9%) Studienjahr. Die Studierenden bearbeiteten eine Liste von 16 Genres unter mehreren Gesichtspunkten. Die Auswahl der Genres orientierte sich an der Taxonomie von Goldman und Bisanz (2002) und den Ergebnissen einer qualitativen Vorstudie. Sie umfasste wissenschaftliche Originalliteratur (z.B. Artikel in Fachzeitschriften) ebenso wie didaktisch orientierte Literatur (z.B. Lehrbuchtexte), informelle Literatur (z.B. Skripte), populärwissenschaftliche Literatur (z.B. Sachbücher für die breite Öffentlichkeit) und internetspezifische Quellen (z.B. interaktive Plattformen). Die Studierenden schätzten jedes Genre u.a. danach ein, (a) für wie geeignet sie es für das Ziel halten, Inhalte wiedergeben zu können, und (b) für wie geeignet sie es für das Ziel halten, ein tieferes Verständnis der Inhalte zu gewinnen (Instrumentalitätsratings mit den Antwortkategorien ungeeignet, eher ungeeignet, eher geeignet, geeignet und sehr geeignet).

## Ergebnisse

Um zunächst zu untersuchen, welche Genres die Studierenden für besonders geeignet bzw. ungeeignet halten, um die jeweiligen Ziele zu erreichen, wurden genrespezifische Abweichungen vom mittleren Instrumentalitätsrating des jeweiligen Ziels auf Signifikanz geprüft. Mit Blick auf das Ziel „Inhalte wiedergeben“ ergaben sich positive Abweichungen für informelle Genres (z.B. Skripte und Mitschriften anderer Studierender und Materialien zu früheren Prüfungen), aber auch für Lehrbücher und Nachschlagewerke, und negative Abweichungen für Artikel in Fachzeitschriften. Mit Blick auf das Ziel „ein tieferes Verständnis gewinnen“ ergaben sich negative Abweichungen für informelle Genres, und positive Abweichungen für wissenschaftliche Originalliteratur (z.B. Monographien, Beiträge in Herausgeberwerken und Artikel in Fachzeitschriften). In einer weiteren Analyse wurden Instrumentalitätsratings zu beiden Zielen miteinander kontrastiert. Hierbei ergab sich, dass informelle Genres für geeigneter gehalten wurden, um das Ziel zu erreichen, Inhalte wiederzugeben, als um das Ziel zu erreichen, ein tieferes Verständnis zu gewinnen. Für wissenschaftliche Originalliteratur ergab sich ein spiegelbildlicher Befund. Schließlich wurde geprüft, ob Studierende in unterschiedlichen Studienabschnitten unterschiedliche Instrumentalitätsprofile aufweisen. Für beide Zieltypen ergab sich eine signifikante Interaktion der Faktoren Genre und Studienjahr. Studierende im zweiten und dritten Studienjahr schätzten die Instrumentalität wissenschaftlicher Originalliteratur unabhängig vom Zieltyp höher ein als Studierende im ersten Studienjahr. Gleichzeitig betrachteten sie informelle Genres und populärwissenschaftliche Literatur als weniger geeignet, um ein tieferes Verständnis zu gewinnen. Insgesamt sprechen die Befunde dafür, dass Studierende differenzierte Überzeugungen über Instrumentalität unterschiedlicher Informationsquellen besitzen, die sich im Verlauf des Studiums weiter ausdifferenzieren scheinen. In der Diskussion wird auch auf Domänenunterschiede und Befunde zur Nutzungshäufigkeit verschiedener Genres eingegangen.

## **Muss die Studienleistung immer so langweilig sein? Generieren von Fragen als Wünschenswertes Erschwernis im universitären Unterricht**

**Roman Abel<sup>1</sup>, Matthias Mai<sup>2</sup>**

<sup>1</sup>Universität Kassel, Deutschland; <sup>2</sup>Leibniz Universität Hannover, Deutschland; [r\\_abel@uni-kassel.de](mailto:r_abel@uni-kassel.de)

### Hintergrund und Fragestellung:

Das Generieren von Fragen gilt als eine effektive Strategie, um elaborative Lernprozesse anzuregen (Foos, Mora & Tkacz, 1994; Rosenshine, Meister & Chapman, 1996; Bugg & McDaniel, 2012). Diese Lernmethode wurde bereits in experimentellen Untersuchungen im Rahmen universitären Unterrichts erfolgreich angewandt (Foos et al., 1994; Levin & Arnold, 2008). Es fehlen jedoch empirische Erkenntnisse zu der systematischen Implementation von Generierung von Fragen als Studienleistung im Seminar. Die Hauptfrage, die in diesem Artikel adressiert werden soll, lautet somit: Wie gut funktioniert das Generieren von Fragen im Vergleich zu den traditionellen Studienleistungen in universitären Lehrveranstaltungen. Kann diese Methode die Studierenden dabei unterstützen, Inhalte zu lernen, die relevant sind fürs Studium?

Die herkömmlichen Studienleistungen in Seminaren sind durch die Verfügbarkeit von zu erarbeitenden Lernmaterialien für alle KursteilnehmerInnen charakterisiert: Präsentationen, Handouts, Protokolle. Die Frage nach optimalen Lernergebnissen sollte folglich sowohl die Erbringer der Studienleistung als auch die Empfänger der Lerninhalte berücksichtigen. So könnten die von den Studierenden generierten Fragen zum Thema der letzten Unterrichtseinheit in Form eines Quiz allen KursteilnehmerInnen zur Verfügung gestellt werden. Im Gegensatz zu Lernmaterialien, die im Rahmen von herkömmlichen Studienleistungen – Präsentationen und Protokolle – erarbeitet werden und nicht unbedingt zu einer tiefen Verarbeitung der Lerninhalte bei Studierenden führen, kann das Beantworten von Fragen als Abruf-Übung begriffen werden (Jonge, Tabbers & Rikers, 2015; Karpicke & Zaroomb, 2010). Der Abruf regt die Metakognition an und damit das Bewusstsein der Lernenden über die individuellen Wissenslücken (Thiede, Anderson & Theriault, 2003). Das Generieren von Fragen als Studienleistung zum Erlangen von Credits ist insbesondere deshalb eine attraktive Methode, da sie sowohl für die Studierenden, die die Fragen generieren als auch für diejenigen, die sie beantworten, positive Effekte verspricht. Erstere profitieren vom sogenannten Generierungseffekt (Bertsch, Pesta, Wiscott & McDaniel, 2007) während letztere ihren Nutzen aus dem Testungseffekt ziehen können (Karpicke & Zarambo, 2010).

### Methode:

An unserer Studie nahmen 87 PsychologiestudentInnen aus vier inhaltlich äquivalenten Seminaren zur Pädagogischen Psychologie teil. Das inhaltliche Thema sowie die Art der Studienleistung wurden den Studierenden zufällig zugeteilt. Die TeilnehmerInnen der Experimentalbedingung bekamen den Auftrag, Fragen zu den Inhalten der vorherigen Sitzung zu generieren, welche von den Mitstudierenden in der darauffolgenden Sitzung im Rahmen eines Quiz beantwortet wurden. Darüber hinaus wurden zwei Kontrollgruppen gebildet, die den Arbeitsauftrag bekamen 1. Ein Protokoll über die Inhalte der vorherigen Stunde zu schreiben 2. Eine Präsentation zu einem angrenzenden Thema zu halten. Der unangekündigte finale Test setzte sich zusammen aus Aufgaben aller inhaltlichen Sitzungen.

### Ergebnisse und Diskussion:

Die Beschäftigung mit den Lerninhalten beim Erbringen der Studienleistung führte per se zu einem tieferen Verständnis der zu verarbeitenden Inhalte  $F(1,83)=4.60, p=.035$ . Wichtiger ist jedoch, dass sich die Lernbedingungen nicht nur im Hinblick auf die durchgearbeiteten Inhalte unterschieden, sondern ebenso im Hinblick auf die rein rezipierten Inhalte  $F(2,83)=5.26, p=.007$ . Daraus lässt sich schlussfolgern, dass die Art der zu erbringenden Studienleistung nicht nur das intentionale, sondern ebenso das inzidentelle Lernen beeinflusst. Wir konnten unsere anfängliche Hypothese nicht bestätigen: Nicht das Generieren von Fragen, sondern das Verfassen von Protokollen erwies sich als die effektivste Lernmethode. Die Studierenden in der Präsentationsbedingung zeigten dagegen die schlechtesten Lernergebnisse.

Eine mögliche Erklärung für die Überlegenheit des Protokollschreibens ist, dass die Studierenden beim Verfassen aller wichtigen Informationen und Kernkonzepte abdecken müssen, während das Generieren von Fragen zu einer eher fragmentierten Elaboration der Lerninhalte führt. Darüber hinaus könnte der Generierungseffekt – beim Erstellen von Fragen – durch den ausgleichenden Testungseffekt bei Mitstudierenden, welche die Fragen beantworteten, verschleiert worden sein (Weinstein, McDermott & Roediger, 2010).

In der Nachfolgestudie sollen die Effekte, die auf Generierung und Testung zurückgeführt werden, entkoppelt sowie die Qualität der zu generierenden Fragen durch Unterstützungsmaßnahmen gesteigert werden.

## C09

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Bernoullianum 32, H223

### **Akute Effekte einer Bewegungspause oder eines Achtsamkeitstrainings auf Befindlichkeit, Aufmerksamkeit und exekutive Funktionen**

**Christian Müller<sup>1</sup>, Manuel Lieb<sup>1</sup>, David Dubiel<sup>1</sup>, Marie Schnell<sup>1</sup>, Eleni Kremeti<sup>2</sup>, Marina Vrancic<sup>1</sup>, Nour Siakir Oglou<sup>3</sup>, Christoph Mickel<sup>1</sup>, Julia Karbach<sup>4</sup>**

<sup>1</sup>Goethe-Universität, Frankfurt, Deutschland; <sup>2</sup>Philipps-Universität Marburg, Deutschland; <sup>3</sup>Ege Universität, Izmir, Türkei;

<sup>4</sup>Universität Koblenz Landau, Deutschland; [mueller@psych.uni-frankfurt.de](mailto:mueller@psych.uni-frankfurt.de)

Aufmerksamkeit und Wohlbefinden bedingen einerseits Lernprozesse (Allport, 1980; Kort, Reilly, & Picard, 2001), schwanken aber insbesondere in Lehr-Settings erheblich (Bunce, Flens, Neiles, 2010). Eine Methode Aufmerksamkeitsprozesse und Wohlbefinden positiv zu beeinflussen sind körperliche Aktivierung (Hillman, 2008) und zunehmend auch Achtsamkeitstrainings (Mak, Whittingham, Cunnington, & Boyd, 2017). Die Studienergebnisse hierzu sind jedoch heterogen und innerhalb universitärer Lehr-Settings bisher kaum untersucht worden (Carmody & Baer, 2009; Tomporowski, 2003). In den vorliegenden Studien wurde daher untersucht, inwieweit sich akute Effekte einer Bewegungspause (Studie 1) oder eines Achtsamkeitstrainings (Studie 2) auf Aufmerksamkeitsprozesse und Wohlbefinden nach einer Sitzung innerhalb eines universitären Lernsettings nachweisen lassen.

Design:

Für Studie 1 (Bewegungspause) und 2 (Achtsamkeitstraining) wurde je ein randomisiertes within-subject-cross-over-Design über zwei Tagen gewählt. Am ersten Tag wurden die Teilnehmer/innen randomisiert in Interventions- oder Kontrollgruppe eingeteilt. Am darauffolgenden Tag wurden die Gruppen getauscht, um Reihenfolgeeffekte und Stichprobenunterschiede zu berücksichtigen. Beide Studien wurden in Blockseminaren für Lehramtsstudierende durchgeführt.

Stichprobe:

In Studie 1 (Bewegungspause) nahmen 68 Personen (Alter M = 25.85 Jahre, SD = 5.06 Jahre, weiblich = 75 %) und in Studie 2 (Achtsamkeitstraining) 28 Personen (Alter M = 23.69 Jahre, SD = 3.11 Jahre, weiblich = 67.86 %) teil.

Bewegungspause/Achtsamkeitstraining:

Die in Kooperation mit Sportwissenschaftlern entwickelte standardisierte Bewegungspause, bestand aus den Phasen: Warm-up, Partner-Balanceübungen, mittlerer kardiovaskulärer Aktivierung und Cool-down.

Das standardisierte Achtsamkeitstraining bestand aus einer kurzen Streck-/Atemübung und einem Body-scan. Beim Body-scan saßen die Teilnehmer/innen auf einer Matte und wurden angeleitet, ihre Aufmerksamkeit auf den eigenen Körper und die eigene Atmung zu richten, ohne zu werten oder abzuschweifen.

Die Interventionen dauerten je 15 Minuten und wurden von einer/m geschulten erfahrenen Trainer/in durchgeführt.

In den Kontrollgruppen wurde jeweils ein Video gezeigt (Studie 1: Fitness-Workout, Studie 2: Dokumentation über Yoga).

Befindlichkeit:

Zur Verlaufsmessung der psychischen Befindlichkeit wurde der mehrdimensionale Befindlichkeitsfragebogen (MDBF) ausgewählt (Steyer, Schwenkmezger, Notz, & Eid, 1997). Dabei wird die aktuelle psychische Befindlichkeit bezüglich gute-schlechte Stimmung (GS), Wachheit-Müdigkeit (WM) und Ruhe-Unruhe (RU) erfasst.

Aufmerksamkeit:

Zur Erfassung selektiver Aufmerksamkeit wurde das Frankfurter Aufmerksamkeits-Inventar-2 (FAIR-2) ausgewählt (Moosbrugger, Oehlschlägel, & Sörensen, 2011), bei dem die Diskrimination visuell ähnlicher Zeichen erfasst wird. Als Skalen werden fehlerkorrigierter Leistungswert (bearbeitete Zeichen in Abhängigkeit der Fehler), Qualitätswert (Fehleranteil in Relation der bearbeiteten Zeichen) und Kontinuitätswert (Aufmerksamkeitsschwankungen) angegeben.

Exekutive Funktionen:

Zur Erfassung eines Teilaspekts exekutiver Funktionen, wurde der Trail Making Test (TMT) ausgewählt (Tombaugh, 2004), bei dem Informationen trotz proaktiver Interferenz und negativem Priming verarbeitet werden sollen. Als Skalen werden die Gesamtzahl der bearbeiteten Zeichen (TMTGZ), die Fehleranzahl (TMTF) und ein kombinierter Wert aus bearbeiteten Zeichen und Fehlern (TMTGZF) ausgegeben. Zur Vermeidung von Übungseffekten wurden Parallelversionen je Messzeitpunkt mit jeweils exakt gleicher Gesamtlängelänge verwendet.

Ergebnis:

Zur Befindlichkeit ergaben sich in Studie 1 und 2 signifikante mittlere bis starke Interaktionseffekte. Die Teilnehmer/innen der Interventionsgruppe berichteten jeweils eine positivere Grundstimmung, höhere Wachheit und mehr innerer Ruhe als Teilnehmer/innen der jeweiligen Kontrollbedingung ( $p < .05$ ,  $d$  zwischen .6 und 1.2).

In Studie 1 ergaben sich zusätzliche Interaktionseffekte beim FAIR-2 (selektive Aufmerksamkeit). Die Teilnehmer/innen der Bewegungspause erreichten bei den Skalen fehlerkorrigierter Leistungswert ( $F(1, 67) = 5.062$ ,  $p = .028$ ,  $d = .549$ ) und Kontinuitätswert ( $F(1, 67) = 5.182$ ,  $p = .026$ ,  $d = .557$ ) bessere Werte als Teilnehmer/innen der Kontrollgruppe. Beim TMT (exekutive Funktionen) ergaben sich keine signifikanten Interaktionseffekte.

In Studie 2 ergaben sich Interaktionseffekte beim TMT. Die Teilnehmer/innen der Interventionsgruppe bearbeiteten mehr Zeichen als Teilnehmer/innen der Kontrollbedingung TMTGZ  $F(1, 27) = 5.346$ ,  $p = .029$ ,  $d = .889$ . In Studie 2 gab es keine Interaktionseffekte beim FAIR-2.

Diskussion:

Die vorliegenden Ergebnisse belegen akute positive Effekte einer Bewegungspause und eines Achtsamkeitstrainings auf die Befindlichkeit und zusätzlich interventionsspezifische Effekte. Die positiven Effekte einer angeleiteten Pause auf relevante Lernbedingungen sollten als notwendige Lernphase insbesondere in der universitären Lehre mehr Aufmerksamkeit erhalten. Im Vortrag werden weitere Implikationen der Ergebnisse vorgestellt und Möglichkeiten einer Kombination beider Interventionen diskutiert.

**Benjamin Niederkofler<sup>1</sup>, Christian Herrmann<sup>2</sup>, Günter Amesberger<sup>3</sup>**

<sup>1</sup>Pädagogische Hochschule Salzburg, Österreich; <sup>2</sup>Universität Basel, Schweiz; <sup>3</sup>Universität Salzburg, Österreich;  
[benjamin.niederkofler@phsalzburg.at](mailto:benjamin.niederkofler@phsalzburg.at), [christian.herrmann@unibas.ch](mailto:christian.herrmann@unibas.ch)

Theoretischer Hintergrund

Professionelle Handlungskompetenz von Lehrkräften wird heute als Kontinuum zwischen Wissen und Handeln angesehen (Blömeke, Gustafsson & Shavelson, 2015; Oser, 2013). Die Diagnosekompetenz (DK) nimmt darin eine zentrale Bedeutung ein. So bildet das diagnostische Urteil im Unterricht eine wesentliche Voraussetzung für adaptives Handeln der Lehrkraft (Brühwiler, 2014; Südkamp & Praetorius, 2017). Im schulischen Diagnoseprozess wird das diagnostische Urteil von Wissen aus Diagnosen abgeleitet und durch Diagnosen gebildet. Im Sportunterricht werden Diagnosen in Diagnosesituationen gebildet, die sich häufig aus der Beobachtung von Bewegungshandlungen ergeben. Letztlich umfasst die DK auch das Wissen und Können zur Diagnostik sowie die diagnostischen Fähigkeiten, welche die Lehrkraft in die Diagnosesituation einbringt. DK ist damit einerseits wissensgenerierend, indem die Lehrkraft Wissen über die Schulkinder generiert. Andererseits ist sie auch bedingend, da sie zentrale Stütze für weitere Unterrichtshandlungen ist.

DK von Sportlehrkräften ist ein bisher kaum beachtetes Forschungsfeld. Dies ist erstaunlich, da sich hochwertige Diagnosen als wichtige Grundlage für eine kognitiv aktivierende Gestaltung des Sportunterrichts zeigen und sie damit potentiell zur Lernwirksamkeit beitragen (Niederkofler & Amesberger, 2016). Im Sportunterricht kann über die Diagnose von Bewegungshandlungen auf das Vorliegen von Kompetenzen geschlossen werden (Herrmann & Seelig, 2017a). So erfordert eine Bewegungshandlung (z.B. Zielwerfen) die Aktivierung eines repräsentierten Handlungskonzepts (Niederkofler & Amesberger, 2016). Ein Schulkind muss beim Zielwerfen motorische Repräsentationen aktivieren, um motorischen Fertigkeiten und Fähigkeiten erfolgreich zu kombinieren. Auch bestimmen wesentliche (z.B. Selbsteinschätzung des motorischen Könnens) und emotional-energetisierende Repräsentationen (z.B. sportliches Fähigkeitsselbstkonzept) den Erfolg beim Zielwerfen. Sportlehrkräfte diagnostizieren beim Zielwerfen also nicht bloß die Zielerreichung. Die Diagnose sollte das gesamte kognitive Handlungskonzept beachten. Erst dadurch können potentiell hilfreiche didaktische Entscheidungen getroffen werden.

Fragestellung

1. Wie sind die Niveau-, Differenzierungs- und Rangkomponente der Diagnosekompetenz bei Sportlehrkräften in der Einschätzung von motorischen Basiskompetenzen ausgeprägt?
2. Kann die motorische (motorische Basiskompetenzen), wissentliche (Selbstwahrnehmung der Kinder über motorische Basiskompetenzen) und emotional-energetisierende (fachliches Fähigkeitsselbstkonzept) Handlungsrepräsentation das diagnostische Urteil (Rangordnungskomponenten) vorhersagen?

Methode

Mittels Fragebogen wurden 7 Sportlehrkräfte (M = 34.4 Jahre, SD = 7.0) gebeten, die motorischen Basiskompetenzen von 100 Kinder (M = 11.7 Jahre, SD = 0.7) aus acht Klassen der 5. und 6. Schulstufe einzuschätzen („Er/Sie kann ein Ziel mit einem Ball treffen.“). Zudem wurden Kinder mittels Fragebogen zur Selbstwahrnehmung ihrer motorischen Kompetenz (SEMOK „Ich kann einen Ball gezielt werfen.“, Herrmann & Seelig, 2017b) und zum sportlichen Fähigkeitsselbstkonzept (FSK, „Ich bin sehr gut im Sport.“, Gerlach, 2008) befragt. Alle Items wurden 5-stufig skaliert. Zudem durchliefen die Kinder den MOBAK-5-Test (Herrmann & Seelig, 2017). MOBAK-5 ist ein valider Schultest zur Erfassung der motorischen Kompetenzbereiche (KB) Sich-Bewegen und Etwas-Bewegen in der 5. und 6. Schulstufe. Die Berechnung der Diagnosekomponenten erfolgte nach Schrader (1989). Die Vorhersage der Rangkomponente durch das repräsentierte Handlungskonzept der Kinder wurde mittels Regression berechnet.

Ergebnisse

Die Niveauelemente zeigt eine Überschätzung der beiden KB (Sich-Bewegen = .09; Etwas-Bewegen = .29). In Bezug auf die Differenzierungskomponente lässt sich feststellen, dass Lehrkräfte die Streuung der Merkmalsausprägung überschätzen. Die Rangkomponente verweist auf eine mittlere Urteilstreue (KB Etwas-Bewegen = .50\*\*\*; KB Sich-Bewegen = .55\*\*\*). Die Vorhersage der Rangkomponente durch das Handlungskonzept der Kinder legt für beide KB Einflussfaktoren offen (KB Sich-Bewegen:  $\beta(\text{MOBAK}) = .41^{***}$ ,  $\beta(\text{SEMOK}) = .08$  n.s.,  $\beta(\text{FSK}) = .24^*$ ,  $R^2 = .31$ ; KB Etwas-Bewegen:  $\beta(\text{MOBAK}) = .42^{***}$ ,  $\beta(\text{SEMOK}) = .22^*$ ,  $\beta(\text{FSK}) = .29^*$ ,  $R^2 = .48$ ).

Diskussion

Diese Ergebnisse decken sich mit Ergebnissen aus anderen Fächern (Südkamp, Kaiser, & Möller, 2012). Sofern mit einer leichten Überschätzung der motorischen Basiskompetenzen auch Aufgabenstellungen verbunden sind, die leicht über dem Niveau der Schulkinder liegen, kann dies durchaus als positiv betrachtet werden (Brühwiler, 2014; Hosenfeld, Helmke, & Schrader, 2002). Die Ausprägung der DK im KB Etwas-Bewegen zeigt, dass selbst Basiskompetenzen, die Kinder in jeder Unterrichtslektion benötigen, für Sportlehrkräfte schwer bewertbar sind.

**Analyse von Wirkungseinflüssen auf Lernleistungen im Sportunterricht**

**Sara Seiler**

Bundesamt für Sport, Schweiz; [sara.seiler@baspo.admin.ch](mailto:sara.seiler@baspo.admin.ch)

Theoretischer Rahmen und Fragestellung:

Die Qualität des Unterrichts ist ein entscheidender Faktor, wenn es um die Frage nach dem Einfluss auf die Entwicklung von Lernleistungen von Schülern und Schülerinnen geht. Dabei ist die durch die Schüler und Schülerinnen wahrgenommene Qualität des Unterrichts entscheidend (Hattie, 2013). Das durch die Lehrperson initiierte Unterrichtsangebot und die Nutzung dessen durch die Schüler lassen sich als Lehr-Lern-Prozess beschreiben. Merkmale der Tiefenstrukturen des Unterrichts beschreiben die Qualität dieses Prozesses. Sie lassen sich zu drei übergeordneten Dimensionen zuordnen (Kunter & Trautwein, 2013).

Der vorliegende Beitrag möchte der Fragestellung nachgehen: inwieweit beeinflussen die durch die Schüler und Schülerinnen eingeschätzte Qualität des Unterrichts ihre Lernleistungen? Dafür liegt das Angebot-Nutzen-Modell, welches die Wirkungskette schulischer Lernleistung, ausgehend vom Unterrichtsangebot der Lehrperson über die Nutzung dessen von den Schülern und Schülerinnen hin zu deren Lernleistung, veranschaulicht, als Orientierung zugrunde (Helmke, 2010).

Methode:

In der durch die Eidgenössische Sportkommission finanzierten IMPEQT-Studie nahmen unter Leitung von Prof. Dr. Erin Gerlach an der Universität Basel zu zwei Messzeitpunkten (2012 - 2013) 1081 Schüler und Schülerinnen der 7. (49.1% ♀, Alter M = 13.2 Jahre) und 8. Jahrgangstufe in 48 Klassen bei 42 verschiedenen Lehrpersonen in 24 Schulen unterschiedlicher Siedlungsgebiete teil. Bei den Erhebungen wurden die Schüler und Schülerinnen in ihren motorischen Lernleistungen (IMPEQT-MOBAQ Testaufgaben; vgl. Seiler, et al. 2016) getestet, sowie motivationale (fachspezifisches Interesse), volitionale (Anstrengungsbereitschaft) und personale Aspekte (fachspezifisches Selbstkonzept) als weitere Lernleistungen des Sportunterrichts durch Fragebogenitems ermittelt. Ebenso wurde in diesem Fragebogen die Einschätzung der Qualität des Sportunterrichts abgefragt (Herrmann, Seiler, Pühse & Gerlach, 2015). Um die geschachtelte Datenstruktur der quantitativ erhobenen Schülerdaten zu berücksichtigen, wurden Mehrebenenanalysen berechnet. Dazu wurden das Programm „Hierarchical Linear Modeling“ (HLM) 7.01 (Bryk, Raudenbush & Congdon, 2013) verwendet, welches es ermöglicht, klassenspezifische Merkmale (wie die wahrgenommene Unterrichtsqualität) auf Aspekte individueller Lernleistungen hin in ihrer Einflussnahme zu analysieren.

Ergebnisse und Diskussion:

Die Qualität des Unterrichts hatte einen Einfluss auf die Ausprägung und die Entwicklung des fachlichen Interesses und auf die Anstrengungsbereitschaft der Schüler und Schülerinnen. Insbesondere das Interesse scheint sich durch die in einer Klasse wahrgenommene Motivation der Lehrperson ( $\beta = .23$ ,  $p < .01$ ) und Zufriedenheit mit der Unterrichtsgestaltung ( $\beta = .24$ ,  $p < .001$ ) positiv zu entwickeln. Im Gegensatz zu anderen Fächern hatten die Merkmale der Klassenführung einen stärkeren Einfluss auf motivational-volitionale Aspekte als die Schülerorientierung. Sportunterrichtliche Spezifika wie der Unterricht in der Sporthalle, die Sichtbarkeit der Aktivitäten und das permanente Monitoring der Lehrperson könnten hierfür ausschlaggebend sein. Diese Ergebnisse sollten vor allem in der Aus- und Weiterbildung von Lehrpersonen berücksichtigt werden und sind in der Diskussion um guten Sportunterricht relevant.

### **Zum Einfluss von Interesse am Sport auf die Wahrnehmung der Unterrichtsqualität im Sportunterricht aus Schülersicht**

**Jan Sohnsmeier<sup>1</sup>, Christian Herrmann<sup>2</sup>, Rüdiger Heim<sup>1</sup>**

<sup>1</sup>Universität Heidelberg, Deutschland; <sup>2</sup>Universität Basel, Schweiz; [sohnsmeier@uni-hd.de](mailto:sohnsmeier@uni-hd.de)

Einleitung

Die Erfassung der Unterrichtsqualität (Unterrichts- und Klassenführung, Schülerorientierung und kognitive Aktivierung) kann durch Lehrer, externe Beobachter oder Schüler erfolgen. Der Perspektivenvergleich zeigt, dass die „Erfassung und Beurteilung nur perspektivenspezifisch valide möglich ist, wobei keine der drei Perspektiven grundsätzlich näher an der ‚wahren‘ Unterrichtsqualität ist als die anderen“ (Rahn, Gruehn, Keune & Fuhrmann, 2016, S. 167). Fokussiert man auf Schülerurteile zur Diagnostik von Unterrichtsqualität, werden verschiedene Biasvariablen diskutiert. Von besonderer Bedeutung ist die Ausprägung des Fachinteresses. Die Befunde zeigen einen positiven Zusammenhang zwischen dem Fachinteresse der Schüler und ihrer Bewertung der Unterrichtsqualität (zsf. Rahn et al., 2016).

Auch für die Evaluation von Sportunterricht werden Schülerurteile als sinnvoll und praktikabel angesehen und vielfältig verwendet (Herrmann, Seiler, Pühse & Gerlach, 2015). Untersuchungen, ob Schülerratings systematischen Verzerrungen unterliegen, wurden in der Sportdidaktik bisher allerdings noch nicht durchgeführt. Diese dürften aufgrund der lebensweltlichen Verstrickung der Heranwachsenden in den außerschulischen Sport jedoch eine besondere Relevanz besitzen. Vor diesem Hintergrund wird im vorliegenden Beitrag analysiert, welchen Einfluss das Interesse am Sport auf die wahrgenommene Qualität des Sportunterrichts hat.

Methode

Die Analyse basierte auf einer Stichprobe von 618 Gymnasialschülern (52% weiblich) der Klassen fünf bis zwölf. Der Altersrange reichte von 8 bis zu 18 Jahren und betrug im Mittel 12.91 Jahre (SD = 2.13). Um das Interesse am Sport zu messen, kam der Fragebogen iSpo zum Einsatz, mit dem das individuelle Interesse am Sport erfasst wird (Heim & Sohnsmeier, 2016; Sohnsmeier & Heim, 2017). Der Fragebogen besitzt eine bifaktorielle Struktur (Morin, Arens & Marsh, 2016) mit dem G-Faktor Interesse am Sport und den gut differenzierten Subskalen Rezeption, Aktivität und Vermittlung mit jeweils sechs Items. Die Fragen zur Unterrichtsqualität wurden auf Basis des von Herrmann et al. (2015) empirisch geprüften Erhebungsinstruments weiterentwickelt. Das Instrument beschreibt neun Subskalen mit jeweils drei Items, die die drei Dimensionen der Unterrichtsqualität bestimmen.

Die Datenauswertung erfolgte anhand latenter Regressionsanalysen. Dazu wurden die drei latent modellierten Qualitätsdimensionen Klassenführung, Schülerorientierung und kognitive Aktivierung durch die Prädiktoren Interesse am Sport (bifaktorielle Modellierung), Geschlecht, Alter und Vereinsmitgliedschaft vorhergesagt. Fehlende Werte wurden modellbasiert (FIML) geschätzt und die hierarchische Datenstruktur durch eine Korrektur der Standardfehler berücksichtigt.

Ergebnisse und Diskussion

Die Strukturgleichungsanalysen ergaben durchgängig akzeptable Modellfits: CFIs  $\geq .915$ ; TLIs  $\geq .903$ ; RMSEAs  $\leq .053$ ; SRMRs  $\leq .076$ . Es kann festgehalten werden, dass der Einfluss auf die wahrgenommene Unterrichtsqualität am größten ist bei sportinteressierten, jüngeren und männlichen Schülern, die nicht in einem Verein Sport treiben. Dieses Muster zeigte sich für die beiden Dimensionen Schülerorientierung und kognitive Aktivierung. Dabei war der Einflussfaktor Interesse am Sport über die anderen unabhängigen Variablen hinweg am bedeutsamsten. Für die Dimension Klassenführung fanden sich keine Verzerrungstendenzen durch die untersuchten Variablen. Die Ergebnisse machen also deutlich, dass ein höheres Interesse am Sport mit einer positiveren Wahrnehmung der Unterrichtsqualität in zwei der drei Dimensionen einhergeht. Es muss jedoch diskutiert werden, ob sportlich interessierte Schüler den Unterricht positiver wahrnehmen, oder ob sich das Lehrerverhalten verändert, wenn sich die Schüler für das Fach interessieren (Bohmann & Weinstein, 2013). Auffällig ist auch die negative Verzerrung durch die Vereinsmitgliedschaft. Demnach haben Vereinsmitglieder einen kritischeren Blick auf ihren Sportunterricht, obwohl ihr Interesse am Sport generell höher ausfällt.

## C10

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus S107

### Lösungswege, Lösungsschritte und Positioning der Lernenden bei der Erarbeitung einer mathematischen Textaufgabe in Kleingruppen

Iris Tanner<sup>1</sup>, Kurt Reusser<sup>1</sup>, Christine Pauli<sup>2</sup>

<sup>1</sup>Universität Zürich, Schweiz; <sup>2</sup>Universität Fribourg, Schweiz; [itanner@ife.uzh.ch](mailto:itanner@ife.uzh.ch)

#### Theoretischer Rahmen

Auf der Grundlage soziokultureller Lehr- und Lernkonzepte wird die Qualität der Unterrichtsinteraktion für den Wissensaufbau als zentral betrachtet. Deshalb propagieren Modelle produktiver Unterrichtskommunikation die Vorstellung ko-konstruktiver Gespräche, bei denen die Lernenden substantiell mit eigenen Beiträgen verantwortungsvoll an der gemeinsamen Wissensgenerierung partizipieren. Denn individuelles Wissen kann aus kollektivem Wissen entstehen (Mercer 2013). Die gemeinsame Verantwortlichkeit der Gesprächsteilnehmenden für die Gemeinschaft der Lernenden, für exaktes Wissen und für präzise Denkvorgänge (Sohmer, Micheals, O'Connor & Resnick, 2009) soll nachhaltiges Lernen ermöglichen. Vor diesem Hintergrund wird der Frage nachgegangen, wie in einer durch die Lehrperson angeleiteten Kleingruppensituation die Lösung einer Textaufgabe ko-konstruiert wird. Dabei werden einerseits die Partizipationsräume der Lernenden (Brandt 2004) und ihr Positioning (Greeno 2006) analysiert und andererseits wird ein Modell des Lösens mathematischer Textaufgaben zugrunde gelegt, das von vier Schritten der Mathematisierung ausgeht (Reusser 1989/1995): Aufbau von Text- und Situationsverständnis, Mathematisierung, Rechnen sowie Formulieren einer situationsbezogenen Antwort. Um diese Schritte kennen zu lernen und mit der Zeit selbstständig vollziehen zu können, müssen die Lernenden Gelegenheit erhalten, an gemeinsamen Problemlösungsprozessen verantwortungsvoll partizipieren zu können. Um die Qualität der Unterrichtsgespräche zu sichern, braucht es eine lernberatende Form der Lehrerrolle (Scaffolding), welche die Lernenden anleitet, unterstützt und dafür besorgt sind, dass ein anspruchsvoller fachbezogener Lernprozess in Gang gehalten wird.

#### Forschungsumfeld

Analysiert wurden 33 videografierte Lehr-Lerngespräche in Kleingruppen von je vier Schülerinnen und Schülern und einer Lehrperson beim Lösen einer inhaltlich standardisierten mathematischen Textaufgabe auf Sekundarstufe I (132 Lernende, ca. 500 Min. Videoaufnahmen). Die Daten stammen aus der vom SNF geförderten Videostudie „Didaktische Kommunikation und Bildungswirkungen im problemorientierten Mathematikunterricht“ (Reusser & Pauli 2012).

#### Fragestellungen

1. Welche unterschiedlichen Lösungswege und Lösungsschritte werden in den Kleingruppen angewendet?
2. Inwiefern ermöglichen die Lehrpersonen den Lernenden substantielle Beiträge beim Wissensaufbau in einer ko-konstruktiven Lehr-Lernsituation einzubringen?
3. Bei welchen Lösungsschritten sind die Lernenden und bei welchen die Lehrpersonen partizipatorisch und inhaltlich am aktivsten?

#### Forschungsmethodik

Um den Mathematisierungsprozess in den unterschiedlichen Gruppen zu untersuchen, wird anhand einer Episodenzuordnung des Videos erfasst, welche Lösungsschritte in den jeweiligen Kleingruppensituationen explizit hervorgehoben werden, wie lange die Behandlung dieser Schritte dauert und wer diese initiiert. Sodann wurden die Beiträge der einzelnen Lernenden zur gemeinsamen Wissenskonstruktion in Anlehnung an das Produktdesign von Krummheuer und Mitarbeitern (Krummheuer & Brandt, 2001) nach dem Grade ihrer Autonomie kodiert und sequenzanalytisch nach den einzelnen Lösungsschritten ausgewertet.

#### Ergebnisse

Die Ergebnisse zeigen auf, dass sich die Zeitdauer der einzelnen Lösungsschritte zwischen den Gruppen stark unterscheidet und dass die Reihenfolge der Lösungsschritte nicht immer gleich ist. Generell lässt sich sagen, dass die Erarbeitung des Situationsverständnisses und der Mathematisierung am meisten Zeit beanspruchen und dass etliche Gruppen auf das Formulieren eines Antwortsatzes verzichten. Zudem kann berichtet werden, dass die Lernenden beim Situationsverständnis mehr substantielle Beiträge einbringen als bei der Mathematisierung. Ein weiteres Ergebnis ist, dass in Gruppen, welche weniger Zeit für die Problemlösung beanspruchen, die Lehrperson viele substantielle Äusserungen übernimmt. Achtet man auf die Grade der Autonomie der Äusserungen der einzelnen Lernenden innerhalb der Gruppen, so lassen sich Schülertypen beschreiben. So stellt derjenige Lernende, welcher beim Lösungsschritt ‚Ausrechnen‘ oft das Wort ergreift, mehrheitlich Fragen. Diese Ergebnisse werden vor dem Hintergrund der Kongressfrage diskutiert, wie Lernprozesse differenziert erschlossen werden können.

### Wie differenziert fallen Schülerurteile über die Unterrichtsqualität in verschiedenen Fächern aus? Ergebnisse zum Mathematik- und Sachunterricht in der Grundschule

Svenja Rieser<sup>1</sup>, Ruven Stahns<sup>2</sup>, Alexander Naumann<sup>3,4</sup>, Heike Wendt<sup>5</sup>

<sup>1</sup>Bergische Universität Wuppertal; <sup>2</sup>Universität Duisburg-Essen; <sup>3</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF); <sup>4</sup>IdEA Forschungszentrum Frankfurt; <sup>5</sup>Technische Universität Dortmund; [svenja.rieser@tu-dortmund.de](mailto:svenja.rieser@tu-dortmund.de)

Die kognitive Aktivierung der Schüler, ein unterstützendes Klima und die Unterrichtsführung sind zentrale Qualitätsdimensionen erfolgreichen Unterrichts in verschiedenen Fächern (Klieme & Rakoczy, 2008). Zur Erfassung der Unterrichtsqualität werden in Studien häufig Schülerurteile genutzt, auch wenn deren Validität angezweifelt werden kann (z.B. Aleamoni, 1999; de Jong & Westerhof, 2001). Für die Validität von Schülerurteilen sprechen jedoch Ergebnisse von Studien, in denen – basierend auf Schülereinschätzungen – die dreidimensionale Struktur der Unterrichtsqualität sowie Zusammenhänge zwischen der Wahrnehmung des Unterrichts und der Motivation bzw. der Leistung nachgewiesen werden konnten (z.B. Fauth et al., 2014a; Kunter & Baumert., 2006).

Ein weiterer Hinweis auf die Validität von Schülerurteilen kann aus dem Vergleich von Einschätzungen der Qualität über verschiedene Fächer hinweg gewonnen werden: Wenn Schüler in der Lage sind, valide Angaben zur Unterrichtsqualität zu machen, so sollten sich in ihren Urteilen über die Qualität in verschiedenen Fächern Unterschiede nachweisen lassen. Dabei muss berücksichtigt werden, ob Klassen in den Fächern von derselben oder von verschiedenen Lehrkräften unterrichtet

werden (Praetorius et al., 2015; Wagner et al., 2013). Da die Unterrichtsqualität stark von Merkmalen der Lehrkraft und der Interaktion zwischen Lehrkraft und Klasse abhängt, kann erwartet werden, dass sich die Unterrichtsqualität zwischen zwei Fächern nur in geringem Maße unterscheidet, wenn eine Klasse in beiden Fächern von derselben Lehrkraft unterrichtet wird. Wird die Klasse in den beiden Fächern dagegen von verschiedenen Lehrkräften unterrichtet, sollten deutlichere Unterschiede sichtbar werden.

1. Lassen sich die drei Dimensionen in Schülereinschätzungen zu verschiedenen Fächern identifizieren?
2. Unterscheiden Schüler in ihren Angaben zur Unterrichtsqualität zwischen zwei Fächern, wenn sie in beiden Fächern...
  - a. von derselben Lehrkraft unterrichtet werden?
  - b. von verschiedenen Lehrkräften unterrichtet werden?

3948 Schülerinnen und Schüler aus 211 Klassen beantworteten in TIMSS 2015 Fragen zur Qualität ihres Mathematik- und Sachunterrichts. Dabei wurden für beide Fächer dieselben Items eingesetzt (4 Items zur kognitiven Aktivierung, 5 Items zum unterstützenden Klima und 5 Items zur Unterrichtsführung).

Zur Beantwortung der ersten Fragestellung wurde für jedes Fach eine Mehrebenen-CFA (ML-CFA) mit drei Faktoren auf zwei Ebenen spezifiziert. Zur Beantwortung der zweiten Fragestellung wurde der Datensatz in zwei Gruppen geteilt: Die erste Gruppe umfasste Klassen, die in beiden Fächern von derselben Lehrkraft (126 Klassen) unterrichtet wurden, die zweite Gruppe Klassen, die in beiden Fächern von unterschiedlichen Lehrkräften unterrichtet wurden (74 Klassen). Für jede dieser Gruppen wurde pro Dimension ein zweifaktorielles ML-CFA Modell spezifiziert. Der erste Faktor bildete die Dimension im Mathematikunterricht ab, der zweite Faktor im Sachunterricht.

1. In Übereinstimmung mit anderen Studien lässt sich das dreidimensionale Modell für den Mathematik- (RMSEA = .036; CFI = .931) und den Sachunterricht (RMSEA = .044; CFI = .927) mit einem akzeptablen Fit abbilden. Das Modell lässt sich somit auf beide Fächer anwenden.

2a: Wird die Klasse in beiden Fächern von derselben Lehrkraft unterrichtet, kann das Modell aufgrund einer zu hohen Korrelation zwischen den Faktoren für zwei der Dimensionen interpretiert werden. Für die dritte Dimension korrelieren die Faktoren auf beiden Ebenen mit  $r \geq .97$  ( $p < .05$ ). Die Schülerangaben differenzieren in diesem Fall also nicht zwischen den Fächern.

2b: Wird die Klasse in den Fächern von verschiedenen Lehrkräften unterrichtet, können für alle Dimensionen die entsprechenden zweifaktoriellen Modelle geschätzt werden. Die Korrelationen zwischen den Faktoren variieren zwischen  $r_{within} = .60$  ( $p < .05$ ) und  $r_{within} = .89$  ( $p < .05$ ) bzw.  $r_{between} = -.23$  (n.s.) und  $r_{between} = .79$  ( $p < .05$ ). Entsprechend unterscheiden sich die Angaben zur Unterrichtsqualität zwischen den Fächern.

Die Ergebnisse liefern einen weiteren Beleg dafür, dass es möglich ist, Unterrichtsqualität durch Schülerangaben valide zu erfassen. Weiterhin deuten die Befunde darauf hin, dass es sich bei den erfassten Merkmalen der Unterrichtsqualität aus Schülersicht um stabile Personenmerkmale von Lehrkräften handeln könnte.

### **„Ich will gut in Mathe sein“ – eine längsschnittliche Überprüfung der Theorie des geplanten Verhaltens zur Vorhersage von Mathematikleistung**

**Christoph Niepel<sup>1</sup>, Jeremy Burrus<sup>2</sup>, Samuel Greiff<sup>1</sup>, Anastasiya A. Lipnevich<sup>3</sup>, Meghan W. Brennan<sup>4</sup>, Richard D. Roberts<sup>2</sup>**

<sup>1</sup>Universität Luxemburg, Luxemburg; <sup>2</sup>ACT, Inc., USA; <sup>3</sup>Queens College & the Graduate Center, City University of New York; USA; <sup>4</sup>Educational Testing Service, USA; [christoph.niepel@uni.lu](mailto:christoph.niepel@uni.lu)

Die Theorie des geplanten Verhaltens (Ajzen, 1991) ist eine der prominentesten Theorien zur Erklärung des Zusammenhangs von Einstellung und Verhalten und hat sich als ein theoretisch bedeutsamer Rahmen erwiesen, um Mathematikleistung von Schülerinnen und Schülern zu erklären (Lipnevich, MacCann, Krumm, Burrus, & Roberts, 2011). Die Theorie des geplanten Verhaltens erklärt Verhalten (welches sich z.B. in guter Mathematikleistung ausdrückt) durch (1) die individuelle Intention, ein Verhalten zu zeigen (gut in Mathematik zu sein), die wiederum gebildet wird durch (2) die individuelle Einstellung (zum Fach Mathematik), (3) der erlebten sozialen Norm (im Fach Mathematik gut zu sein) und (4) der erlebten Verhaltenskontrolle (Kompetenzerleben im Fach Mathematik). Allerdings fehlen bislang Studien, die die prädiktive Validität der Theorie des geplanten Verhaltens zur Vorhersage von Mathematikleistung im Längsschnitt untersucht haben. Die vorliegende Studie untersuchte daher den Zusammenhang zwischen den vier Konstrukten der Theorie des geplanten Verhaltens – also Intention, Einstellung, soziale Normen und Verhaltenskontrolle (jeweils bezogen auf das Fach Mathematik) – mit Mathematikleistung über einen Zeitraum von circa einem halben Jahr. Hierfür untersuchten wir N=752 (T1 im Juni 2012; 50% männlich) bzw. N=514 (T2 im November 2012; 49.3% männlich) amerikanische Schülerinnen und Schüler, die zum Zeitpunkt der Befragung acht verschiedene Middle Schools aus sechs verschiedenen US-Bundesstaaten besuchten. Intention, Einstellung, soziale Norm und erlebte Verhaltenskontrolle im Fach Mathematik wurden mit Hilfe des Mathematics Attitude Questionnaire (Lipnevich et al., 2011) erfasst. Als Indikatoren für Mathematikleistung dienten die von den Schulen berichteten Mathematiknoten der Schülerinnen und Schüler. Für unsere Analysen berechneten wir Strukturgleichungsmodelle im Statistikprogramm Mplus 7. Um fehlende Werte in den Analysen zu berücksichtigen, nutzten wir multiple Imputation. Wir wählten ein cross-lagged latent variable design, in welchem die untersuchten vier Konstrukte der Theorie des geplanten Verhaltens sowie die Mathematikleistung jeweils als Prädiktoren (T1) und Kriterien (T2) eingingen. Weiterhin kontrollierten wir in unseren Analysen individuelle Unterschiede der Schülerinnen und Schüler im schlussfolgernden Denken (Quantitative Reasoning; Comprehensive Testing Program 4; ERB, 2012). Als eine Voraussetzung für die Analyse längsschnittlicher Zusammenhänge, konnte metrische Messinvarianz für alle latenten Konstrukte etabliert werden. Die Ergebnisse der cross-lagged Analyse zeigten, dass – auch nach Kontrolle der vorherigen Leistung (Mathematiknote zu T1) – die Intention, besser in Mathematik werden zu wollen (T1) und spätere Mathematikleistung (T2) einen bedeutsamen Zusammenhang ( $\beta = .12$ ,  $p < .05$ ) aufwiesen. Daneben haben wir die längsschnittlichen Zusammenhänge der vier Konstrukte der Theorie des geplanten Verhaltens untereinander explorativ untersucht. Hierbei zeigten Intention (T1) und Verhaltenskontrolle (T2) einen positiven Zusammenhang über die Zeit ( $\beta = .23$ ,  $p < .01$ ), Einstellung (T1) und Verhaltenskontrolle (T2) einen negativen Zusammenhang (nach Kontrolle geteilter Varianzen;  $\beta = -.13$ ,  $p < .05$ ). Intention und soziale Norm zeigten einen positiven und reziproken längsschnittlichen Zusammenhang (mit  $\beta = .20$ ,  $p < .01$  [Norm auf Intention] und  $\beta = .09$ ,  $p < .05$  [Intention auf Norm]). Insgesamt sprechen unsere Ergebnisse für die inkrementelle prädiktive Validität der Theorie des geplanten Verhaltens zur Vorhersage von Mathematikleistung über Reasoning hinaus. Die längsschnittlichen Zusammenhänge der Konstrukte der Theorie des geplanten Verhaltens untereinander werden vor dem Hintergrund akademischer Motivationstheorien diskutiert.

## Formatives Assessment im Mathematikunterricht der Grundschule

**Birgit Schütze<sup>1</sup>, Jannis Kraiss<sup>2</sup>, Elmar Souvignier<sup>1</sup>, Ralf Benölken<sup>1</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, Deutschland; <sup>2</sup>University of Twente; [birgit.schuetze@uni-muenster.de](mailto:birgit.schuetze@uni-muenster.de)

### Theoretischer Hintergrund

Formatives Assessment gilt als eines der wirksamsten Rahmenkonzepte zur Förderung schulischen Lernens (Black & Wiliam, 1998; Hattie, 2009). Zentrale Merkmale von formativem Assessment sind die Diagnose des individuellen Lernstands, eine darauf basierende Leistungsrückmeldung und eine differenzierte Förderung (Wiliam & Thompson, 2008). Zur Gestaltung und Implementation dieser Merkmale mangelt es jedoch an empirischen Befunden sowie an praktikablen und evaluierten Instrumenten (Bennett, 2011). Einen Ansatz zur Unterstützung des Diagnose- (nicht aber Förder-) -prozesses im Mathematikunterricht liefern Instrumente zur Lernverlaufdiagnostik. Empirische Befunde deuten darauf hin, dass die Nutzung solcher Diagnostikinstrumente insbesondere dann leistungsförderlich ist, wenn den Lehrkräften ergänzend förderbezogene Hilfen zur Verfügung gestellt werden (Stecker, Fuchs & Fuchs, 2005). In der vorliegenden Studie wird die Wirkung einer onlinebasierten Lernverlaufdiagnostik (quop; Souvignier, Förster, Schulte, 2014) allein und in Kombination mit einer Rückmelde- und Förderkomponente (quop+Förder) im Vergleich zu herkömmlichen Mathematikunterricht (KG) untersucht. Während sich ein Großteil einschlägiger Studien bislang auf Leistungseffekte konzentriert, beschäftigt sich diese Studie darüber hinaus mit der vermittelnden Funktion von motivationalen (wahrgenommene Kompetenzunterstützung) und selbstregulatorischen (Selbstevaluationsfähigkeit) Variablen.

### Fragestellungen

(1) Wie wirkt sich quop bzw. quop+Förder im Vergleich zur KG auf die Entwicklung der erlebten Kompetenzunterstützung und (2) auf die Entwicklung der Selbstevaluationsfähigkeit von Schüler(inne)n aus? (3) Haben die Entwicklung der erlebten Kompetenzunterstützung bzw. Selbstevaluationsfähigkeit einen Effekt auf die Leistungsentwicklung? (4) Gibt es einen indirekten Leistungseffekt vermittelt über die Entwicklung der erlebten Kompetenzunterstützung bzw. die Entwicklung der Selbstevaluationsfähigkeit?

### Methode

Im Rahmen einer quasi-experimentellen Feldstudie wurden 27 Klassen der dritten Jahrgangsstufe (N = 614 Schüler(innen)) randomisiert einer der drei Bedingungen (quop+Förder, quop, KG) zugeordnet. In der quop-Bedingung bearbeiteten die Schüler(innen) ein Schuljahr lang alle drei Wochen Mathematiktests am Computer. Nach jeder Testung wurden den Lehrkräften und Schüler(inne)n die Ergebnisse in Form von Lernverläufen computerbasiert rückgemeldet. In der quop+Förder-Bedingung erhielten die Schüler(innen) darüber hinaus eine differenzierte mündliche Rückmeldung durch die Lehrkraft sowie individuell auf ihr jeweiliges Testergebnis abgestimmte Fördermaterialien. Die Mathematikleistung der Schüler(innen) wurde mit dem DEMAT2/3 (Prä-/Posttest; Cronbachs  $\alpha=.89/.82$ ) erfasst. Die erlebte Kompetenzunterstützung (Cronbachs  $\alpha=.77$ ) bzw. die Selbstevaluationsfähigkeit (Cronbachs  $\alpha=.83$ ) wurden mit Schülerfragebögen zu einem Prä- und Postmesszeitpunkt erhoben. Die statistische Auswertung erfolgte unter Berücksichtigung der Mehrebenenstruktur mittels Pfadanalysen in MPlus.

### Ergebnisse

(1) Schüler(innen) der quop- bzw. quop+Förder-Bedingung zeigen eine stärkere positive Entwicklung ihrer erlebten Kompetenzunterstützung als Schüler(innen) der KG,  $\beta=.46$ ,  $p<.01$  bzw.  $\beta=.40$ ,  $p<.01$ . (2) Zudem ergibt sich im Vergleich zur KG in der quop+Förder-Bedingung ein größerer Zuwachs in der Selbstevaluationsfähigkeit,  $\beta=.26$ ,  $p=.01$ ; dies gilt aber nicht für die quop-Bedingung,  $\beta=.03$ ,  $p=.80$ . (3) Während die Entwicklung der erlebten Kompetenzunterstützung positiv mit der Leistungsentwicklung assoziiert ist,  $\beta=.11$ ,  $p=.04$ , gibt es keinen Zusammenhang zwischen der Entwicklung der Selbstevaluationsfähigkeit und der Entwicklung der Leistung,  $\beta=.06$ ,  $p=.16$ . (4) Im Vergleich zur KG zeigt sich kein indirekter Effekt von quop auf die Leistungsentwicklung vermittelt über die Entwicklung des Kompetenzerlebens,  $\beta=.05$ ,  $p=.10$ , bzw. die Entwicklung der Selbstevaluationsfähigkeit,  $\beta=.002$ ,  $p=.82$ . Ebenso hat die quop+Förder-Bedingung verglichen mit der KG keinen indirekten Effekt auf die Leistungsentwicklung über die Entwicklung des Kompetenzerlebens,  $\beta=.04$ ,  $p=.12$ , bzw. die Entwicklung der Selbstevaluationsfähigkeit,  $\beta=.02$ ,  $p=.28$ .

### Diskussion

Zusammengefasst deuten die Befunde darauf hin, dass sich Schüler(innen) in Klassen mit Lernverlaufdiagnostik (mit und ohne ergänzende Rückmelde- und Förderkomponente) zunehmend in ihrem Kompetenzerleben unterstützt fühlen. Die Kombination von quop mit Rückmelde- und Fördermaterialien scheint zusätzlich die Entwicklung der Selbstevaluationsfähigkeit zu begünstigen. Die Leistungen der Schüler(innen) konnten durch keine der beiden Interventionen signifikant beeinflusst werden. Die Befunde sollen unter Berücksichtigung implementationspraktischer Fragen diskutiert werden.



## C11

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus, Aula 33

### Die Thin Slices Technik zur Videoanalyse in der Unterrichtsforschung

Chair(s): **Susanne Kuger** (Deutsches Institut für Internationale Pädagogische Forschung, Deutschland)

DiskutantIn(nen): **Felicita Thiel** (Freie Universität Berlin)

In der Unterrichtsforschung gelten Beurteilungen von Videos durch neutrale Dritte als objektive, reliable, valide und informative jedoch aufwändige und teure Methode der Erfassung von unterrichtlichem Handeln und relevanten Prozessen des pädagogischen Geschehens (Pauli & Reusser, 2006; Helmke, 2014). Die video-basierte Unterrichtsforschung deckt dabei ein breites Feld relevanter Themenschwerpunkte ab und entwickelte sich in den letzten Jahren vielfältig weiter (Seidel & Thiel, 2017).

Für die Beschreibung unterrichtlichen Handelns via Videoaufzeichnungen werden zumeist ganze Unterrichtsstunden aufgezeichnet, (mindestens einmal) vollständig von ausführlich geschulten Beobachtern angesehen und dabei eingeschätzt (Jordan & Seidel, 2009). So können die im Unterricht stattfindenden Prozesse, oberflächliche Strukturen und auch komplexe Konstrukte differenziert und adäquat beurteilt werden. Gleichzeitig ist bekannt, dass Beobachterurteile, die nur aufgrund sehr geringer Informationsmenge, d. h. aufgrund des ersten Eindrucks gefällt werden, hinsichtlich unterschiedlicher Eigenschaften relativ zuverlässig ausfallen können (Ambady, Bernieri & Richeson, 2000). Dieser erste Eindruck wird durch das Ansehen relativ kurzer Videoausschnitte (Thin Slices) von 5-60 Sekunden Länge erreicht und erwies sich in früheren Studien zur Unterrichtsforschung als relevante und zuverlässige Informationsquelle (Ambady & Rosenthal, 1993; Babad, 2005; Begrich, Fauth, Kunter & Klieme, 2017; Pretsch, Flunger, Heckman & Schmitt, 2013).

Das Projekt Thin Slices als Verbundvorhaben der Goethe Universität Frankfurt und des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) untersuchte inwiefern der erste Eindruck einer Unterrichtssituation zu zuverlässigen und relevanten Urteilen kommt, inwiefern diese Urteile mit ausführlichen Videobeurteilungen und Schüleraussagen übereinstimmen und welche Rolle die mittels Thin Slices erhobenen Informationen über den Unterricht in der Prädiktion der Motivations- und Leistungsentwicklung spielen. Neben dem Umgang mit Fehlern werden auch die drei Basisdimensionen guten Unterrichts betrachtet. Der Umgang mit Fehlern ist ein alltägliches Unterrichtsereignis, welches bislang in seiner Bedeutung als Unterrichtsmerkmal wenig erforscht wurde. Aus theoretischer Sicht wird ein positiver und konstruktiver Umgang mit Fehlern auf individueller Schülerebene mit einer günstigen Leistungs- und Motivationsentwicklung in Verbindung gebracht. Das Projekt re-analysiert für diesen Zweck Daten der Studien DESI und der TIMSS.

Das Symposium fasst die Ergebnisse des mittlerweile abgeschlossenen Projekts in drei Beiträgen zusammen. Der erste Beitrag widmet sich auf konzeptioneller Ebene der Dimensionalität der mittels Thin Slices Einschätzungen über den Unterricht eingeholten Informationen. Er untersucht inwiefern die von den Beurteilern abgegebenen Einschätzungen der Thin Slices ein Globalurteil über Unterricht oder unterschiedliche Facetten abbilden und bindet diese die Forschung zur Unterrichtsqualität und zum Umgang mit Fehlern im Unterricht an. Der zweite Beitrag konzentriert sich auf die Reliabilität von Unterrichtseinschätzungen mittels Thin Slices Technik und vergleicht die so entstandenen Daten mit den Ergebnissen systematischer ausführlicher Videobeobachtungen sowie mit Schüler- und Lehrerangaben aus Befragungen. Im dritten Beitrag stehen Fragen der prädiktiven Validität der Thin Slices Einschätzungen im Mittelpunkt. Hierfür werden mit DESI Daten längsschnittliche Mehrebenenstrukturgleichungsmodelle modelliert, um die Vorhersagekraft der Thin Slices Einschätzungen für die Leistungs- und Motivationsentwicklung der Schülerinnen und Schüler zu analysieren. Insbesondere wird die prädiktive Kraft mit der von Schülerangaben und ausführlichen Videobeurteilungen verglichen. Zum Abschluss werden die Beiträge diskutiert und die Bedeutung der Thin Slices Technik für die video-basierte Unterrichtsforschung kritisch hinterfragt.

#### Beiträge des Symposiums

### Dimensionalität von Unterrichtsmerkmalen in der Thin Slices Technik – Beziehungen zwischen Fehlerumgang und Basisdimensionen guten Unterrichts

**Susanne Kuger<sup>1</sup>, Lukas Begrich<sup>2</sup>, Julia Käfer<sup>1</sup>, Eckhard Klieme<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland, <sup>2</sup>Johann Wolfgang Goethe-Universität

Theoretischer Hintergrund:

Zur Beurteilung der Güte des pädagogischen Geschehens im Klassenraum haben sich in der Unterrichtsforschung die drei Basisdimensionen guten Unterrichtes (Klieme, Schümer & Knoll, 2001) etabliert (Fauth, Decristan, Rieser, Klieme & Büttner, 2014). Zusätzlich werden häufig Dimensionen erfasst, die studienspezifische Aspekte des Unterrichts abdecken (Decristan et al., 2015; Harks, Rakoczy, Hattie, Besser & Klieme, 2014).

Der Umgang mit Fehlern im Unterricht ist nach konstruktivistischen Lerntheorien ein wichtiger Bestandteil des Lernprozesses auf Individualebene, jedoch nicht explizit ein Bestandteil der Basisdimensionen. Schüler können aus Fehlern Erkenntnisse darüber erlangen, welche Fehlvorstellungen in ihrem Wissen existieren und ein positiver Umgang mit ihren Fehlern kann zu einer Revision eventueller Fehlvorstellung und damit zu Lernen führen (Oser & Spychiger, 2005). In verschiedenen Studien bewies ein positiver Umgang mit Fehlern bzw. eine positive „Fehlerkultur“ auf Individualebene positive Zusammenhänge mit dem Leistungsfortschritt und/oder der motivationalen Entwicklung der Schülerinnen und Schüler (Kreutzmann, 2014; Steuer & Dresel, 2015). Dabei wurde das Fehlerklima jeweils mehrdimensional erhoben und umfasst auf der einen Seite kognitive und affektive Aspekte des Umgangs mit einer Situation, in der ein Schüler im Unterricht einen Fehler macht, sowie auf der anderen Seite Reaktionen des Schülers selbst, der Mitschüler sowie der Lehrkräfte (z. B. Steuer, 2014; Oser & Spychiger, 2005). Die Anzahl der angenommenen Dimensionen variiert zwischen den Studien. Derzeit ist unklar, ob diese Facetten auch als unterrichtliche Prozesse auf Klassenebene identifizierbar sind und damit Unterrichtsmerkmale darstellen, die gezielt im Unterricht manipuliert und verbessert werden können.

Ein konstruktiver Umgang mit Fehlern sollte mit zwei der Basisdimensionen stärker und mit einer weniger stark assoziiert sein. Zum einen kann der Umgang mit den eigenen Fehlvorstellungen kognitiv aktivierend wirken (Klieme, 2006), zum anderen stellt ein angemessener Umgang mit Schülerfehlern eine Facette des unterstützenden und motivierenden Unterrichts dar (Seidel & Prenzel, 2004). Die Dimension Struktur und Regelklarheit sollte weniger stark mit dem Umgang mit Fehlern im Unterricht interagieren. Entsprechende Beziehungen wurden bislang kaum untersucht bzw. mehrheitlich auf individueller Ebene bzw. Anhand von Schülerangaben überprüft.

Ziel der Studie:

Der Beitrag untersucht drei verbundene Fragestellungen anhand der im Thin Slices Projekt erhobenen Daten. Konfirmatorisch werden die drei Basisdimensionen guten Unterrichts in den Einschätzungen, die auf den Thin Slices Videoausschnitten beruhen, nachgewiesen. Diese Analyse stellt eine erste Validierung der Thin Slices Methode dar. In einem zweiten Schritt werden explorativ vier Dimensionen des Umgangs mit Fehlern in den Thin Slices Beobachtungsdaten identifiziert. Schließlich werden die Dimensionen des Umgangs mit Fehlern und die drei Basisdimensionen in einem Zweiebenenstrukturgleichungsmodell zueinander in Beziehung gesetzt.

Methode und Ergebnisse:

Die Daten stammen aus einer Zusatzerhebung, welche die Unterrichtsvideos aus TIMSS und der DESI-Studie neu analysiert. Aus den vorhandenen Videos wurden je drei Thin Slices Ausschnitt von je 10 Sekunden Länge geschnitten. Diese Ausschnitte wurden jeweils 30 ungeschulten Beobachtern präsentiert, die ihren ersten Eindruck des Unterrichts hinsichtlich einer Reihe von Items einschätzen sollten. Die Items bezogen sich auf die drei Basisdimensionen sowie auf den Umgang mit Fehlern im Unterricht. Die Mehrebenen-CFA zum Nachweis der drei Basisdimensionen erreichte gute Fitwerte. Eine explorative Mehrebenenfaktorenanalyse weist auf vier Subdimensionen des Fehlerumgangs hin, wobei zwei Subdimensionen das Lehrkraftverhalten und zwei das (Mit-)Schülerverhalten ansprechen. Ebenso fassen jeweils zwei Subdimensionen Aspekte des kognitiven Umgangs mit Fehlern zusammen und zwei Subdimensionen eher Beziehungsaspekte. Im Mehrebenenstrukturgleichungsmodell zeigen sich theoriekonforme Korrelationen zwischen den Faktoren der Basisdimensionen und denen des Umgangs mit Fehlern. Die bestätigten und gefundenen Dimensionen bilden damit sowohl den Ausgangspunkt für die beiden weiteren Beiträge des Symposiums zur Validierung der Thin Slices Technik als auch eine Erweiterung der in der Unterrichtsforschung zu berücksichtigenden Merkmale des pädagogischen Geschehens. Besonders diese Erweiterung wird in der Diskussion des Beitrags berücksichtigt.

### **Thin Slices als zuverlässiges und präzises Instrument zur Erfassung von Unterrichtsmerkmalen?**

**Lukas Begrich<sup>1</sup>, Julia Käfer<sup>2</sup>, Susanne Kuger<sup>2</sup>, Mareike Kunter<sup>1</sup>**

<sup>1</sup>Johann Wolfgang Goethe-Universität, <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland

Hintergrund

Die Thin Slices Technik stammt ursprünglich aus der Persönlichkeitspsychologie. Sie beruht auf dem ersten Eindruck, den ungeschulte Beobachter von einer Person und/oder einer Situation haben und nutzt daher andere Beurteilungskomponenten und -prozesse als ausführliche Einschätzungen in herkömmlichen Videoanalysen (vgl. Ambady, Bernieri, & Richeson, 2000; Wood, 2014). Sie konnte schon erfolgreich eingesetzt werden, um Merkmale der Lehrkraft und des Unterrichts reliabel und valide einzuschätzen (Ambady & Rosenthal, 1993; Babad, 2005; Begrich, Fauth, Kunter, Klieme, 2017). Einer Studie von Ambady und Rosenthal untersuchte Einschätzungen des ersten Eindrucks von Lehrkräften hinsichtlich breiter Persönlichkeitseigenschaften sowie deren Zusammenhänge mit den Effizienzbeurteilungen dieser Lehrkräfte. Weitere Studien von Babad (2005) sowie Begrich et al. (2017) fanden Hinweise darauf, dass spezifischere Unterrichtsmerkmale von ungeschulten Beobachtern abgegeben werden konnten, selbst wenn die Videogrundlage nur wenige Sekunden umfasste. Im Rahmen des Thin Slices Projekts wurde diese Möglichkeit in einer größer angelegten Studie hinsichtlich mehrerer Unterrichtsmerkmale weiter exploriert. Hierfür wurden zwei Pilotierungs- und eine Hauptstudie durchgeführt, um bestmögliche Bedingungen zur Erfassung von Unterrichtsmerkmalen mithilfe der Thin Slices Technik zu eruieren und schließlich verschiedene Unterrichtsmerkmale in den Videos aus TIMSS und der DESI Studie einzuschätzen.

Der Beitrag dieser Studie

Die vorliegende Studie berichtet aus diesen Kalibrierungsstudien und analysiert die Ergebnisse der Pilotierungs- sowie der Hauptstudie hinsichtlich verschiedener Merkmale der Reliabilität und Validität der mithilfe der Thin Slices Technik erhobenen Beurteilungsdaten. Im Zentrum stehen Auswertungen, welche die Daten der Thin Slices Beobachtungen mit solchen von ausführlichen Videoratings und Schülerangaben zu eben diesem Unterricht vergleichen.

Ergebnisse

Pilotstudie 1 wurde genutzt, um die notwendige Informationsmenge und die Ziehungsart zu bestimmen. Systematisch variiert wurden sowohl die Länge der aus den vollständigen Unterrichtsvideos gezogenen Slices (30 Sekunden versus 90 Sekunden) als auch die Art ihrer Ziehung (event-basiert zu Beginn der Unterrichtsstunde versus randomisiert aus jedem Drittel der Unterrichtsstunde). Die Ergebnisse zeigen bezüglich der Reliabilität und der Validität weder Unterschiede der Darbietungsdauer noch der Ziehungsart der gezeigten Videosequenzen. In Pilotstudie 2 wurde die Spezifität der gezeigten Videosequenz variiert und untersucht welchen Einfluss das Zeigen einer für das zu beurteilende Zielkonstrukt spezifischen Unterrichtssituation auf die Eindrucksbildung hat. In zwei Versuchsgruppen beurteilten ungeschulte Rater Videos der gleichen Lehrkräfte. Einer der Gruppen wurden Unterrichtssituationen gezeigt, in denen eine Fehlerreaktionen der Lehrkräfte zu sehen war. Die andere Gruppe sah randomisiert gezeigte Videosequenzen. Hinsichtlich der Reliabilität und der Validität der abgegebenen Urteile zeigten sich keine bedeutsamen Unterschiede zwischen den beiden Versuchsgruppen.

Die Ergebnisse der Pilotstudien trugen zur Gestaltung der Hauptstudie bei, in der 30 studentische Rater insgesamt 201 Lehrkräfte auf der Grundlage ihres ersten Eindrucks hinsichtlich der Qualität ihres Unterrichts (drei Basisdimensionen guten Unterrichts mit 9 Items) sowie ihres Fehlerumgangs (13 Items) einschätzten. Erste Ergebnisse zeigen gute Reliabilität, d. h. hohe Übereinstimmungen zwischen den Ratern. Weiterhin korrelieren die Thin Slices Urteile teilweise substanziell sowohl mit Urteilen geschulter Rater, die auf Grundlage der vollständigen Unterrichtsvideos abgegeben wurden, als auch mit den Schülerurteilen aus den betreffenden Klassen. Spezifische Korrelationen finden sich vor allem für die Basisdimension der Klassenführung, weniger spezifische, jedoch trotzdem signifikant positive Korrelationen zeigten die Eindrucksratings zur kognitiven Aktivierung und zur konstruktiven Unterstützung. Die Eindrucksratings des instruktionalen Fehlerumgangs korrelierten signifikant mit Schülerurteilen zu der im Unterricht vorherrschenden Fehlerkultur, dem Fehlerfeedback und Fehlerfreundlichkeit ihrer Lehrkraft, sowie mit der Schülereinschätzung hinsichtlich der Nützlichkeit von Fehlern. Anhand dieser und weiterer Analysen soll im Rahmen des Vortrags ein möglichst vollständiger Abgleich der verschiedenen Erhebungsinstrumente zur Erfassung der genannten Unterrichtsmerkmale vorgestellt und so ein möglicher Mehrwert von eindrucksbasierten Erfassungen von Unterrichtsmerkmalen zur Diskussion gestellt werden.

### **Bedeutung eines konstruktiven Umgangs mit Fehlern im Unterricht: Inkrementelle Validität von Thin Slices Ratings**

**Julia Käfer<sup>1</sup>, Susanne Kuger<sup>1</sup>, Lukas Begrich<sup>2</sup>, Eckhard Klieme<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland, <sup>2</sup>Johann Wolfgang Goethe-Universität

#### Theoretischer Hintergrund:

Fehler sind natürliche Bestandteile des Lernens und stellen alltägliche Unterrichtsereignisse dar. In der bildungswissenschaftlichen Forschung herrscht Konsens darüber, dass ein verständnisvoller und kognitiv anregender Umgang mit Fehlern ein wichtiges Unterrichtsmerkmal darstellt, das geeignet ist, um Lernende in ihrer Kompetenz- und Motivationsentwicklung zu unterstützen (Hascher & Hagenauer, 2010; Kunter & Trautwein, 2013). Es wird angenommen, dass der Umgang mit Fehlern eine „soziale und kognitive Doppelfunktion“ (Klieme, 2006, S. 772) aufweist, da er sowohl affektive und motivationale wie auch kognitive und behaviorale Reaktionen hervorruft (Dresel et al., 2013; Link, 2012). Folglich können für motivational-affektive Aspekte des Fehlerumgangs einerseits und kognitive Aspekte andererseits spezifische Effekte auf Leistung und Motivation von Schülerinnen und Schülern angenommen werden.

Bei der Erfassung des Fehlerumgangs im Unterricht werden meist Schüler- und/oder Lehrereinschätzungen herangezogen. Häufig werden auch Einschätzungen geschulter Beobachter genutzt, die den Fehlerumgang anhand von Unterrichtsvideos beurteilen. Studien zur Thin-Slices-Technik in der Unterrichtsforschung konnten allerdings zeigen, dass auch Einschätzungen, die durch das Ansehen kurzer Videoausschnitte („thin slices“) gefällt werden, ebenfalls relevante und zuverlässige Informationsquellen für unterschiedliche Unterrichtsprozesse darstellen können (Babad, 2005; Begrich, Fauth, Kunter & Klieme, 2017). Die Erfassung des Fehlerumgangs im Unterricht wurde mit dieser Technik jedoch noch nicht untersucht.

#### Fragestellung:

Der vorliegende Beitrag beschäftigt sich mit der Bedeutung des Fehlerumgangs für die Leistungs- und Motivationsentwicklung von Schülerinnen und Schülern, und vergleicht dabei drei unterschiedliche Datenquellen: Schülerangaben, hoch-inferente Ratings und Thin-Slices-Ratings. Der Schwerpunkt des Beitrags liegt auf der Untersuchung der prädiktiven Kraft der Thin-Slices-Ratings, die mit den anderen beiden (bereits etablierten) Datenquellen zum Fehlerumgang im Unterricht verglichen werden.

#### Methode:

Als Datengrundlage dient die deutsche Large-Scale-Studie Deutsch-Englisch-Schülerleistungen-International (DESI; DESI-Konsortium, 2008). In der DESI-Studie wurden Neuntklässler zu Beginn und am Ende des Schuljahres 2003/04 getestet bzw. befragt. Im Schülerfragebogen wurde die Schülereinschätzung hinsichtlich verschiedener Aspekte des Fehlerumgangs im Unterricht erfasst (Fehlerfreundlichkeit der Lehrkraft, Rückmeldung der Lehrkraft zu Fehlern, wahrgenommene Nützlichkeit von Fehlern). In ausgewählten Klassen wurde zudem der Unterricht videografiert und die Videos mithilfe hoch-inferenter Ratings ausgewertet. Bei der Beurteilung der Videos durch geschulte Rater wurde zwischen dem Fehlerumgang auf Lehrerseite und dem Fehlerumgang auf Schülerseite unterschieden.

Für die Erfassung des Fehlerumgangs mithilfe der Thin-Slices-Technik wurden die in der DESI-Studie videografierten Unterrichtsstunden erneut beurteilt. Dieses Mal jedoch durch ungeschulte Rater, die nur kurze Ausschnitte des Unterrichts (3x10 Sekunden) zu sehen bekamen. Auch bei den Thin-Slices-Ratings wurden verschiedene Aspekte des Fehlerumgangs eingeschätzt (u.a., Fehlerumgang der Lehrkraft, Negatives Fehlerklima in der Klasse).

Für die Analyse der Ratings wurde eine DESI-Substichprobe von 1140 Schülerinnen und Schülern aus 101 Klassen genutzt. Längsschnittliche Mehrebenenstrukturgleichungsmodelle (ML-SEM; Marsh et al., 2012; Lüdtke, Marsh, Robitzsch, & Trautwein, 2011) wurden verwendet, um die Prädiktivität der unterschiedlichen Datenquellen zum Fehlerumgang für die Leistungs- und Motivationsentwicklung von Schülerinnen und Schülern zu untersuchen.

#### Ergebnisse:

Hinsichtlich der Schülereinschätzungen zeigt sich, dass die Aspekte „Fehlerfreundlichkeit der Lehrkraft“ und „Wahrgenommene Nützlichkeit von Fehlern“ die mittlere Leistung der Schülerinnen und Schüler bedeutsam vorhersagen, wenngleich die Effekte eher gering ausfallen. Weitaus größer sind die Effekte auf die Klassenmotivation, die von allen drei untersuchten Aspekten des Fehlerumgangs vorhergesagt wird.

Die Untersuchung der hoch-inferenten Ratings hat ergeben, dass weder der eingeschätzte Fehlerumgang seitens der Lehrkraft noch seitens der Schüler die Leistung der Klasse vorhersagen. Allerdings lässt sich die Motivation der Klasse durch den eingeschätzten Fehlerumgang der (Mit-)Schüler vorhersagen.

Die Ergebnisse zur prädiktiven Kraft der Thin-Slices-Ratings zeigen keine bedeutsamen Effekte, für keine der vier untersuchten Aspekte. Die Einschätzung des Fehlerumgangs anhand kurzer Ausschnitte des Unterrichtsgeschehens scheint folglich nicht geeignet zu sein, um die Leistungs- und Motivationsentwicklung von Schülerinnen und Schülern vorherzusagen.

## C12

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H119

### Voraussetzungen von frühpädagogischen Fachkräften für die Weiterentwicklung von sprachlichen Bildungsangeboten und Formen der Zusammenarbeit mit Familien

Chair(s): **Nadine Wieduwilt** (Freie Universität Berlin, Deutschland)

DiskutantIn(nen): **Wilfried Smidt** (Universität Innsbruck)

Pädagogische Fachkräfte in Kindertageseinrichtungen sehen sich heute einem vielfältigen Aufgabenspektrum ausgesetzt, das sich verstärkt auf die Kompensation von Benachteiligung von Kindern aus sozial schwachen Familien und Kindern mit Migrationshintergrund bezieht (Anders, 2012). Die Fachkräfte tragen eine hohe Verantwortung für sprachliche Bildungsarbeit, sind angehalten, Eltern in die Bildungsarbeit einzubeziehen und die besonderen Bedürfnisse von Kindern und Familien mit Fluchterfahrung zu berücksichtigen. Der Kita-Alltag ist gekennzeichnet durch vielfach nicht standardisierbare Situationen, die ein schnelles Finden von jeweils angemessenen Handlungsoptionen erfordern (Fröhlich-Gildhoff et al., 2011). Um diesen Herausforderungen gerecht zu werden, benötigen pädagogische Fachkräfte eine Vielzahl an professionellen Kompetenzen, wie beispielsweise Wissen und pädagogische Überzeugungen. Dabei wird davon ausgegangen, dass die Kompetenzfacetten Einfluss auf die Interaktion zwischen Fachkräften und Kindern bzw. auch zwischen Fachkräften und Eltern haben (Roux & Tietze, 2007). Zudem deuten Untersuchungen darauf hin, dass das Führungsverständnis von Leitungskräften Einfluss auf den Umgang mit den einzelnen Fachkräften und deren Unterstützung zur fachlichen Weiterentwicklung hat (Ballaschk & Anders, eingereicht). Es wird angenommen, dass professionelle Kompetenzen erlernbar sind und durch Aus-, Fort- und Weiterbildung verändert werden können (Anders, 2012).

Bundesweite Initiativen, wie die Programme „Schwerpunkt-Kitas Sprache & Integration“ (BMFSFJ, 2015) oder „Sprach-Kitas: Weil Sprache der Schlüssel zur Welt ist“ (BMFSFJ, 2016), greifen den Gedanken der Kompetenzentwicklung durch Fort- und Weiterbildung auf und unterstützen pädagogische Fachkräfte durch spezifische Funktionsstellen, wie SprachexpertInnen und zusätzliche Fachberatungen. Damit werden verschiedene Unterstützungsebenen in den Blick genommen und sowohl Fachkräfte als auch Leitungskräfte in Bezug auf die Themenschwerpunkte alltagsintegrierte sprachliche Bildung, inklusive Pädagogik und Zusammenarbeit mit Familien fachlich weiterentwickelt. Neben den SprachexpertInnen, die die Themenschwerpunkte gezielt mit den Fachkräfteteams bearbeiten sollen, haben die zusätzlichen Fachberatungen insbesondere die Aufgabe, Leitungskräfte untereinander zu vernetzen, sie in der Arbeit mit ihren Teams zu unterstützen und gleichzeitig Leitungskräfte in ihrem professionellen Selbstverständnis als Führungskräfte einer Einrichtung weiterzuentwickeln. Auch SprachexpertInnen und zusätzliche Fachberatungen sollten eine Vielzahl an professionellen Kompetenzen bereitstellen, um ihrer Rolle als professionelle Unterstützungssystem gerecht zu werden.

Obzwar die Bedeutung professioneller Kompetenzen frühpädagogischer Fachkräfte hervorgehoben wird, sind diese in der frühkindlichen empirischen Bildungsforschung insgesamt betrachtet noch wenig erforscht. Das Symposium hat daher zum Ziel, verschiedene Kompetenzfacetten auf unterschiedlichen Akteursebenen zu beleuchten. Mit diesen Schwerpunkten reiht sich das Symposium inhaltlich in das Tagungsthema "Professionelles Handeln als Herausforderung für die Bildungsforschung" ein. Über den Einsatz quantitativer und qualitativer Methoden werden unterschiedliche Handlungsperspektiven bildungswissenschaftlicher Forschung aufgezeigt. Beitrag eins befasst sich mit sprachpädagogischen Überzeugungen von SprachexpertInnen, die für die Umsetzung alltagsintegrierter sprachlicher Bildung als relevant erachtet werden. Der zweite Beitrag beschäftigt sich spezifisch mit multikulturellen Überzeugungen von zusätzlichen Fachberatungen und spricht eine zweite Akteursebene an. Dabei wird der Frage nachgegangen, welchen Einfluss das Führungsverständnis von Fachberatungen auf deren Umgang mit Multikulturalität hat, um Rückschlüsse auf deren Handlungsstrategien zur Weiterentwicklung der Fachkräfteteams ziehen zu können. Ein dritter Beitrag thematisiert Wissen als Facette professioneller Handlungskompetenz. Im Fokus steht die Frage, welches Wissen zusätzliche Fachberatungen im Kompetenzbereich Zusammenarbeit mit Familien haben. Der vierte Beitrag befasst sich schließlich mit der Bedeutsamkeit der Zusammenarbeit mit Eltern für die häusliche Lernumwelt und kindliche Sprachentwicklung und fokussiert damit auf die Adressaten frühkindlicher Bildung.

*Beiträge des Symposiums*

### Dimensionalität sprachpädagogischer Überzeugungen von frühpädagogischen Fachkräften

**Nadine Wieduwilt<sup>1</sup>, Simone Lehl<sup>2</sup>, Yvonne Anders<sup>1</sup>**

<sup>1</sup>Freie Universität Berlin, <sup>2</sup>Otto-Friedrich-Universität Bamberg

#### Theoretischer Hintergrund

Im Rahmen institutioneller frühkindlicher Bildung wird der sprachlichen Bildung eine große Bedeutung zugesprochen, wobei unterschiedliche Ansätze sprachpädagogischer Arbeit diskutiert werden (Fried, 2013). Vor dem Hintergrund einer geringen/kurzfristigen Wirksamkeit additiver Sprachförderprogramme wurde in den vergangenen Jahren der Ansatz der alltagsintegrierten sprachlichen Bildung stark forciert (BMFSFJ, 2015; 2016), der – anders als die trainingsbasierten Sprachförderprogramme mit stark instruktivistischem Charakter – das Potenzial qualitativ hochwertiger Dialoge in Routinesituationen des frühpädagogischen Alltags fokussiert. Die Ausgestaltung und Umsetzung einer alltagsintegrierten sprachlichen Bildung bedürfen vielfältiger professioneller Kompetenzen. Als eine zentrale Facette professioneller Handlungskompetenz werden pädagogische Überzeugungen beschrieben, worunter implizite, handlungsleitende Sichtweisen von frühpädagogischen Fachkräften verstanden werden (Levin et al., 2015). Überzeugungen frühpädagogischer Fachkräfte sind insgesamt und für den sprachlichen Bildungsbereich im Speziellen noch wenig erforscht. Bisherige Forschungsarbeiten konnten zeigen, dass pädagogische Überzeugungen zum kindlichen Lernen frühpädagogischer Fachkräfte hinsichtlich konstruktivistischer und instruktivistischer (Benz, 2012) oder zusätzlich ko-konstruktivistischer Überzeugungen (Mischo et al., 2012) unterschieden werden können. Hinsichtlich der Überzeugungen speziell zur sprachpädagogischen Arbeit liegen solche Befunde bislang jedoch nicht vor.

#### Fragestellung

Vor diesem Hintergrund wird in der vorliegenden Studie die Dimensionsstruktur sprachpädagogischer Überzeugungen frühpädagogischer Fachkräfte untersucht. Konkret wird dabei geprüft, ob sich sprachpädagogische Überzeugungen hinsichtlich der Dimensionen „Sprachförderung“ und „alltagsintegrierte sprachliche Bildung“ unterscheiden lassen. Darüber hinaus werden strukturelle Merkmale der Kita-Gruppe (z.B. Anteil der Kinder mit Migrationshintergrund) sowie Merkmale auf Seiten der pädagogischen Fachkräfte (z.B. Qualifikation) in den Blick genommen, die die Variation in den Dimensionen vorhersagen.

## Methode

Die Daten stammen aus der Evaluationsstudie zum Bundesprogramm „Schwerpunkt-Kitas Sprache & Integration“ und wurden im Frühjahr und Sommer 2015 im Rahmen einer Onlinebefragung erhoben. Befragt wurden die zusätzlichen Fachkräfte für sprachliche Bildung aus dem Bundesprogramm (N = 231), die die primäre Aufgabe hatten, das Kita-Team für die alltagsintegrierte sprachliche Bildungsarbeit zu qualifizieren. Die sprachpädagogischen Überzeugungen wurden über insgesamt 16 Items mittels einer Zustimmungsskala erfasst. Als Antwortoptionen wurden die Kategorien „Stimme gar nicht zu“, „Stimme eher nicht zu“, „Stimme eher zu“ und „Stimme voll zu“ vorgegeben. Die Aussage „Kinder lernen Sprache am besten, wenn sie dazu ermutigt werden, über ihre eigenen alltäglichen Erfahrungen zu berichten“ stellt ein Beispiel für die Dimension „alltagsintegrierte sprachliche Bildung“ dar. Das Item „Mit speziellen Lehrmitteln oder Förderprogrammen zur Sprachentwicklung werden die Kinder am besten unterstützt“ gilt als ein Beispiel der Dimension „Sprachförderung“. Erste Analyseschritte sahen Reliabilitätsanalysen sowie die explorative faktorenanalytische Überprüfung der Dimensionsstruktur sprachpädagogischer Überzeugungen vor. Für die Analyse wurde die Statistiksoftware SPSS 23 genutzt.

## Ergebnisse

Die Deskriptiv-Statistik verweist auf eine hohe Zustimmung (M = 3,7) der zusätzlichen Fachkräfte in Bezug auf die Aussagen zur alltagsintegrierten sprachlichen Bildung. Mit Blick auf die Aussagen zur Sprachförderung liegt die Zustimmung niedriger (M = 2,3). Über Reliabilitätsanalysen wurden in einem ersten Schritt die internen Konsistenzen der beiden Skalen geprüft. Dabei hat sich für die Dimension „Sprachförderung“ eine gute interne Konsistenz gezeigt (7 Items,  $\alpha = .798$ ), während dies für die Dimension „alltagsintegrierte sprachliche Bildung“ nicht ermittelt werden konnte. Daraufhin wurde in einem zweiten Schritt eine explorative Faktorenanalyse durchgeführt. Die Hauptkomponentenanalyse mit obliminer Rotation hat eine anderweitige Faktorenstruktur aufgezeigt. Folgende vier Faktoren konnten vorläufig extrahiert werden: (1) Sprachförderung mittels spezifischer Programme, (2) Sprachförderung mit erklärendem Charakter durch die Fachkraft, (3) sprachliche Bildung mit Fokus auf die Eigeninitiative des Kindes und (4) sprachliche Bildung durch gezielte Fragen der Fachkraft. Weiterführende Analysen sollen die Dimensionsstruktur sprachpädagogischer Orientierungen erneut mittels einer konfirmatorischen Faktorenanalyse prüfen. Darüber hinaus werden Merkmale auf Seiten der pädagogischen Fachkräfte und Strukturmerkmale der Gruppe in den Blick genommen, um die Variation in den Dimensionen vorherzusagen.

## **Führungskonzepte im Umgang mit kultureller Diversität**

**Itala Ballaschk, Csaba Kurucz**  
Freie Universität Berlin

### Theoretischer Hintergrund

Kindertageseinrichtungen haben aktuell große Herausforderungen zu bewältigen. Die in den letzten Jahren angestiegene Zahl an Kindern aus asylsuchenden Familien führt zu einer steigenden soziokulturellen Diversität in vielen Einrichtungen. Dabei sind Kinder und Familien mit Fluchthintergrund spezifischen Risikolagen ausgesetzt, die sich insbesondere im Bereich der Bildungs- und Chancengleichheit manifestieren. Ein Faktor, der den Bildungserfolg maßgeblich positiv beeinflussen kann, ist die sprachliche Ausdrucksfähigkeit. Sprache gilt als Schlüsselkompetenz in der kindlichen Entwicklung (Weinert, 2011) und als Voraussetzung für gesellschaftliche Teilhabe. Kindertageseinrichtungen können einen entscheidenden Beitrag zur Realisierung einer inklusiven Pädagogik leisten, um Bildungsbenachteiligungen frühzeitig auszugleichen. Die Evaluationsstudie zum Bundesprogramm „Schwerpunkt-Kitas Sprache & Integration“ konnte aufzeigen, dass eine kontinuierliche fachliche Unterstützung der Einrichtungen und ihrer Fachkräfte für eine Implementierung sprachlicher Bildung im professionellen pädagogischen Alltag von großer Bedeutung ist (Anders et al., 2015). Das Nachfolgeprogramm „Sprach-Kitas: Weil Sprache der Schlüssel zur Welt ist“ bezieht als weitere Unterstützungsressource Fachberatungspersonal ein. Fachberatungen sollen als zentrale Beratungsinstanzen agieren und gezielt Qualitätsentwicklungsprozesse mit Fokus auf inklusive Bildung unterstützen. Zudem sollen sie für Leitungskräfte als Rollenvorbilder für effektive Führung im Sinne von Leadership dienen. Forschungsbefunde weisen auf die Bedeutung des Führungsverständnisses von Leitungskräften für die Qualitätsentwicklung in Kindertageseinrichtungen hin (Ballaschk, Anders & Flick, in Druck). Im Rahmen der Evaluationsstudie zum Bundesprogramm „Sprach-Kitas“ wird an diese Befunde anknüpfend davon ausgegangen, dass die Bedeutung, die Fachberatungen als externe Führungskräfte einer inklusiven Pädagogik zusprechen, Effekte auf Einrichtungsebene nach sich zieht. Bisher fehlt es jedoch an evidenzbasierter Forschung zum Nutzen von Fachberatung für die Qualitätsentwicklung in Kindertageseinrichtungen. Auch gibt es bislang keine Befunde zum Führungsverständnis von Fachberatungen und dessen Auswirkungen auf den Umgang mit Fachkräften und kultureller Diversität. Die vorliegende Studie untersucht daher das Führungsverständnis von Fachberatungen und deren Auswirkungen auf den Umgang mit Fachkräfteteams zum Themenschwerpunkt kulturelle Diversität.

### Fragestellungen

Im Rahmen der vorliegenden Studie wird zum einen untersucht, welches Verständnis von Führung sich auf Ebene der Fachberatungen zeigt. Zum anderen wird betrachtet, welche Bedeutung die befragten Fachberatungen dem Thema kultureller Diversität zusprechen. Zudem wird untersucht, welche Handlungsstrategien sich aus den einzelnen Führungskonzepten erüieren lassen, die Rückschlüsse auf den Umgang mit den Fachkräften im Hinblick auf den Themenschwerpunkt kulturelle Diversität zulassen.

### Methode

Die Studie ist als Mixed-Methods-Untersuchung angelegt. Es wurden bundesweit von Dezember 2016 bis März 2017 101 Fachberatungen online befragt, die am Bundesprogramm „Sprach-Kitas: Weil Sprache der Schlüssel zur Welt ist“ teilnehmen. Mit diesen Daten wurde eine Clusteranalyse durchgeführt, um Leadershipprofile zu erfassen. Vertiefend dazu wurden 17 Fachberatungen qualitativ mit einem halb-standardisierten Leitfadens episodisch befragt, um einem impliziten Führungsverständnis sowie professionellen Einstellungen und Handlungsstrategien zum Thema kulturelle Diversität möglichst nahe zu kommen (Flick, 2014). Ausgewertet wurde das Datenmaterial über die Methode des thematischen Kodierens (Flick, 2011).

### Ergebnisse

Die Clusteranalyse identifiziert zwei Leadershipprofile. Während die Fachberatungen des zweiten Profils die einzelnen Leadership-Aspekte im Durchschnitt als eher zutreffend und als voll und ganz zutreffend einschätzen, fallen die Einschätzungen der Fachberatungen des ersten Profils in allen Aspekten niedriger aus. Das bedeutet, dass zwar alle Fachberatungen ihre Rolle darin sehen, Austausch und Reflexion für die Teams zu ermöglichen, aber Unterschiede insbesondere bei den Aspekten zur gemeinsamen Zielorientierung bestehen. Die Analysen der qualitativen Interviews zeigen ein ambivalentes Verhältnis der Fachberatungen zum Führungsbegriff. Sie lehnen einen negativ konnotierten Führungsbegriff

ab. Gleichzeitig deuten ihre Handlungsstrategien auf ein Konzept von Führung im Sinne von Leadership hin. Kulturelle Diversität wird als klarer Bestandteil des eigenen professionellen Handelns der befragten Fachberatungen wahrgenommen. Die hohe kulturelle Diversität in den betreuten Einrichtungen wird als Herausforderung benannt, mit der es konstruktiv umzugehen gilt. Gleichzeitig bleiben konkrete Handlungsstrategien jedoch vage.

## **Welches Wissen haben Fachberatungen im Kompetenzbereich Zusammenarbeit mit Familien?**

**Theresia Hummel, Yvonne Anders**  
Freie Universität Berlin

### Theoretischer Hintergrund

Die Anforderungen an Kindertageseinrichtungen sind in den letzten Jahren in Deutschland stark gestiegen. Neben einer qualitativ hochwertigen Bildung, Betreuung und Erziehung der Kinder soll heute auch eine entsprechend gute Zusammenarbeit mit den Familien realisiert werden. Die Umsetzung einer intensiven Zusammenarbeit mit Familien in der pädagogischen Praxis bedarf vielfältiger professioneller Kompetenzen. Als eine zentrale Facette professioneller Handlungskompetenz wird das Professionswissen beschrieben (Anders, 2012). Das Professionswissen frühpädagogischer Fachkräfte ist insgesamt und für das Themenfeld Zusammenarbeit mit Eltern im Speziellen noch wenig erforscht. Bisherige Studien verweisen lediglich auf allgemeine Kompetenzeinschätzungen im Bereich Zusammenarbeit mit Eltern und liefern nur bedingt positive Ergebnisse (Behr & Walter, 2012; Schneewind et al., 2012). Ein nicht unerheblicher Teil der pädagogischen Fachkräfte fühlt sich im Bereich Zusammenarbeit mit Eltern überfordert (Textor, 1992; Sturzbecher & Bredow, 1998) und unzureichend ausgebildet (Viernickel et al., 2013; Fuchs-Rechlin, 2007). Das Bundesprogramm „Sprach-Kitas: Weil Sprache der Schlüssel zur Welt ist“ reagiert auf diesen Qualifizierungsbedarf mit der Finanzierung einer zusätzlichen Fachberatung als Unterstützungsressource. Diese hat die Aufgabe, das Team auch im Themenfeld Zusammenarbeit mit Familien fachlich zu begleiten und zu beraten. Der Fokus liegt hierbei insbesondere auf der Weitergabe und Weiterentwicklung von Professionswissen.

### Fragestellung

Vor dem Hintergrund einer mangelnden Qualifizierung frühpädagogischer Fachkräfte im Kompetenzbereich Zusammenarbeit mit Familien wird in der Studie untersucht, welches Wissen zusätzliche Fachberatungen in diesem Themenfeld haben. Dafür wird einerseits der Frage nachgegangen, welche Methoden der Zusammenarbeit mit Familien die zusätzlichen Fachberatungen kennen. Darüber hinaus wird untersucht, inwieweit sich unterschiedliche Wissensprofile identifizieren lassen. Schlussendlich wird der Frage nachgegangen, ob sich Unterschiede im Wissen vor dem Hintergrund der fachlichen Voraussetzungen der zusätzlichen Fachberatungen ergeben

### Methode

Im Rahmen der Evaluation des Bundesprogramms „Sprach-Kitas“ wurden im Zeitraum von Dezember 2016 bis März 2017 zusätzliche Fachberatungen (N = 99) standardisiert und onlinebasiert befragt. Das Wissen der zusätzlichen Fachberatungen wurde über ein offenes Antwortformat erfragt, um die Vielfalt möglicher Antworten abbilden zu können. Die offenen Antworten wurden mit Hilfe der Computersoftware EXCEL kategorisiert und quantifiziert. Mithilfe einer Clusteranalyse wird untersucht, inwieweit bestimmte Methodenformen im Wissen eher miteinander einhergehen und inwieweit diese Typenbildung auf die fachlichen Voraussetzungen auf Fachberatungsebene (Berufserfahrung, Berufsausbildung und Fortbildungsaktivität) zurückzuführen sind.

### Ergebnisse

Bisherige Analysen verweisen auf eine große Vielfalt an Methoden, die von den zusätzlichen Fachberatungen im Bereich der Zusammenarbeit mit Familien genannt werden. Die Antworten konnten in folgende Kategorien (Methodenbereiche) eingeteilt werden: Einzelkontakte, Angebote zur Partizipation von Eltern, Angebote unter Beteiligung von Eltern und Erzieherinnen, Angebote nur für Eltern, Angebote für bestimmte Elterngruppen, Begleitung von Übergängen, informative Angebote, Elternbildung und sozialraumorientierte Kooperation. Im Durchschnitt nennen die zusätzlichen Fachberatungen 10-11 Methoden und differenzieren zwischen 4-5 Methodenbereichen. Mit knapp 92% werden am häufigsten Formen aus der Kategorie Einzelkontakte genannt. 86% der zusätzlichen Fachberatungen nennen Angebote zur Partizipation der Eltern. Methoden im Bereich sozialraumorientierte Kooperation werden hingegen nur von 16% der zusätzlichen Fachberatungen genannt. Die realisierte Stichprobe spiegelt jedoch nicht nur hinsichtlich Wissen sondern auch fachlicher Voraussetzungen auf Fachberatungsebene eine große Varianz wider. Die Mehrheit der zusätzlichen Fachberatungen verfügt über einen (Fach-)Hochschulabschluss. Ferner können die meisten befragten Fachberatungen als erfahrene Fachkräfte angesehen werden, die im Durchschnitt über 15 Jahre Berufserfahrung in Kindertageseinrichtungen sammeln konnten, jedoch etwa 2 ½ Berufsjahre nur als Fachberatung tätig waren. In nachfolgenden Analysen soll daher der Frage nachgegangen werden, ob sich Unterschiede in Wissensprofilen vor dem Hintergrund der fachlichen Voraussetzungen der zusätzlichen Fachberatungen ergeben.

## **Direkte und indirekte Bedeutsamkeit der Zusammenarbeit mit Eltern für die kindliche Sprachentwicklung**

**Simone Lehr<sup>1</sup>, Manja Flöter<sup>2</sup>, Yvonne Anders<sup>3</sup>**

<sup>1</sup>Otto-Friedrich-Universität Bamberg, <sup>2</sup>PädQUIS® gGmbH, <sup>3</sup>Freie Universität Berlin

### Theoretischer Hintergrund

Die Zusammenarbeit von pädagogischen Fachkräften und Eltern ist neben der direkten Interaktion mit dem Kind und der Gestaltung der Lernumwelt Kindergarten ein zentrales Element pädagogischer Praxis in der institutionellen frühkindlichen Bildung und Betreuung. Auf der einen Seite besteht weitestgehend Einigkeit darüber, dass die Zusammenarbeit zwischen pädagogischen Fachkräften und Eltern einen wichtigen Aspekt zur Unterstützung der sozialen und kognitiven Entwicklung der Kinder darstellt (z.B. Fröhlich-Gilthoff et al., 2011). Auf der anderen Seite, ist die empirische Basis zu Dimensionen, Bedingungen und Bedeutsamkeit der Zusammenarbeit mit Eltern im frühkindlichen Bereich insbesondere in Deutschland schmal. Die empirische Befundlage stützt sich hier eher auf Formen und Ausgestaltung der Zusammenarbeit (z.B. Hildenbrand & Köhler, 2010; Lemke et al., 2007). Die wenigen Studien, die sich mit dem Zusammenhang der Zusammenarbeit mit Eltern und kindlicher Kompetenzentwicklung beschäftigen deuten auf eine inkonsistente Befundlage hin, je nachdem, ob Querschnitt-, oder Längsschnittdaten verwendet wurden. Anders et al. (2016) z.B. konnte im Längsschnitt zeigen, dass sich einige Aspekte der Zusammenarbeit mit Eltern als positiv assoziiert mit der kindlichen Sprachentwicklung erweisen, während Durda & Buntins (2016) im Querschnitt keine oder sogar negative Beziehungen fanden. Darüber hinaus ist bislang ungeklärt, welche dahinterliegenden Mechanismen zur Erklärung eines möglichen positiven Effekts herangezogen werden können. Möglich wäre, dass die Zusammenarbeit mit Eltern die häusliche Lernumwelt anreichert, oder dass sie lediglich Ausdruck einer insgesamt

anregenden und wertschätzenden Lernumwelt im Kindergarten ist. Für beide Aspekte wurde in bisherigen Studien gezeigt, dass sie positiv mit der kindlichen Kompetenzentwicklung in Beziehung stehen (z.B. Lehl, Kluczniok & Roßbach, 2016).

#### Fragestellung

Vor diesem Hintergrund wird in der vorliegenden Studie untersucht, ob die von Anders et al. (2016) gefundenen Zusammenhänge von unterschiedlichen Dimensionen der Zusammenarbeit mit Eltern und der Sprachentwicklung von Kindergartenkindern durch die Qualität der häuslichen Lernumwelt mediiert werden oder/ und durch die Qualität der Einrichtungen moderiert werden.

#### Methode

Datengrundlage sind Daten von 1.331 Kindern und ihren Eltern sowie von den von ihnen besuchten 335 Einrichtungen mit ihren pädagogischen Fachkräften aus der Evaluation zum Bundesprogramm „Schwerpunkt-Kitas Sprache & Integration“. Sie wurden über einen Zeitraum von 4 Jahren ab Ende 2012 begleitet. Die pädagogischen Fachkräfte und die Eltern wurden online (bzw. persönlich) zu Möglichkeiten und Formen der Zusammenarbeit mit Eltern befragt (z.B. Weitergabe von Medien, spezifische Hinweise zur Anreicherung der häuslichen Lernumwelt). Mittels standardisierter Ratingverfahren wurden die Einrichtungen zusätzlich hinsichtlich der Qualität der Einrichtungen beobachtet. Die häusliche Lernumwelt wurde ebenfalls durch standardisierte Befragung der Eltern und Beobachtung erfasst. Zur Erfassung der sprachlichen Kompetenzen wurden die Kinder in Einzeltestungen in den Einrichtungen mit Hilfe standardisierter Verfahren hinsichtlich des Satzverstehens und des Wortschatzes untersucht (Dunn & Dunn, 2007; Grimm, Aktas & Frevert, 2000; 2001).

Zur Beantwortung der Fragestellung wurden Pfadmodelle mit Hilfe von MPlus 6.1 (Muthén & Muthén, 1998-2012) spezifiziert, die sowohl direkte, indirekte als auch moderierende Effekte untersuchen.

#### Ergebnisse

Erste Befunde zeigen, dass sprach-spezifische Zusammenarbeit mit Eltern stärker mit der Sprachentwicklung assoziiert ist als allgemeine Methoden der Elternzusammenarbeit. Die Effekte sind unabhängig von der Qualität des Kindergartens.

Weiterführende Analysen sollen zeigen, welche Effekte die Zusammenarbeit mit Eltern für die Sprachentwicklung in Kindergärten mit variierender Qualität und für bestimmte Subgruppen von Kindern aufweist.

## C13

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H118

### **Reflexion bei (angehenden) Lehrkräften als Antrieb für professionelle Weiterentwicklung – Eine Herausforderung für Bildungsangebote und die empirische Bildungsforschung**

Chair(s): **Iika Benner** (Justus-Liebig-Universität Gießen, Deutschland), **Amina Fraij** (Justus-Liebig-Universität Gießen, Deutschland), **Jochen Wissinger** (Justus-Liebig-Universität Gießen, Deutschland)

DiskutantIn(nen): **Tobias Leonhard** (Fachhochschule Nordwestschweiz, Pädagogische Hochschule, Schweiz)

Der Umgang mit nicht abschließend planbaren Situationen erfordert, dass Lehrkräfte ihr eigenes Handeln und dessen Wirkung auf der Basis theoretischen und empirischen Wissens reflexiv hinterfragen. Reflexion dient dabei als Motor der (Weiter-)Entwicklung professionellen Handelns (u. a. Roters, 2012; Wyss, 2013) und kann mit Blick auf die pädagogische Praxis als vertieftes, lösungsorientiertes Nachdenken und strukturiertes Analysieren über die in der Praxis auftretenden unklaren und inkohärenten Situationen beschrieben werden (Clarà, 2015; Copeland, Birmingham, de la Cruz, & Lewin, 1993; Hatton & Smith, 1995). Reflexion wird oft als Prozess mit spezifischen Komponenten gesehen (u. a. Dewey, 2009; Korthagen, 1985), von dem angenommen werden kann, dass in ihm zwischen den eigenen Dispositionen (bspw. Wissen, Fähigkeiten, Einstellungen) und der eigenen Performanz in der Situation eine Verbindung hergestellt wird, um die eigenen Dispositionen und/oder eigene Performanz (weiter-)zuentwickeln (Copeland et al., 1993; Forster, 2014).

Eine hohe Reflexivität im Kontext der pädagogischen Arbeit bietet die Grundlage für professionelle pädagogische Handlungen und Haltungen (Fiebert & Solzbacher, 2014). Helsper schlägt diesbezüglich eine doppelte Professionalisierung von Lehrkräften vor: Einerseits sei ein „praktisch-pädagogischer Habitus“ als primärer Habitus im Schul- und Unterrichtsalltag zu entwickeln. Notwendig für eine qualitativ hochwertige Arbeit sei aber zusätzlich ein „wissenschaftlich-reflexiver“ Habitus als Sekundärhabitus (Helsper, 2008). Eine mögliche Verbindung beider Habitus könnte sich im Leitbild des „reflective practitioner“ (Schön, 1983) ergeben; entsprechende Kompetenzen werden vielfach als relevant beschrieben bzw. im Kontext der Lehrerbildung gefordert (Altrichter & Posch, 2007). Zu Fragen, in welcher Weise sich Reflexivität als Kompetenz im Laufe der Lehrerbildung entwickelt und zeigt und wie sie auch über die Phasen hinweg im Aufbau unterstützt werden kann, besteht gegenwärtig noch erkennbarer Forschungsbedarf. Daher widmet sich das Symposium aus verschiedenen Perspektiven der Frage, wie sich Reflexivität bei (angehenden) Lehrkräften in unterschiedlichen Professionalisierungsphasen initiieren, erfassen und (weiter-)entwickeln lässt.

Das Symposium nähert sich schrittweise verschiedenen Aspekten von Reflexion. In einem ersten Beitrag wird zunächst das Konstrukt der Reflexionskompetenz kritisch diskutiert und auf den aktuellen Stand der Forschung bezogen sowie mögliche Operationalisierungen mit Blick auf die Erfassung von Reflexivität vorgestellt. Die zwei folgenden Beiträge befassen sich mit den Reflexionsprozessen von (angehenden) Lehrkräften in unterschiedlichen Settings, zum einen in der reflexiven Arbeit vor und während des Studiums, zum anderen in der Verwendung von Schülerrückmeldungen als Ausgangspunkt für Unterrichtsreflexionen. Hier entsteht eine Schnittstelle zum letzten Beitrag, der die Professionalisierung von Lehrkräften durch die Anregung von Reflexion im Rahmen unterschiedlicher universitärer Fortbildungsformate in den Blick nimmt.

#### *Beiträge des Symposiums*

### **Reflexionskompetenz angehender Lehrkräfte – Theoretische Modellierung und empirische Erfassung eines komplexen Konstrukts**

**Amina Fraij, Claudia von Aufschnaiter, Daniel Kost, Sophie Kirschner**  
Justus-Liebig-Universität Gießen, Deutschland

#### Theoretischer Hintergrund

Ausgangspunkt unserer Arbeit bildet die Frage nach der theoretischen Modellierung und der empirischen Erfassbarkeit von Reflexionskompetenz. Wir gehen davon aus, dass Kompetenz die Verknüpfung von Dispositionen, situationsspezifischen Fähigkeiten und Performanz umfasst (Blömeke, Gustafsson, & Shavelson, 2015). Den reflexionsbezogenen Dispositionen können u. a. Wissensbestände darüber, was eine (hochwertige) Reflexion kennzeichnet (z. B. Larivee, 2008) und welche Prozesskomponenten sie enthält (z. B. Korthagen, 1985; 2001), zugeordnet werden sowie z. B. Einstellungen über das Reflektieren. Die situationsspezifischen Fähigkeiten umfassen die Wahrnehmung von Situationen, die Reflexion zulassen bzw. ermöglichen, die Interpretation dieser Situationen sowie das Finden von Entscheidungen. Damit zusammenhängend entsteht Performanz als (beobachtbare) reflexionsbezogene Handlung, die wiederum Rückschlüsse auf die eigenen zugrundeliegenden Skills und Dispositionen zulässt.

In der Forschung finden sich unterschiedliche Versuche, Reflexion(skompetenz) zu erfassen, die von Selbstauskünften über verschiedene Facetten eigener Dispositionen (u. a. Korthagen & Wubbels, 1985; Felten, 2005; Niggli, 2004) bis zu eher verhaltensnäheren Versuchen reichen, Reflexion über spezifische Stimuli (Bild-, Video- oder Textvignetten) oder eigene Aktivität zu erfassen (u. a. Leonhardt et al. 2010; Weinberger & Seyfried, 2009; Wyss, 2013). In keinem Fall scheint ein Kompetenzverständnis zugrunde zu liegen, das sowohl Dispositionen als auch Skills und Performanz gleichermaßen in die Untersuchung einbezieht. Zudem konnten wir kaum Arbeiten identifizieren, die explizit für Reflexion notwendige Dispositionen beschreiben und diese jenseits von Selbstauskünften mit Tests zu erfassen versuchen.

#### Fragestellung

Wie lassen sich einzelne Bereiche eines integrierten Kompetenzmodells (Dispositionen, Skills, Performanz) vor dem Hintergrund geteilter theoretischer Überlegungen erfassen?

#### Methode

Um das Modell umfassend abzubilden, haben wir in einem ersten Schritt sowohl einen Zugang ausgehend von den Dispositionen als auch einen ausgehend von der Performanz gewählt: In den Dispositionen wurden mit bereits bestehenden und neu entwickelten standardisierten Instrumenten das reflexionsbezogene Wissen (eigene Entwicklung), die reflexionsbezogenen Orientierungen (u. a. Korthagen & Wubbels, 1995), die reflexionsbezogene Selbstwirksamkeit (eigene Entwicklung) und die reflexionsbezogenen Einstellungen (angelehnt an Niggli, 2004) mithilfe eines Tests erfasst. Der Test umfasst auch einen kleineren Teil, der auf situationsspezifischen Skills gerichtet ist, die mit Vignetten und offenen Fragen einen Reflexionsprozess initiieren, der schriftlich festgehalten wird. Das gesamte Instrument wurde mit ca. 190 Studierenden, überwiegend des zweiten Fachsemesters, erprobt. Die Performanz wurde im Rahmen von Nachgesprächen und Interviews



zum eigenen Unterricht im schulischen Praktikum am Beispiel Physik erfasst (N=20). Weil es sich in beiden Fällen um Pilotierungen von Zugängen vor dem Hintergrund eines geteilten Modells handelt, sind die Datenquellen noch nicht aufeinander abgestimmt. In zukünftigen Erhebungen mit überarbeiteten Instrumenten ist diese Abstimmung jedoch geplant.

Die quantitativen Teile des Tests wurden mittels Raschanalyse sowie explorativer Faktorenanalyse ausgewertet. Die qualitativen Teile des Tests wurden ebenso wie die transkribierten Nachgespräche und Interviews kodiert sowie teilweise in Cluster zusammengefasst und/oder mit weiteren sequenzanalytischen Verfahren analysiert. Es wurden zudem Korrelations- und Pfadanalysen genutzt, um Zusammenhänge unterschiedlicher Facetten innerhalb der Dispositionen sowie zwischen Dispositionen (als unabhängige Variable) und Skills (als abhängige Variable) zu ermitteln.

Ergebnisse

Korrelationsanalysen zeigen, dass vor allem die reflexionsbezogenen Orientierungen, die Selbstwirksamkeit und die Einstellungen hoch miteinander korrelieren ( $r$ =zwischen .385\*\* und .524\*\*). Mit dem reflexionsbezogenen Wissen korrelieren die reflexionsbezogenen Selbstwirksamkeitserwartung am stärksten ( $r$ =.257\*\*). In der Performanz zeigt sich, dass Studierende zwar einen analytischen Zugang zu ihrem Unterricht finden, darin jedoch eher selten ihr eigenes Verhalten und Denken hinterfragen.

Im Vortrag werden die Erträge, aber auch die Limitationen unserer Erkenntnisse für das Reflexionsmodell und die Erfassung von Reflexivität diskutiert.

### **Entwicklungsaufgabe „Perspektivenwechsel“. Wie kann ein für Reflexionsprozesse förderlicher Studierendenhabitus initiiert und unterstützt werden?**

**Mandy Geisler, Bianka Kaufmann**  
Justus-Liebig-Universität Gießen, Deutschland

Theoretischer Hintergrund

Universitäre Lehrerbildung steht vor der besonderen Herausforderung, dass Studierende durch ihre eigene Schulzeit starke Prägungen haben, die sich unmittelbar auf die Motive für die Berufswahl Lehramt und auch auf den Ausbildungsprozess selbst auswirken. Während der ca. 13000 erlebten und inkorporierten Schulstunden (Reinhardt, 2002) bauen junge Menschen bereits einen „Lehrer\*innenhabitus“ auf, der sich in einem reziproken Wechselspiel mit dem eigenen „Schüler\*innenhabitus“ manifestiert (Helsper, 2014). „Von daher ist die Irritation des alltäglichen Wissens von Lehrern und vor allem der impliziten Wissensbestände von Studierenden des Lehramts, im Rahmen der handlungsentlastenden Studienphase, eine der zentralen Herausforderungen für die Lehrerbildung [...]“ (Helsper, 2002, S. 78). Entsprechend dieser hochschuldidaktischen Herausforderung lässt sich die Entwicklungsaufgabe (Hericks, 2006) „Perspektivenwechsel“ für Lehramtsstudierende definieren. Gelingen kann dieser, indem eine reflexive Distanz zu den eigenen Deutungsmustern aufgebaut und bestehende subjektive Theorien (Groeben, Wahl, Schlee, & Schlee, 1988) re- und dekonstruiert werden.

Um den Perspektivenwechsel zu unterstützen, wurde ein umfassendes Beratungs- und Trainingsangebot konzipiert. Auf diese Weise wird eine intensive Auseinandersetzung mit der eigenen schulischen Lerngeschichte, mit familiär geprägten impliziten Deutungsmustern zu Schule und Unterricht sowie den bisherigen Erfahrungen im pädagogischen Handlungsfeld gefördert. Dies wird u. a. über eine forschungsbasierte und berufsbiographisch orientierte Lernwerkstatt organisiert, in dem der Perspektivenwechsel über den intensiven Austausch mit Peers initiiert, moderiert und in einem Lerntagebuch dokumentiert wird. In einer ersten Phase wurde und wird das Angebot mit Studierenden verschiedener Semester pilotiert, um Aussagen zu dessen Lage im Studienverlauf und zum Grad der Verbindlichkeit ableiten zu können.

Fragestellung

(1) Zu welchem Zeitpunkt innerhalb der Studienphase sollte die Lernwerkstatt stattfinden und unterscheidet sich die Bereitschaft zur Reflexion in Abhängigkeit von der Teilnahme am Angebot (freiwillig versus obligatorisch)? (2) Unter welchen Bedingungen lässt sich im Angebot ein fragend-forschender Studierendenhabitus initiieren und mit welche Widerständen muss dabei gerechnet werden?

Methode

Für die Beantwortung der Forschungsfragen kann auf Daten von 114 Studierenden des zweiten bis achten Semesters zurückgegriffen werden, die im Rahmen von Lehrveranstaltungen verpflichtend an dem Angebot teilgenommen haben. Im Wintersemester 2017/2018 wird das Angebot auf freiwilliger Basis sowie erneut verpflichtend in Seminaren mit Studierenden stattfinden, wodurch sowohl ein Vergleich verschiedener Studierendensemester als auch ein Vergleich der Teilnahmebedingungen möglich ist. Es wurden zwei unterschiedliche methodische Zugänge gewählt. (1) Mithilfe von Befragungen zu Beginn und zum Ende der Lernwerkstatt wurden u. a. Daten zur Teilnahmemotivation, zur Bereitschaft der Auseinandersetzung mit den eigenen subjektiven Theorien sowie am Ende zur Einschätzung der Qualität des Angebots, dessen erlebter subjektiver Nutzen und dessen wahrgenommener Anregung zur Reflexion erhoben. Darüber hinaus wurden die Studierenden zu ihrer Studienmotivation fürs Lehramt sowie der Zufriedenheit und Sicherheit mit der Studienwahl befragt, um insbesondere zu untersuchen, ob sich im freiwilligen Angebot bestimmte Profile vermehrt finden lassen (Nolle, 2013). (2) In einem weiteren Schritt werden biographisch-narrative Interviews mit Studierenden geführt (Beginn im Wintersemester 2017/2018), um individuelle Deutungsmuster zu rekonstruieren, sowie förderliche und einschränkende Einstellungen für Reflexionsprozesse nachzeichnen zu können.

Ergebnisse

Erste Auswertungen zeigen ein ambivalentes Bild bei der Bedeutungszuschreibung der Studierenden im Hinblick auf den Sinn und Zweck hinter der Auseinandersetzung mit ihren subjektiven Theorien. Auch wenn ca. 80% der Studierenden angeben, dass die Reflexion ihrer eigenen schulischen Erfahrungen wichtig ist, scheint das Einlassen in reflexiv angelegte Angebote eher gering, was sich vor allem in den offenen Antworten im Fragebogen zeigt, wie bspw. „...weil ich mir bewusst bin, warum ich auf Lehramt studiere und ich mich auch so schon sehr gut selbst reflektieren kann.“ oder „Zu viel Reflexion, die im ersten Semester fehl am Platz ist.“. Eine Diskussion zu den möglichen Ursachen und hochschuldidaktischen Konsequenzen hieraus soll durch den Beitrag angeregt werden.

### **Schülerrückmeldungen als Ausgangspunkt für Unterrichtsreflexion – Erkenntnisse einer empirischen Studie bei angehenden Sekundarlehrpersonen**

**Corinne Wyss, Simone Biaggi**  
Pädagogische Hochschule Zürich, Schweiz

## Theoretischer Hintergrund

Während angehende Lehrpersonen in der Ausbildung darin geschult werden ihren Schülerinnen und Schülern Rückmeldungen zu geben, wird umgekehrt Schülerfeedback in der Ausbildung und der Unterrichtspraxis noch wenig systematisch verwendet (Buhren, 2015; Kämpfe, 2009). Die Einschätzungen von Schülerinnen und Schülern sind bedeutsam, da sie eine Innensicht des Unterrichts darstellen und die Wirkung von Unterricht und Lehrerhandeln auf ihr Lernen gut beurteilen können (Hascher, Baillod, & Wehr, 2004; Clausen, 2002). Verschiedene Studien haben bislang außerdem gezeigt, dass die Rückmeldungen von Schülerinnen und Schülern zum Unterricht sehr differenziert ausfallen können (Buhren, 2015). Für die Lehrperson bietet Schülerfeedback eine wichtige Zusatzinformation zur eigenen Unterrichtsbeurteilung und kann damit sehr wertvoll für die Reflexion des Unterrichts, die Unterrichtsqualität sowie die Entwicklung der professionellen Kompetenz der Lehrperson sein (Hattie, 2009). Sofern Schülerfeedback sinnvoll und geplant eingesetzt wird, kann es sowohl zu einer qualitativen Verbesserung der Unterrichtsqualität beitragen als auch die Lernleistungen der Schülerinnen und Schüler positiv zu beeinflussen (Buhren, 2015).

Im Projekt „Schülerrückmeldungen zur Förderung der Unterrichtsreflexion“ (Selfreflex) werden angehende und erfahrene Praxislehrpersonen der Sekundarstufe I der Pädagogischen Hochschule Zürich angeleitet, Schülerfeedback gezielt für die Reflexion von Unterricht einzuholen und zu verwenden. Das Projekt wird in Anlehnung und Kooperation mit dem Projekt ScRiPS (Leitung Prof. Dr. Kerstin Göbel, Universität Duisburg-Essen) durchgeführt. Die zentrale Forschungsfrage lautet:

### Fragestellung

Inwiefern lässt sich die Reflexionsfähigkeit zum eigenen Unterricht von Studierenden im Quartalspraktikum und von erfahrenen Lehrpersonen durch eine systematische Reflexion von Schülerrückmeldungen in unterschiedlichen Reflexionssettings schulen und entsprechend abbilden?

### Methode

Konkret ist ein Einbezug von rund 100 - 120 Studierenden der Sekundarstufe I sowie 25 Praxislehrpersonen pro Durchführung geplant, wobei eine Durchführung im Frühlingsemester 2017 und eine zweite im Frühlingsemester 2018 vorgesehen ist. Im Rahmen des Quartalspraktikums befragen die Studierenden ihre Schülerinnen und Schüler zu ihrem Unterricht zu zwei Zeitpunkten und verarbeiten das Feedback im Anschluss in zwei unterschiedlichen Settings. Je die Hälfte der Studierenden reflektiert die Schülerfeedbacks: a) mit einer Studienkollegin bzw. einem -kollegen, oder b) individuell schriftlich im Lernprotokoll. Es wird einerseits geprüft, welches der zwei Settings die Reflexion am besten unterstützt. Andererseits werden die verschiedenen Reflexionssettings unter Berücksichtigung der motivationalen und volitionalen Voraussetzungen der Studierenden untersucht. Es wird hierzu eine schriftliche Pre- und Posttest Befragung der Studierenden durchgeführt (Neuber & Göbel, 2016) wie auch die schriftlichen und mündlichen Reflexionen eingesammelt und analysiert. Die Reflexionen der Praxislehrpersonen werden in gleicher Weise erhoben, um einen Vergleich zwischen Novizen (Studierende) und erfahrenen Lehrpersonen zu ermöglichen. Zudem werden auch die Schülerinnen und Schüler einbezogen, indem sie zu ihren Erfahrungen mit Unterrichtsfeedback und spürbaren Veränderungen im Unterricht befragt werden. Die Auswertung der quantitativen Daten erfolgt mittels der Analysesoftware SPSS, die schriftlichen und mündlichen Reflexionen werden mittels qualitativer Inhaltsanalyse (Mayring, 2015) in Bezug auf die inhaltliche Struktur und qualitative Merkmale analysiert.

### Ergebnisse

Derzeit liegen die Daten des ersten Erhebungszeitpunktes bei den Studierenden (N=117) im Frühlingsemester 2017 vor. Die Ergebnisse der Pre- und Posttest Befragung zeigen, dass die angehenden Lehrpersonen grundsätzlich positiv gegenüber der Reflexion von Unterricht eingestellt sind und die Rückmeldungen von Schülerinnen und Schülern als Reflexionsunterstützung wertschätzen. Vorgegebene Strukturierungshilfen für und die Verschriftlichung von Unterrichtsreflexion werden hingegen als weniger sinnvoll beurteilt. Aus den Daten werden auch einige geschlechtsspezifische Unterschiede ersichtlich: weibliche Studierende (N=64) bewerten die Reflexion insgesamt positiver und schätzen die Rückmeldungen von Lernenden und Kolleginnen und Kollegen mehr, als es die männlichen Studierenden (N=51) tun.

Auf Basis der theoretischen und empirischen Erkenntnisse werden in diesem Beitrag die Möglichkeiten und Herausforderungen der systematischen Reflexion von Schülerrückmeldungen erläutert und Implikationen für die Aus- und Weiterbildung diskutiert.

## **Weiterentwicklung des professionellen Handelns von Lehrkräften durch die Anregung von Reflexion im Rahmen von Fortbildungsangeboten**

**Ilka Benner, Stephanie Brombach, Katharina Hombach**  
Justus-Liebig-Universität Gießen, Deutschland

### Theoretischer Hintergrund

Die Professionalisierung von Lehrkräften endet nicht mit dem Erreichen des für den Beruf qualifizierenden Abschlusses, sondern erstreckt sich kontinuierlich über die gesamte Berufslaufbahn (Hippel, 2011, S. 249). Verdichtete Herausforderungen für die Professionalisierungsbestrebungen zeigen sich – aus Perspektive der Lehrkräfte – im Schulalltag: die Erwartungen an Unterricht, der Umgang mit einer heterogenen Schülerschaft sowie die (zuweilen widersprüchlichen) gesellschaftlichen und politischen Anforderungen. U. a. diese Aspekte erfordern eine Reflexion des eigenen professionellen Handelns (Tenorth, 2006, S. 585f.).

Die Universität, als Anbieterin wissenschaftlicher Fortbildungen und wissenschaftlich geleiteter Reflexion von Schule und Unterricht, ermöglicht Lehrkräften, die vor Ort in der Einzelschule oder der Unterrichtssituation vorgefundene Praxis wissenschaftlich zu reflektieren, systematisch zu bearbeiten und die eigene Praxis weiterzuentwickeln. In der Ausgestaltung dieses Angebotes kann die Universität ein breites Spektrum an Fort- und Weiterbildungsformaten zur Verfügung stellen (Leuphana Universität Lüneburg, 2012). Wir fokussieren im Rahmen unserer Studie drei Formate. Das erste Format ist eine klassische Fortbildung zur Berufs- und Studienorientierung (Angebot a), die insgesamt an vier Tagen über das Schuljahr verteilt stattfindet. Das zweite Angebot (Angebot b) verzahnt Schulentwicklung mit Unterrichtsforschung anhand ausgewählter Multiplikator\*innen. Dabei werden Forschungsfragen aus ihrer konkreten Schulpraxis generiert und universitär begleitete praxisorientierte Forschungsprojekte durchgeführt. Das dritte Angebot stellt ein organisiertes Vernetzungsangebot zwischen Universität und Schule in Bezug auf Schulpraktika (v. a. Mentor\*innen) dar (Angebot c). Dieses Angebot widmet sich der Aufgabe von Mentor\*innen, u. a. Praktikant\*innen mittels Unterrichtsanalysen und Reflexionsgesprächen auf ihrem Weg zur professionell handelnden Lehrkraft zu begleiten (KMK, 2011a; KMK, 2011b). Um dieser Aufgabe professionell nachkommen zu können, bedürfen Mentor\*innen selbst einer ausgeprägten Reflexivität (Kuhl, Solzbacher, & Zimmer, 2017).

Alle Angebotsformate haben – mit unterschiedlicher Herangehensweise – das Ziel, die Teilnehmer\*innen zur Reflexion ihrer eigenen Praxis anzuregen und Reflexionsprozesse zu initiieren. Angelehnt an dieses Ziel wird folgende Fragestellung bearbeitet.

#### Fragestellung

Bieten die unterschiedlichen Formate Anlässe zur Reflexion und welche Aspekte initiieren die Reflexionsprozesse der Teilnehmenden im besonderen Maße?

#### Methode

Zur Beantwortung der Fragestellung werden sowohl standardisierte Befragungen eingesetzt als auch leitfadengestützte Interviews mit den Teilnehmenden geführt. Im Fragebogen wird u. a. erfasst, inwiefern die Teilnehmenden das jeweilige Angebot als Anregung der Reflexion ihrer eigenen Praxis wahrgenommen haben (eigene Entwicklung in Anlehnung an Kember, Doris, & Leung (2000) und Schippers, Den Hartog, & Koopman (2007)). Die bereits für andere Angebote getestete Skala beinhaltet fünf Items (5-Stufiges Antwortformat, 1=nie bis 5=immer,  $\alpha=.802$ ). Mittels der Interviews sollen vertiefte Einblicke darin gewonnen werden, welche Aspekte, ggf. in Wechselwirkung mit der eigenen schulischen Praxis, besonders relevant für die Initiierung eigener Reflexionsprozesse sind. Die Auswertung der Interviews erfolgt mit der qualitativen Inhaltsanalyse nach Mayring (2015).

Beide Erhebungen werden über individualisierte Codes miteinander verbunden, sodass die quantitativen und qualitativen Befunde wechselseitig aufeinander bezogen werden können.

#### Ergebnisse

Für die Fortbildung (Angebot a) liegen erste Ergebnisse von acht Teilnehmenden vor. Erhoben wurde die Skala zu Reflexionsanlässen. Erste Analysen zeigen, dass – nach Angaben der Teilnehmenden – das Angebot ihnen im Durchschnitt häufig Anlässe zur Reflexion geboten hat ( $AM=3,9$ ;  $SD=0,245$ ). Die Durchführung der Interviews mit diesen Teilnehmenden dauert gegenwärtig noch an. Erste Eindrücke aus den Interviews zeigen, dass die Elemente des kollegialen Austauschs und die wissenschaftliche Rückbindung schulischer Erfahrungen Aspekte zur Initiierung von Reflexionsprozessen darstellen. Die Angebote für die Multiplikator\*innen (Angebot b) und die Mentor\*innen (Angebot c) werden im Wintersemester 2017/18 durchgeführt, bei denen zwischen 8 und 15 Personen teilnehmen werden.

Aufgrund der insgesamt eher schmalen Datenlage – eine typische Herausforderung für Datenerhebungen im Kontext von Lehrerfortbildungen – wird eine Ableitung generalisierbarer Prinzipien für Fortbildungen nicht möglich sein. Die Daten können jedoch Hinweise auf Möglichkeiten und Grenzen der Anregung reflexiver Prozesse geben, die kritisch diskutiert werden sollen.

## C14

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H117

### Wie gelingt Sprach- und Leseförderung? Befunde aus BiSS-Evaluationsprojekten im Schulbereich

Chair(s): **Jörg Jost** (Universität zu Köln), **Elmar Souvignier** (Universität Münster)

DiskutantIn(nen): **Michael Becker-Mrotzek** (Universität zu Köln)

In der bundesweiten Initiative „Bildung durch Sprache und Schrift (BiSS)“ werden die im Rahmen von Verbänden in den Bundesländern eingeführten Angebote zur Sprachförderung, Sprachdiagnostik und Leseförderung wissenschaftlich überprüft und weiterentwickelt. Dazu wird knapp die Hälfte der Verbände durch Evaluationsprojekte begleitet, in denen die Qualität der eingesetzten Materialien, die Umsetzung der Förderkonzepte im Unterricht und die Wirksamkeit der Maßnahmen untersucht werden.

Die in den BiSS-Evaluationsprojekten behandelten Fragen spannen einen Bogen von „Wie gut sind die Materialien?“ über „Wie gut gelingt die Implementation in den Unterricht?“ und „Welche Fähigkeiten haben die Lehrkräfte?“ bis hin zu „Wie fallen die Leistungsentwicklungen der Schülerinnen und Schüler in den BiSS-Verbänden aus?“. Für die Evaluationsforschung typische Wirksamkeitsprüfungen werden mithin um Bedingungsanalysen ergänzt, die Hinweise auf das Zustandekommen (bzw. das Ausbleiben) von Effekten geben können.

In dem Symposium sollen vier inhaltliche Beiträge sowie ein übergreifender Diskussionsbeitrag präsentiert werden. In dem Beitrag von Ehl et al. wird das in der schulischen Praxis häufig eingesetzte Verfahren der Sprachprofilanalyse nach Griebhaber auf seine Aussagekraft hin untersucht. Diese Arbeit illustriert exemplarisch die Notwendigkeit, vorliegende Materialien auf ihre Güte zu prüfen, um auf einer wissenschaftlichen Grundlage fundierte Empfehlungen für die schulische Praxis zu geben. Die Studie von Meudt et al. setzt bei der Implementation von Fördermaßnahmen an, indem durch Unterrichtsbeobachtungen und Lehrerbefragungen die tatsächliche Nutzung der Materialien untersucht wird. Indem zusätzlich potenzielle Bedingungen für die Nutzung der Materialien (Einstellungen, Motivation, kollegiale Kooperation) erfragt wurden, lassen sich Ansatzpunkte für eine Optimierung der Implementation benennen. Rank et al. setzen sich in ihrer Studie mit dem fachdidaktischen Wissen und den Einstellungen der Lehrkräfte als Voraussetzungen für eine wirksame Sprachförderung auseinander. Auf Seiten der Kinder werden (bildungs)sprachliche und fachliche Fähigkeiten erhoben und es wird analysiert, in welchem Zusammenhang Merkmale der Lehrkräfte mit den Leistungsentwicklungen der Schülerinnen und Schüler stehen. In der Studie von Schlitter et al. wird schließlich in einem Vergleich mit Kontrollschulen ohne spezifische Leseförderprogramme geprüft, ob die Lesekompetenzen und die sprachlichen Fähigkeiten von Kindern aus BiSS-Schulen am Ende der Grundschulzeit überlegen sind. Zudem wird der Frage nachgegangen, ob leistungsschwächere Kinder in besonderem Maße von den BiSS-Konzepten profitieren. Der abschließende Diskussionsbeitrag (Becker-Mrotzek) ordnet die Beiträge in den Kontext der BiSS-Initiative ein, beleuchtet sie kritisch und thematisiert noch offene Fragen bzw. Desiderate.

Insgesamt machen die Beiträge des Symposiums deutlich, dass es im Hinblick auf die Auswahl an Materialien, deren Umsetzung im Unterricht und die lehrerseitigen Voraussetzungen für Maßnahmen zur Sprach- und Leseförderung Optimierungsbedarf gibt. Gleichzeitig wird deutlich, dass enorme Unterschiede zwischen einzelnen Lehrkräften vorliegen, was sich letztlich in den Effekten auf Seiten der Kinder spiegelt. Das Symposium versucht, Hinweise auf Bedingungen einer gelingenden Sprach- und Leseförderung zu geben.

*Beiträge des Symposiums*

### Eignung der Profilanalyse nach Griebhaber zur Bestimmung von Wortschatzkompetenzen

**Birgit Ehl, Michèle Paul, Gunnar Bruns, Michael Grosche**

Bergische Universität Wuppertal

Die Profilanalyse nach Griebhaber (2013) ist ein häufig verwendetes diagnostisches Instrument zur Bestimmung der grammatischen Fähigkeiten von Grundschul\*innen. Die Profilanalyse bewertet die grammatische Komplexität von Äußerungen eines Kindes in sogenannten Erwerbsstufen. In einer ersten Studie evaluierten wir gemeinsam mit den Lehrkräften im Sinne einer symbiotischen Implementationsstrategie einige Testgütekriterien und optimierten die Profilanalyse daraufhin (Ehl et al. in Begutachtung). Jedoch wird weiterhin kontrovers diskutiert, ob die diagnostizierten grammatischen Erwerbsstufen der Profilanalyse weitere diagnostische Informationen liefern können. Griebhaber (2013) konstatiert, dass die Erwerbsstufe „gleichzeitig Hinweise auf weitere Bereiche des Sprachstands gibt, ohne dass diese Bereiche direkt gemessen werden“ (Griebhaber, 2013, S. 6). Beispielsweise kontrastiert er in einer Tabelle (S. 7) die verschiedenen Erwerbsstufen mit Meilensteinen der Wortschatzentwicklung und legt damit nahe, dass aus den Erwerbsstufen auch diagnostische Informationen über den Wortschatz abgeleitet werden können. „Zumindest indirekt lässt die Profilanalyse durchaus Rückschlüsse auf den schon erworbenen Wortschatz zu. Unter der Voraussetzung, dass sich die erworbenen sprachlichen Strukturen auf der Basis eines entsprechend erworbenen Lexikons entwickeln, setzt z.B. das Erreichen der Stufe 3 auch einen der Stufe 3 entsprechend umfangreichen Wortschatz voraus, ohne dass dazu genauere Auskünfte gegeben werden“ (Griebhaber, 2005, S. 16). Die in unserem symbiotischen Projekt beteiligten Lehrkräfte und Verbundkoordinator\*innen würden diese zusätzlichen diagnostischen Informationen gerne für die Planung der Sprachförderung nutzen baten uns jedoch zuvor um eine entsprechende Evaluation. Es ist zu erwarten, dass grammatische Fähigkeiten und Wortschatz im Sinne einer allgemeinen Sprachkompetenz zwar zusammenhängen, aber dennoch grundsätzlich voneinander dissozierbar sind. Um den Wortschatzumfang repräsentativ erfassen zu können, müssen zahlreiche Kriterien beachtet werden (Glück 2011, S. 11). Daraus wäre schlusszufolgern, dass allein aus den Erwerbsstufen der Profilanalyse keine (oder nur sehr wenige) diagnostischen Informationen über den Wortschatz ableitbar sind. Daher erwarten wir im Sinne der Konstruktvalidität der Profilanalyse eine zwar positive, jedoch eher geringe Korrelation zwischen den Erwerbsstufen der Profilanalyse und anerkannten Wortschatztests.

556 ein- und mehrsprachige Grundschul\*innen (52,2% weiblich) der Klassen 1 bis 4 (Alter MW 7,98 Jahre,  $s = 1,27$ ) aus fünf Grundschulen nahmen an der Erhebung teil. 161 erwarben Deutsch als Muttersprache, 235 wuchsen simultan mehrsprachig (Kontaktdauer zu Deutsch MW = 7,25 Jahre  $s = 1,62$ ) und 160 sukzessiv mehrsprachig (Kontaktdauer zu Deutsch MW = 4,14 Jahre  $s = 1,51$ ) auf. Sie durchliefen die Wortschatztests AWST-R (Kiese-Himmel 2005) und WWT 6-10 (Glück 2011) sowie die Profilanalyse nach Griebhaber. Als Erzählimpuls verwendeten wir ein selbstentwickeltes Bilderbuch.

Die Korrelation zwischen der Erwerbsstufe aus der Profilanalyse und dem Wortschatz in der deutschen Sprache beträgt in der Gesamtgruppe  $r = .37$  ( $p < .001$ ). Die Korrelationen in den verschiedenen Sprachgruppen sind mit Werten von  $r = .29$ ,  $.33$  und  $.43$  (alle  $p < .001$ ) sehr ähnlich und unterscheiden sich nicht signifikant voneinander ( $V[2] = 2.22$ ,  $p > .05$ ).

Zwar hängen grammatische Fähigkeiten und Wortschatz im Sinne einer allgemeinen Sprachkompetenz signifikant zusammen, jedoch spricht der lediglich mittelhohe Zusammenhang deutlich für die grundsätzliche Dissoziierbarkeit beider Fähigkeiten. Wortschatz und grammatische Fähigkeiten teilen nur 14 % identische Informationen ( $r^2$ ). Eine Ableitung diagnostischer Informationen über den Wortschatz aus den Ergebnissen der Profilanalyse ist demnach (und im Widerspruch zu den theoretischen Annahmen Grießhabers) bei ein- und mehrsprachigen Grundschulkindern nicht zu empfehlen. Daher sollten Wortschatzkompetenzen mit spezifisch dafür vorgesehenen Testverfahren ermittelt werden. Einschränkend ließe sich in weiteren Studien überprüfen, ob die Anzahl an unterschiedlichen Wörtern je Wortart in der Profilanalyse präzisere Informationen für eine Diagnostik des Wortschatzes liefern könnte. Eine solche qualitative Auswertung wäre jedoch sehr aufwändig und damit für Lehrkräfte voraussichtlich nicht praktikabel.

### **Bedingungen erfolgreicher Implementation von Leseförderkonzepten in der Sekundarstufe I**

**Sarah-Ines Meudt<sup>1</sup>, Anke Schmitz<sup>2</sup>, Fabiana Karstens<sup>2</sup>, Nina Zeuch<sup>1</sup>, Elmar Souvignier<sup>1</sup>, Jörg Jost<sup>2</sup>**

<sup>1</sup>Universität Münster, <sup>2</sup>Universität zu Köln

#### Theoretischer Hintergrund

Das Verständnis darüber, wie innovative Unterrichtskonzepte – beispielsweise wissenschaftlich gesicherte Erkenntnisse zur Leseförderung – Eingang in das Bildungssystem finden, ist eher unzureichend (Kretlow & Helf, 2013; Philipp & Souvignier, 2016). Im Hinblick auf die konkrete Nutzung von innovativen Unterrichtskonzepten zeigt sich in Implementationsstudien, dass sich Lehrkräfte sowohl in der Umsetzungshäufigkeit als auch in der Genauigkeit der Umsetzung voneinander unterscheiden (McGraw et al., 2000). Bislang liegen jedoch nur vereinzelt empirische Befunde darüber vor, welche Einflussfaktoren für den nachhaltigen Einsatz (= Nutzung) evidenzbasierter Innovationen bedeutsam sind. Auf der individuellen Ebene der Lehrkraft ist anzunehmen, dass sich neben der Einstellung und Motivation auch die Selbstwirksamkeit und Akzeptanz gegenüber den Materialien förderlich auswirken (vgl. Lakin & Shannon, 2015). Darüber hinaus gibt es auf der schulischen Ebene Hinweise darauf, dass der Implementationserfolg von der Verfügbarkeit kooperativer Netzwerkstrukturen abhängt (z.B. Höke, 2013).

Vor diesem Hintergrund untersucht das BiSS-Projekt EILe (Evaluation der Implementation von Konzepten zur Leseförderung in der Sekundarstufe I) die Implementation von Leseförderkonzepten. Ziel ist es, am Beispiel von drei BiSS-Verbänden Bedingungen für eine gelingende Implementation zu ermitteln. Der Fokus dieses Beitrags liegt zunächst auf der Analyse des Nutzungsverhaltens (Unterrichtsbeobachtung und Befragung). In einem zweiten Schritt werden Einflussfaktoren auf die Nutzung von Innovationen untersucht.

#### Fragestellungen

Anhand von Unterrichtsbeobachtungen im Deutsch- und Fachunterricht soll zunächst der Einsatz des Lesefördermaterials analysiert werden, das den Lehrkräften zur Verfügung gestellt wurde. Ergänzt wird diese Bestandsaufnahme zur Nutzung durch eine Lehrerbefragung, in der die Nutzungshäufigkeit erhoben wird. Darüber hinausgehend untersucht die vorliegende Studie den Einfluss individueller Faktoren (Einstellung, Motivation, Selbstwirksamkeit, Akzeptanz) und schulischer Faktoren (Kooperation) auf die Nutzungshäufigkeit.

#### Methode

Im Forschungsprojekt EILe werden über zwei Schuljahre 20 Schulen bei der Implementation von Konzepten zur Leseförderung begleitet. Das formative Evaluationskonzept basiert auf objektiven Datenerhebungen, die Lehrerbefragungen und Unterrichtsbeobachtungen einschließen. Auf einer siebenstufigen Likert-Skala wurden  $N = 100$  Lehrkräfte zu individuellen und schulischen Einflussfaktoren sowie der Nutzungshäufigkeit befragt (etablierte Skalen, Cronbach's Alpha zwischen .74-.93). Die Beobachtung der Materialnutzung im Deutsch- und Fachunterricht (Biologie und Mathematik) erfolgte durch zwei geschulte Rater in 62 Unterrichtsstunden (31 Klassen).

#### Ergebnisse

Die Beobachtungen zeigen, dass die vorgesehenen Fördermaterialien über die Verbände hinweg in nur 9,7% der Stunden (6 von 62 Stunden) genutzt werden. Häufiger (in 13 beobachteten Stunden) werden die verbundspezifischen Fördermaterialien adaptiert. Noch öfter (18 Stunden) wurde anderes Material ohne erkennbaren Bezug zu dem Fördermaterial der Verbände genutzt. In der Mehrheit der Stunden (40,3%) wurde auf jegliches Material zur Anleitung des strategischen Lesens verzichtet. Unter Berücksichtigung des Unterrichtsfaches zeigt sich, dass das verbundspezifische Fördermaterial im Deutschunterricht am häufigsten, in den Fächern Mathematik und Biologie dagegen in keiner beobachteten Stunde genutzt wird.

Die Angaben der Lehrkräfte zur Nutzungshäufigkeit der Fördermaterialien weisen eine deutliche Variation auf einem eher niedrigen Niveau aus ( $M = 2,77$ ,  $SD = 1,23$ ). Die Zusammenhänge zwischen den in die Untersuchung einbezogenen Prädiktoren und der Nutzungshäufigkeit sind hoch ( $r = .52 - .68$ ). Die vier Prädiktoren auf der Ebene der individuellen Lehrkraft (Einstellung, Motivation, Selbstwirksamkeit, Akzeptanz) klären gemeinsam 47 Prozent der Varianz in der Nutzungshäufigkeit auf, während für sich genommen auf die kollegiale Kooperation (Schulebene) 31 Prozent entfallen. Eine Moderatoranalyse legt nahe, dass der Zusammenhang zwischen Selbstwirksamkeit und Nutzungshäufigkeit durch die Akzeptanz vermittelt wird.

Zusammenfassend weist die Bestandsaufnahme zum Nutzungsverhalten und zur Nutzungshäufigkeit darauf hin, dass die Implementation von Förderkonzepten in den BiSS-Verbänden noch deutlichen Optimierungsbedarf aufweist. Gleichzeitig wird deutlich, dass – theoretisch plausible – Bedingungen auf individueller und auf schulischer Ebene Unterschiede im Nutzungsverhalten in einem hohen Maße erklären können. Diese Prädiktoren stellen mithin geeignete Ansatzpunkte für eine Optimierung von Implementationsprozessen dar.

### **Evaluation im Primarbereich – Erfahrungen mit Sprachförderung im Projekt „Eva-Prim“**

**Astrid Rank<sup>1</sup>, Isabell Deml<sup>1</sup>, Anja Wildemann<sup>2</sup>, Anita Schilcher<sup>1</sup>, Stefan Krauss<sup>1</sup>, Linda Lenske<sup>2</sup>, Alexandra Merkert<sup>2</sup>, Magdalena Schulte<sup>1</sup>, Karin Binder<sup>1</sup>, Lena Bien-Miller<sup>2</sup>**

<sup>1</sup>Universität Regensburg, <sup>2</sup>Universität Koblenz-Landau

#### Theoretischer Hintergrund

Es ist seit einigen Jahren Konsens, dass jedem Fach die Aufgabe zukommt, nicht nur fachliches Wissen zu vermitteln, sondern auch sprachliche Fähigkeiten zu fördern (Becker-Mrotzek et al. 2013, 7). Zur Verbesserung der Sprachkompetenz der Schülerinnen und Schüler ist bei den Lehrkräften anzusetzen, da sie durch ihre fachspezifischen Kompetenzen die Schülerleistung beeinflussen (Kunter et al. 2011; Helmke 2012; Krauss et al. 2017). Neben dem fachdidaktischen Wissen (hier um Sprachförderung und Mathematik) sind auch die Einstellungen und Haltungen der Lehrkräfte entscheidend (Kunter et al.

2011). So ist z.B. davon auszugehen, dass Sprachförderung nur dann im Unterricht stattfindet, wenn Lehrkräfte ihr einen Platz im Fachunterricht zugestehen.

#### Fragestellung

Bei der Bund-Länder-Initiative BiSS sollen in Verbänden die in Schulen stattfindenden Methoden zur Sprachförderung konzeptualisiert und daraus Professionalisierungsmöglichkeiten für Lehrkräfte abgeleitet werden. Diesen Professionalisierungsprozess untersucht das Projekt „Eva-Prim“ in vier Verbänden der BiSS-Initiative am Beispiel des Mathematikunterrichts, der sich in besonderer Weise zur Identifizierung von Sprachfördermaßnahmen anbietet. Denn die Inhalte des Fachs Mathematik erfordern eine Sprache, die geprägt ist von eigener Fachterminologie, spezifischen Satzstrukturen und Textsorten (Maier/ Schweiger 1999). Im Fokus der Evaluation steht erstens die Frage, welche Kompetenzen in den Bereichen Sprachförderung und Mathematik die Lehrkräfte mitbringen, zweitens, welche Fähigkeiten bei den Kindern im (bildungs)sprachlichen und fachlichen Bereich vorliegen und schließlich, wie sich diese in Abhängigkeit von den Lehrerkompetenzen entwickeln.

#### Methode

Die Evaluation umfasst drei Messzeitpunkte (Schuljahre 2015/2016, 2016/2017, 2017/2018), so dass die Schülerinnen und Schüler (N = 809) von der zweiten bis zur vierten Jahrgangsstufe begleitet werden und auch die Lehrkräfte (N = 83) an mehr als einem Messzeitpunkt zu ihrem Sprach- und Fachwissen getestet werden können. Zur Feststellung der bildungssprachlichen Fähigkeiten wurde der Hörverstehenstest „BiSpra“<sup>1</sup> und ab dem zweiten Messzeitpunkt auch das in einem Promotionsprojekt entwickelte Testverfahren zur sprachlichen Ausdrucksfähigkeit in Mathematik „SAMT“ eingesetzt. Die Mathematikkompetenz wurde mit dem „DEMAT“ ermittelt, Kenntnisse des rezeptiven Wortschatzes anhand einer Gruppenversion des „PPVT“ und die Lesefertigkeit durch Einsatz des „SLS 2-9“; auch die Deutsch- und Mathematiknote wurde erfragt. Kinderspezifische Merkmale, wie kognitive Fähigkeiten, erhoben durch den „CFT 1-R“, und soziodemographischer Hintergrund, wurden außerdem berücksichtigt. Zur Ermittlung des fach-didaktischen Wissens der Lehrkräfte wurden zwei eigenentwickelte Testverfahren („FaWi-S“ und „FaWi-M“) eingesetzt. Zudem wird bei einer Teilstichprobe die Umsetzung im Unterricht videografiert und auf sprachförderliche Interaktion hin analysiert.

#### Ergebnisse

Auf Schülerebene zeigen die Ergebnisse des ersten Messzeitpunktes in allen vier Verbänden unauffällige Werte bei den kognitiven Fähigkeiten. Bei der fachlichen Leistung in Mathematik, der Lesefertigkeit und dem rezeptiven Wortschatz reichen die Ergebnisse von ausbaufähig zu durchschnittlich bis hin zu überdurchschnittlich in den verschiedenen Verbänden. Am zweiten Messzeitpunkt fallen die Leistungen bei der Lesefertigkeit etwas ab, bleiben bei den bildungs-sprachlichen Fähigkeiten dagegen im mittleren Bereich.

Die Lehrkräfte erreichen durchschnittliche Ergebnisse beim fachdidaktischen Wissen in den Bereichen Sprachförderung<sup>2</sup> (M = 19,46, SD = 3,80) und Mathematik<sup>3</sup> (M = 4,54, SD = 1,54). Beim Sprachförderwissen gibt es enorme Unterschiede. Im Bereich der Einstellungen erzielen alle vier Verbände hohe Werte bezüglich einer positiven Grundhaltung gegenüber der Sprachförderung im Mathematikunterricht. Skeptischer hingegen zeigten sich die Lehrkräfte der vier Verbände, was den Einbezug von Herkunftssprachen in den Unterricht betrifft. Die Gestaltung des Unterrichts ist sehr lehrerdominiert, was sich in einem geringen Maß komplexerer Äußerungen der Schülerinnen und Schüler widerspiegelt.

1 Hörverstehenstest der Humboldt-Universität zu Berlin, IQB, Otto-Friedrich-Universität Bamberg; entstanden im Rahmen eines BMBF-Projektes von Stanat & Weinert: „Bildungssprachliche Anforderungen: Anforderungen, Sprachverarbeitung und Diagnostik“.

2 Es konnten 34 Punkte erreicht werden.

3 Es konnten 8 Punkte erreicht werden.

### **Sprachliche Kompetenzen im vierten Grundschuljahr: Erste Ergebnisse zu Effekten der Initiative BiSS**

**Theresa Schlitter<sup>1</sup>, Franziska Schwabe<sup>1</sup>, Jennifer Iglar<sup>1</sup>, Annika Ohle<sup>1</sup>, Annika Teerling<sup>2</sup>, Regine Asseburg<sup>2</sup>, Nele McElvany<sup>1</sup>, Olaf Köller<sup>2</sup>**

<sup>1</sup>TU Dortmund, <sup>2</sup>IPN Kiel

Lesekompetenz gilt als zentrale Voraussetzung für erfolgreiche Bildungskarrieren und ist darüber hinaus bedeutsam für gesellschaftliche Teilhabe (Becker-Mrotzek, Hasselhorn, Roth & Stanat, 2016). Eine zentrale Einflussgröße der Lesekompetenz ist unter anderem auch der Wortschatz (Quinn, Wagner, Petscher & Lopez, 2015). Wiederholt wurde für Deutschland festgestellt, dass ein substantieller Teil der Kinder am Ende der Grundschulzeit noch nicht über hinreichende Lesekompetenz für einen erfolgreichen Schulbesuch verfügen (z.B. Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Zudem wird angenommen, dass geringere Wortschatzkompetenzen die Entwicklung der Lesefähigkeit behindern können (Tannenbaum, Torgesen & Wagner, 2006). Ein Ansatzpunkt ist daher systematische Lese- und Sprachförderung. Studien weisen auf positive Effekte von Wortschatz- und Leseförderung auf Leseverständnis hin, insbesondere auch bei leistungsschwächeren Schülerinnen und Schülern (z.B. Quinn, Wagner, Petscher & Lopez, 2015). Das Projekt Evaluation von Konzepten und Maßnahmen der fachübergreifenden Leseförderung im Primarbereich (BiSS-EvalLesen) zielt im Rahmen der Bund-Länder Initiative „Bildung durch Sprache und Schrift“ (BiSS) in diesem Kontext unter anderem auf die Effektivitätsprüfung der in den beteiligten Schulen eingesetzten Leseförderprogrammen ab.

Der vorliegende Beitrag untersucht mit einem quasiexperimentellen Design, ob Lernende, die BiSS-Schulen besuchen, im Gegensatz zu Kindern von Kontrollschulen ohne spezielle Leseförderprogramme am Ende der Grundschulzeit unter Kontrolle der Ausgangsbedingungen zu Beginn der vierten Klassenstufe über (a) höhere Lesekompetenzen und (b) einen sichereren Umgang mit Wortschatz verfügen. Weiterhin gilt es zu klären, ob leistungsschwache Kinder differenziell von einem BiSS-Konzept profitieren. Betrachtet werden N = 604 Kinder, die Klassen an Schulen besuchen, die ein durch BiSS-initiiertes Sprachförderkonzept verfolgen, sowie N = 400 Kinder aus Klassen an Schulen, die kein strukturiertes Sprachförderkonzept verfolgen. Die Kinder wurden im Herbst 2016 (Messzeitpunkt [MZP] I) und im Sommer 2017 (MZP II) befragt und getestet. Sie bearbeiteten eine adaptierte Version des Lesetests aus der Lernausgangslagenuntersuchung (LAU; Behörde für Schule und Berufsbildung, 2011; vgl. Tabelle 1) sowie die Skala Wortanalogien des Kognitiven Fähigkeitstests (KFT; Heller & Perleth, 2000; vgl. Tabelle 1). Beide Instrumente wiesen sowohl in der BiSS- als auch in der Kontrollgruppe zu MZP I akzeptable bis zufriedenstellende Reliabilitäten auf ( $\alpha > .67$ ). Die Daten des MZP II werden im Oktober 2017 vorliegen. Die Gruppen unterschieden sich nicht nach Anteil Mädchen (BiSS: 47.4%; Kontrolle: 50.0%;  $\chi^2 = .62$ ,  $df = 1$ ,  $p = .43$ ). Allerdings bestanden Unterschiede in Bezug auf das Alter (BiSS: MAlter = 10.88 Jahre, SDAlter = 0.52; Kontrolle: MAlter = 10.80 Jahre, SDAlter = 0.50;  $t[927] = -2.08$ ,  $p < .05$ ), den Anteil Kinder mit einer (auch) von Deutsch verschiedenen Familiensprache (BiSS: 39.6%;

Kontrolle: 32.3%;  $\chi^2 = 4.79$ ,  $df = 1$ ,  $p < .05$ ) und tendenziell dem sozialen Hintergrund (Anteil Kinder mit mehr als 100 Büchern im Haushalt; BiSS: 47.9%; Kontrolle: 41.6%;  $\chi^2 = 3.15$ ,  $df = 1$ ,  $p = .08$ ). Nach Kontrolle der Unterschiede in Alter, Familiensprache und sozialem Hintergrund zeigte sich zwischen Kindern der BiSS- und der Kontrollgruppe kein Unterschied zu Beginn der Klassenstufe 4 in der Lesekompetenz ( $F[1, 767] = 0.80$ ,  $p = .76$ ) oder dem kompetenten Umgang mit Wortschatz ( $F[1, 767] = 7.68$ ,  $p = .41$ ).

Basierend auf dieser Ausgangslage wird der Beitrag zur Beantwortung der ersten Forschungsfrage kovarianzanalytische Verfahren anwenden, um unter Berücksichtigung der Ausgangsleistungen die sprachlichen Kompetenzen am Ende der Grundschulzeit in beiden Gruppen vergleichend zu untersuchen. Hierbei werden bedeutsame Merkmale auf individueller und Klassenebene als Kovariaten einfließen. Zur Beantwortung der zweiten Forschungsfrage nach differenziellen Effekten für schwache Schülerinnen und Schüler wird ein Interaktionsterm aus Gruppe und Fähigkeit neben dem Haupteffekt der Gruppe aufgenommen.

Die Ergebnisse stellen bedeutsame Erkenntnisse für die schulische Praxis und Forschung in dem höchst relevanten Feld sprachlicher Bildung bereit und lassen darüber hinaus eine Diskussion zur breiten Implementation von bottom-up entwickelten schulischen Förderkonzepten zu.

## C15

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H116

### Wie nutzen Schülerinnen und Schüler angebotene Lerngelegenheiten? Innovative Ansätze zur Erfassung von Wahrnehmungs- und Verarbeitungsprozessen

*Chair(s):* **Katrin Rakoczy** (Deutsches Institut für Internationale Pädagogische Forschung), **Marc Kleinknecht** (Leuphana Universität Lüneburg), **Svenja Vieluf** (Deutsches Institut für Internationale Pädagogische Forschung), **Anna-Katharina Praetorius** (Universität Zürich), **Pietsch Marcus** (Leuphana Universität Lüneburg)

*DiskutantIn(nen):* **Jasmin Decristan** (Bergische Universität Wuppertal)

Um das erfolgreiche Zusammenspiel von Lehr- und Lernprozessen differenziert zu erschließen, haben sich in der Professions- und Unterrichtsforschung Angebots-Nutzungsmodelle bewährt. Diese Modelle unterscheiden u.a. zwischen Angebotsstrukturen bzw. Lerngelegenheiten, Nutzungsprozessen und Lernergebnissen (siehe z.B. Helmke, 2011). Die konzeptionelle Herausforderung besteht darin, geeignete vermittelnde Prozesse der individuellen Wahrnehmung und Verarbeitung in das Rahmenmodell zu integrieren, mit deren Hilfe die Wirkung einzelner Aspekte von Unterrichtsqualität auf den Lernerfolg erklärbar wird. Ihr wird entsprechend der multikriterialen Konzeption von Lernerfolg durch Einbezug vielfältiger psychologischer Theorien zum individuellen Lernen begegnet. Die empirische Erfassung des Zusammenspiels von Wahrnehmungs- und Verarbeitungsprozessen der Lernenden beschränkt sich bisher jedoch häufig darauf, solche Prozesse retrospektiv, d.h. nach Ablauf der Unterrichtsstunde oder sogar der Unterrichtseinheit, per Fragebogen zu erfassen. Damit geht häufig eine reduzierte Aussagekraft der Befunde einher.

Wir möchten in dem Symposium Beiträge vorstellen, die Wahrnehmungs- und Verarbeitungsprozesse von Lernenden auf innovative Weise untersuchen. Während die ersten beiden Studien videobasierte Beobachtungen nutzen, um das Wechselspiel zwischen Angebot und Nutzung zu erfassen, setzen Studie 3 und 4 Eye-tracking-Verfahren bzw. Pupillometrie ein, um Prozesse der Aufgabenbearbeitung sichtbar zu machen.

Im ersten Beitrag werden die Ergebnisse zur didaktischen Kommunikation in tutoriellen Situationen in Mathematik vorgestellt. Dazu wurden in videografierten tutoriellen Situationen typische Dialogmuster identifiziert. Die Ergebnisse der Interaktionsanalysen zeigen, dass sich die Abfolge von Lehrer- und Schüleräußerungen in erfolgreichen und nicht erfolgreichen tutoriellen Dialogen unterscheidet.

Die Autorinnen und Autoren des zweiten Beitrags präsentieren Ergebnisse einer Studie, die untersucht, wie sich verschiedene Maßnahmen der Klassenführung auf die individuelle Lernzeitnutzung (On- und Off-Task-Verhalten) von Lernenden mit und ohne sonderpädagogischen Förderbedarf auswirken. Hierzu wurden Aufnahmen von mehreren Minikameras in fünf Inklusionsklassen analysiert. Die Ergebnisse zeigen u.a., dass Allgegenwärtigkeit bei gleichzeitigem Erhalten des Unterrichtsflusses mit häufigerem On-Task-Verhalten assoziiert ist.

Im dritten Beitrag stellt der Autor Ergebnisse aus einer Serie von Laborexperimenten vor, in denen der Frage nachgegangen wird, inwiefern Veränderungen der Pupillengröße als Prozessmaß zur Operationalisierung von kognitiver Aktivität einsetzbar sind. Die Ergebnisse zeigen u.a., dass die durch Pupillometrie gemessene Überraschungsreaktion in positivem Zusammenhang mit der Lernleistung steht.

Im vierten Beitrag werden Befunde aus einer Laborstudie präsentiert, in der ein Experten-Novizen-Vergleich bei der Bearbeitung von Aufgaben zur Erfassung von Bildkompetenz durchgeführt wurde. Blickbewegungsdaten zeigen Unterschiede hinsichtlich der Anzahl und Dauer der Fixationen, der Blicksprünge und der Sakkaden und zeigen damit, welche Vergleichsparameter bedeutsam sind, um lösungsrelevante Vorteile von Experten aufzuschlüsseln und didaktisch nutzbar zu machen.

Die DiskutantIn der Beiträge erörtert abschließend den Erkenntnisgewinn der Studien und das Potenzial der methodischen Ansätze zur Erforschung von Nutzungsprozessen im Unterricht.

*Beiträge des Symposiums*

### Was macht den Unterschied zwischen lernwirksamen und nicht-lernwirksamen Lehrer-Schüler-Dialog aus?

**Anke Wischgoll<sup>1</sup>, Christine Pauli<sup>1</sup>, Kurt Reusser<sup>2</sup>**

<sup>1</sup>Universität Freiburg, Zentrum für Lehrerinnen- und Lehrerbildung, Schweiz, <sup>2</sup>Universität Zürich, Institut für Erziehungswissenschaft, Schweiz

Der Lehrer-Schüler-Dialog ist ein zentrales Instrument des Unterrichts. Um Schülerinnen und Schüler beim Lernen effizient unterstützen zu können, muss seine Wirksamkeit verstanden werden. Tutorielle Dialoge bieten sich zur Analyse des Lehrer-Schüler-Dialogs an, da sie die interaktiven Prozesse eines Dialogs weitestgehend ohne äussere Einflüsse abbilden.

Schüler in ihrem Verstehensprozess zu unterstützen, erfordert den augenblicklichen Stand ihres Verstehens zu kennen. Um diesen in einem Lehrer-Schüler-Dialog feststellen zu können, braucht es die Beteiligung des Schülers/der Schülerin. Wie setzen Lehrpersonen Fragen, Erklärungen und Beispiele in einer Abfolge ein, um das Verstehen zu überprüfen und ein Verstehensmoment zu erarbeiten? Die Forschung weist bereits einzelne Strategien aus, mit denen Lehrpersonen ihre Schüler bei der Erarbeitung von Verstehensmomenten in einem Lehrer-Schüler-Dialog unterstützen können. Durch geschlossene und offene Fragen, Erklärungen und Beispiele können Lehrpersonen den Lehrer-Schüler-Dialog steuern (Mercer & Howe, 2012). Diese Strategien haben unterschiedliche Funktionen und sind je nach Verwendung förderlich für den Verstehensprozess: Geschlossene Fragen dienen der Lehrperson festzustellen, dass etwas gelernt wurde. Offene Fragen fordern den Schüler/die Schülerin auf, Erkenntnisse selbst zu formulieren, und lassen dadurch erkennen, welche Verstehenslücken noch bestehen. An diesen Verstehenslücken ansetzend kann die Lehrperson den Schüler durch Nachfragen unterstützen, den Lerngegenstand besser zu durchdringen (Wolf, Crosson & Resnick 2006). Durch Erklärungen kann die Lehrperson den Fokus für die Bearbeitung des Lerngegenstands auf relevante Aspekte lenken oder auch Hauptaussagen paraphrasieren (Roscoe & Chi 2008). Anhand von Beispielen kann die Lehrperson den Lerngegenstand vereinfacht abbilden und so dem Schüler/der Schülerin den Zugang zum Lerngegenstand erleichtern (Duit, Roth, Komorek & Wilbers, 2001). Der Schüler/die Schülerin trägt zum Dialog Antworten und Selbsterklärungen bei (Roscoe & Chi, 2008).

Ziel der Studie war, anhand von Dialogmustern Unterschiede zwischen erfolgreichen und nicht erfolgreichen tutoriellen Dialogen in der Dialogführung zu bestimmen. Konkret wurden folgende Fragestellungen analysiert: (1) Sind typische Abfolgen



von Lehreräußerungen und Schüleräußerungen nach dem Auftreten eines Fehlers beobachtbar? Unterscheiden sich diese Abfolgen in erfolgreichen und nicht erfolgreichen tutoriellen Dialogen? (2) Werden die Dialoge nach einer korrekten Antwort des Schülers/der Schülerin in erfolgreichen und nicht erfolgreichen tutoriellen Dialoge unterschiedlich weitergeführt?

Der aus der schweizerisch-deutschen Videostudie (Klieme, Pauli & Reusser, 2009) stammende Datensatz bestand aus 26 videografierten tutoriellen Lehrer-Schüler-Dialogen von 26 Lehrpersonen aus der Schweiz und aus Deutschland (8./9. Schuljahr). Eine vorgegebene algebraische Textaufgabe wurde bearbeitet. Anschließend lösten die Schülerinnen und Schüler selbstständig eine ähnliche Aufgabe. Anhand des Lösungserfolgs bei dieser Transfer-Aufgabe fand eine Unterteilung in „erfolgreiche“ und „nicht erfolgreiche“ tutorielle Dialoge statt. Der tutorielle Dialog wurde mittels einer Aufgabenanalyse in Segmente, die jeweils die Behandlung der einzelnen Verstehelemente der Aufgabe umfassten, unterteilt. Innerhalb dieser Segmente wurden die Äußerungen der Lehrperson und des Schülers/der Schülerin als Frage, Antwort, Erklärung und Beispiel kategorisiert. Aus dieser Analyse wurden typische Dialogmuster herausgearbeitet und erfolgreichen bzw. nicht erfolgreichen tutoriellen Dialogen zugeordnet.

Die Ergebnisse zeigen, dass sich die Abfolge von Lehrer- und Schüleräußerungen in erfolgreichen und nicht erfolgreichen tutoriellen Dialogen unterscheidet. Dies gilt sowohl für den Umgang mit dem auftretenden Fehler wie auch für den Umgang mit der richtigen Antwort des Schülers/der Schülerin. Konkret weisen die Ergebnisse u.a. aus, dass (1) in erfolgreichen tutoriellen Dialogen Fehler durch eine wiederholte Abfolge von Lehrerfragen und Schülerantworten aufgearbeitet werden, wohingegen in nicht erfolgreichen tutoriellen Dialogen eine wiederholte Abfolge von Lehrerfragen und SchülerSelbsterklärungen zu beobachten ist, und dass (2) in erfolgreichen tutoriellen Dialogen erarbeitete Verstehelemente durch Erklärungen konsolidiert werden, wohingegen in nicht erfolgreichen tutoriellen Dialogen eine Konsolidierung nicht zu beobachten ist.

### **Welche Klassenführungsmerkmale führen zur Lernzeitnutzung? Eine Multi-Video-Studie zum Angebot-Nutzungszusammenhang in Inklusionsklassen**

**Marc Kleinknecht, Kira Weber, Christopher Prilop**

Leuphana Universität Lüneburg

Von einer effektiven Klassenführung auf Angebotsseite des Unterrichts wird erwartet, dass sie auf einer Nutzungsseite zu einer höheren aktiven Lernzeit (time on task) führt. Die aktive Lernzeit von Schülerinnen und Schüler gilt wiederum als ein zentraler Prädiktor für deren Lernerfolg (u.a. Hattie, 2009). Bereits die Studie von Kounin (1976) konnte zeigen, dass eine wirksame Klassenführung sehr unterschiedliche Maßnahmen, wie den Umgang mit Störungen, das Monitoring unterschiedlicher Interaktionen, den strukturierten Ablauf des Unterrichts oder die Aktivierung der Lernenden umfasst.

Aktuell dominieren in der Unterrichtsqualitätsforschung methodische Zugänge, die Maßnahmen der Klassenführung durch hoch-inferente videobasierte Analysen des Unterrichtsangebots und deren Lernwirkung durch Befragung der Schülerinnen und Schüler erfassen (u.a. Lipowsky & Faust, 2013). Insbesondere Befragungen der Lernenden zur Klassenführung haben dabei den Nachteil, dass lediglich die wahrgenommene Störungsarmut bezogen auf die gesamte Lerneinheit und weniger einzelne Maßnahmen in ihrer individuellen oder klassenbezogenen Wirksamkeit in den Blick genommen werden. Zu vermuten wäre etwa, dass sich die hohe Bedeutung der Allgegenwärtigkeit für den Lernerfolg (Hattie, 2009, S. 102) auch in einer hohen Lernzeitnutzung vieler Schülerinnen und Schüler niederschlägt. Andere Maßnahmen, könnten insbesondere bei heterogenen Lerngruppen zu unterschiedlicher Lernzeitnutzung führen.

Unsere explorative Studie zielt auf eine mikrologische Erfassung von Angebot-Nutzungszusammenhängen und untersucht, wie sich verschiedene Maßnahmen der Klassenführung (Allgegenwärtigkeit/Monitoring von Störungen, Überlappung, Reibungslosigkeit und Schwung, Rechenschaftsprinzip, Gruppenmobilisierung) auf die individuelle Lernzeitnutzung (On- und Off-Task-Verhalten) von Schülerinnen und Schüler mit und ohne sonderpädagogischen Förderbedarf auswirken. Hierzu wurden Aufnahmen von jeweils fünf bis sieben Minikameras (Go-Pros) in fünf Inklusionsklassen der Sekundarstufe I herangezogen, um für Plenums- und Arbeitsphasen Zusammenhänge zwischen Maßnahmen der Lehrkraft und Reaktionen der einzelnen Lernenden aufzuzeigen. Für die Kodierung wurden Kategoriensysteme basierend auf den Arbeiten von Kounin (1976) und Helmke & Renkl (1992) in einem niedrig-inferenten Time-Sampling-Verfahren eingesetzt. Die Daten wurden mit dem Videoanalyseprogramm Interact ausgewertet.

Die ersten Analysen zeigen, dass sich die Allgegenwärtigkeit im Sinne eines rechtzeitigen und korrekten Reagierens auf Störungen bei gleichzeitigem Erhalten des Unterrichtsflusses positiv auf die Lernzeitnutzung auswirkt. Sowohl bei den förderbedürftigen Schülerinnen und Schüler, als auch den meisten anderen Schülerinnen und Schülern lässt sich ein On-Task-Verhalten beobachten. Dagegen wirkt es sich negativ auf die Lernzeitnutzung (hohes Off-Task-Verhalten) aus, wenn die Lehrkraft den Unterricht nicht reibungslos und schwungvoll gestaltet. Dies führt bei Schülerinnen und Schüler mit Förderbedarf und den meisten anderen Lernenden in der Klasse zu einem Off-Task-Verhalten.

Im Beitrag werden weitere Analyse zum Zusammenhang von Klassenführungsmaßnahmen und Lernzeitnutzung auf Individual- und Klassenebene präsentiert und diskutiert.

### **Veränderungen in der Pupillengröße als Maß für kognitive Aktivität?**

**Garvin Brod**

Deutsches Institut für Internationale Pädagogische Forschung

Kognitiv aktivierender Unterricht beschreibt einen fachlich anspruchsvollen Unterricht, der Schülerinnen und Schüler zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand anregt (Klieme & Rakoczy, 2008). Hierfür müssen den Lernenden Aufgaben gestellt werden, die an ihr Vorwissen anknüpfen und sie zu einer vertiefenden Beschäftigung mit den Lerninhalten anregen. Kognitive Aktivierung wird oft über das Anregungspotential von Lehrerverhaltensweisen oder Aufgaben gemessen (z. B. Baumert et al., 2010; Lipowsky, 2006). Nach dem Angebots-Nutzungs-Modell (z.B. Helmke, 2009) ist es jedoch ebenso zentral, inwiefern das Potential von den Lernenden genutzt wird, d.h. wie kognitiv aktiv sie sind. Es stellt sich daher die Frage, wie kognitive Aktivität bei Lernenden objektiv gemessen werden kann. In diesem Beitrag soll der Frage nachgegangen werden, ob sich die Erfassung von Veränderungen in der Pupillengröße mittels Eye-Tracking – kurz: Pupillometrie – als direktes Maß für die individuelle kognitive Aktivität von Lernenden eignet. Hierzu werden zunächst die zwei Haupteinsatzgebiete der Pupillometrie in der Lernforschung vorgestellt: zur Messung von cognitive load sowie lernrelevanter Emotionen wie Überraschung oder Neugier. Es wurde eine Untersuchungsreihe durchgeführt, in der das Potential von lerner-generierten Vorhersagen zur Aktivierung von Vorwissen und zur Induzierung von kognitiven Konflikten erforscht wurde. Vorhersagen generieren setzt voraus, dass Lernende auf ihr Vorwissen zurückgreifen und es mit der zu lernenden Information verknüpfen (Fielding, Anderson, & Pearson, 1990). Stellt sich eine Vorhersage als falsch heraus, führt dies den Lernenden bestehende Wissenslücken / Fehlkonzepte vor Augen. Die erlebte Diskrepanz sollte

bei den Lernenden zunächst Überraschung und im Anschluss einen kognitiven Konflikt hervorrufen, der aufgelöst werden muss (D'Mello, Lehman, Pekrun, & Graesser, 2014; Mandler, 1990). Das Generieren von Vorhersagen sollte daher zu einer hohen kognitiven Aktivität der Lernenden führen und einen positiven Effekt auf deren Lernleistung haben, der über die Effekte erfolgreicher Vorwissensaktivierung hinausgeht. Diese Annahmen wurden in einer Serie von Laborexperimenten mit Studierenden (n = 36) überprüft, in der Wissen über europäische Geographie vermittelt wurde. Verglichen wurde die Lernleistung zwischen einer Bedingung, in der die Lernenden zunächst eine Vorhersage treffen mussten und dann die korrekte Information erhielten, und einer Bedingung, in der sie sofort die korrekte Information erhielten und anschließend ihre Erwartung angeben sollten. Als Prozessmaß wurde Pupillometrie eingesetzt, wobei insbesondere die Pupillenreaktion auf unerwartete Informationen mit der Pupillenreaktion auf erwartete Informationen verglichen wurde ( $\Delta > 0$  = Überraschungsreaktion). Lernerfolg wurde gemessen als Prätest-Posttest-Veränderung im Wissen über die Einwohnerzahlen der Länder und mittels eines Hierarchietests erfasst. Es zeigte sich, dass das Generieren einer Vorhersage zu einer verbesserten Lernleistung führte und dass unerwartete Informationen nur in der Vorhersagebedingung Überraschung hervorriefen. Letzteres spricht für eine erfolgreiche Aktivierung von Vorwissen. Überdies stand die durch Pupillometrie gemessene Überraschungsreaktion in positivem Zusammenhang mit der Lernleistung. Diese Ergebnisse liefern Evidenz für das Potential von Pupillometrie zur Messung individueller Lernaktivität und zur Spezifizierung erfolgreicher Lernprozesse. Zum Abschluss des Vortrags werden Herausforderungen bei der Durchführung von Pupillometrie diskutiert.

### **Wie werden Aufgaben zur Bildkompetenz gelöst? Ein Experten-Novizen Vergleich von Blickbewegungsdaten**

**Miles Tallon<sup>1</sup>, Katrin Rakoczy<sup>2</sup>, Ulrich Frick<sup>3</sup>, Susanne Weiß<sup>2</sup>**

<sup>1</sup>HSD University of Applied Sciences, <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, <sup>3</sup>University of Applied Sciences

Die Spezifikation eines Kompetenzmodells für die im Kunstunterricht essentielle Zielvariable „Bildkompetenz“ stellt ein wichtiges Forschungsdesiderat dar, wofür der Europäische Referenzrahmen zu „Visual Literacy“ wichtige Ansatzpunkte bietet (Wagner & Schönau, 2016). Um Aufschluss darüber zu erhalten, welche Wahrnehmungs- und Verarbeitungsprozesse mit einer hohen vs. niedrigen Bildkompetenz verbunden sind, eignen sich Daten über Blickbewegungen während des Aufgabenlösungsprozesses. Bisherige Studien zeigen, dass künstlerische Experten im Vergleich zu Novizen objektiv messbare Unterschiede bei der Bearbeitung visueller Aufgaben aufweisen (z.B. Engelbrecht, Betz, Klein & Rosenberg, 2010; Vogt & Magnussen, 2011). Experten weisen gegenüber Novizen u.a. folgende lösungsrelevante Vorteile auf: Kürzere Dauer der Fixationen, längere Blicksprünge, mehr Fixationen auf aufgabenrelevante Bereiche und schnellere Erfassung von AOs (areas of interest), die zur Lösung beitragen (Gegenfurtner, Lehtinen, Säljö, 2011). Es zeigen sich allerdings unterschiedliche Ausprägungen je nach Expertendomäne, Aufgabenart und Aufgabenmaterial. So finden Vogt und Magnussen (2011) bei Künstlern weniger Fixationen mit längerer Dauer als bei Kunstlaien, wenn neue Bilder betrachtet werden. In der vorliegenden Studie wird untersucht, welche Unterschiede zwischen Experten und Novizen beim Lösen von visuellen Aufgaben zur Erfassung von Bildkompetenz relevant sind.

Dazu wurden aus einem Tablet-basierten Set visueller Aufgaben, welche im Rahmen des Projekts „Bildkompetenz in der kulturellen Bildung – Was ist und wie fördert man Bildkompetenz?“ (Frick, Rakoczy, Tallon & Weiß, 2017) entwickelt wurden, 23 Items ausgewählt, die Bilderkennung, Sortierung nach semantischen und/oder semiotischen Merkmalen sowie Suchaufgaben umfassen. Im Sinne eines initialen Proof of Principle wurde die Validität dieser Items durch einen Experten-Novizen-Vergleich überprüft. Dazu wurden die Blickbewegungen eines Kunstexperten (studierter Kunstpädagoge, männlich, berufstätig) und einer Novizin (Psychologin, Kunstlaie, weiblich, berufstätig) beim Lösen der Aufgaben vermessen.

Die Kunstlaie benötigte je nach Aufgabe 25% – 74% mehr Fixationen für die Bearbeitung der Aufgaben. Der Experte wies über alle drei Aufgabenblöcke hinweg signifikant weniger Fixationen pro Sekunde auf (Visual Intake Frequency [count/s]),  $t(40.41)=7.34$ ,  $p<.001$ ,  $M_{\text{Experte}}= 2.46$  (SD= 0.32),  $M_{\text{Novize}}= 3.07$  (SD= 0.24), die Fixationen dauerten beim Experten insgesamt länger (Visual Intake Duration Average [ms]),  $t(40.02)=-8.20$ ,  $p<.001$ , und  $M_{\text{Experte}}= 347.51$  (SD= 38.55),  $M_{\text{Novize}}= 266.23$  (SD= 27.81), der Experte wies kürzere Blicksprünge auf (Saccade Duration Average [ms]),  $t(29.44)=7.21$ ,  $p<.001$ ,  $M_{\text{Experte}}=42.48$  (SD= 4.37),  $M_{\text{Novize}}= 58.51$  (SD= 10.46) und benötigte insgesamt weniger Sakkaden pro Sekunde als die Novizin (Saccade Frequency [count/s]),  $t(40.06)=7.36$ ,  $p<.001$ ,  $M_{\text{Experte}}= 2.29$  (SD= 0.36) und  $M_{\text{Novize}}= 2.97$  (SD= 0.26). Die signifikanten Unterschiede ( $\alpha= 0.05$ ) zur Fixationsanzahl und Länge decken sich mit Untersuchungen von Vogt und Magnussen (2007).

Beim Sortieren der Bilder nach Abstraktionsgrad, benötigte der Experte rund 65% weniger Fixationen als die Laiin (82 Fixationen im Vergleich zu 259 Fixationen). Beim Suchen von Details fixierte der Experte insgesamt 177-mal, die Laiin 267-mal. Über vier Teilaufgaben hinweg, benötigte der Experte nur knapp die Hälfte der vergleichbaren Fixationen der untersuchten Novizin, d.h er war schneller und effektiver in der Aufgabenbewältigung.

Durch Vergleich eines Experten und einer Novizin konnten bedeutsame Vergleichsparameter identifiziert werden, um die lösungsrelevanten Vorteile von Experten aufzuschlüsseln und beim Bearbeiten von visuellen Aufgaben zur Erfassung von Bildkompetenz empirisch messbar zu machen. Können diese Ergebnisse auf gruppenstatistischer Ebene bestätigt werden, stützt dies die inhaltliche Validität der visuellen Aufgaben zur Messung von Bildkompetenz. Durch die Analysen können Wahrnehmungs- und Verarbeitungsprozesse sichtbar gemacht werden, indem Lösungsstrategien für Aufgaben herausgearbeitet werden. Anwendungsmöglichkeiten für die Unterrichtssituation wurden von Gegenfurtner et. al. (2011) aufgrund einer Metaanalyse zu Experten-Novizen Vergleichen bereits vorgedacht: Lernenden könnte das Lösungsverhalten von Experten in Form von kurzen Videos der Augensakkaden von z.B. Kunstlehrkräften vorgeführt werden, analog zu „thinking aloud“ nur mit Videos von Blickbewegungen. Darüber hinaus hilft das Sichtbarmachen von Lösungsstrategien mittels Blickbewegungsmessung bei Lernenden Lehrkräften, die individuelle Aufgabenbearbeitung ihrer Schülerinnen und Schüler besser zu verstehen und sie gezielt zu unterstützen.

## C16

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H115

### Akkulturationsorientierungen, kulturelle Identität und schulische Integration

Chair(s): **Aileen Edele** (Humboldt-Universität zu Berlin), **Georg Lorenz** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB))

DiskutantIn(nen): **Petra Stanat** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB))

Akkulturationsorientierungen von Heranwachsenden mit Zuwanderungshintergrund gelten als wichtige Determinante ihrer schulischen Integration (z.B. Motti-Stefanidi et al., 2012). Insbesondere sehen prominente Akkulturationsansätze in der Akkulturationsorientierung ein wichtiges Bindeglied zwischen den Akkulturationsbedingungen, denen Zuwanderer und ihre Nachkommen ausgesetzt sind, und dem Ausgang des Akkulturationsprozesses (Arends-Toth & Van de Vijver, 2006). Meist wird die Akkulturationsorientierung einer Person anhand von zwei Dimensionen beschrieben, nämlich ihrer Orientierung am Herkunftskontext und ihrer Orientierung am Aufnahmekontext (Berry, 1997). Ein verwandter Ansatz nimmt die Beurteilung von outgroups (positiv vs. negativ) und die Einstellungen gegenüber Gruppenunterschieden (sollen betont vs. überwunden werden) in den Blick, um Akkulturationsorientierungen von Personen mit und ohne Zuwanderungshintergrund zu charakterisieren (Hahn et al., 2011). Die kulturelle Identität, also wie sehr sich eine Person mit der eigenethnischen Gruppe (ethnische Identität) bzw. mit der Majoritätsgruppe (nationale Identität) identifiziert, stellt einen zentralen Aspekt der Akkulturationsorientierung dar (Liebkind, 2006; Phinney et al., 2001; Ward, 2001). Für den Bildungserfolg und die soziale Einbettung in die Majoritätsgruppe werden die Orientierung am Aufnahmekontext bzw. die nationale Identität als besonders relevant erachtet, für das psychologische Wohlbefinden und die soziale Einbindung in die ethnische Gemeinschaft hingegen eher die Orientierung am Herkunftskontext bzw. die ethnische Identität (Berry, 1997; Schotte, Stanat & Edele, 2017). Bisher wurden in der Akkulturationsforschung vor allem klassische Aufnahmekontexte wie die USA betrachtet und es ist unklar, inwiefern sich existierende Annahmen und Befunde auf den Aufnahmekontext Deutschland übertragen lassen. Zudem hat sich die bisherige Forschung auf bestimmte Akkulturationsoutcomes konzentriert, etwa schulbezogenen Einstellungen, während anderen, etwa den Bildungsaspirationen oder Freundschaftsbeziehungen, weniger Beachtung geschenkt wurde.

Vor diesem Hintergrund bündelt unser Symposium Beiträge, die sich mit Akkulturationsorientierungen in Deutschland befassen, ihre Bedingungen analysieren und ihre Zusammenhänge mit Aspekten der (schulischen) Integration untersuchen. Aufgrund der Vielschichtigkeit des Themas und des hohen Forschungsinteresses an diesem Forschungsfeld in verschiedenen Fachdisziplinen planen wir ein Symposium mit vier Einzelbeiträgen und zwei Diskussionen. Während sich die ersten beiden Beiträge mit der Prävalenz und den Bedingungen von Akkulturationsorientierungen beschäftigen, fokussieren der dritte und vierte Beitrag auf die Effekte der kulturellen Identität als zentralem Aspekt von Akkulturationsorientierungen.

Der Beitrag von Jansen und Kollegen untersucht, welche Akkulturationsorientierungen bei Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund vorliegen und inwiefern sie in unterschiedlichen schulischen Kontexten bzw. nach unterschiedlichen individuellen Merkmalen variieren.

Moffit und Kolleginnen untersuchen Akkulturationsorientierungen vor dem Hintergrund verschiedener Formen der kulturellen Sozialisation im Klassenraum. Dabei stehen mögliche Zusammenhänge mit der kulturellen Identität und der intrinsischen schulischen Motivation sowie den Einstellungen zu Diversität von Heranwachsenden im Vordergrund des Interesses.

Schotte und Salikutluk analysieren Bildungsaspirationen von Schülerinnen und Schülern mit Zuwanderungshintergrund und wie sich diese mit dem Verbleib im Aufnahmeland verändern. Als mögliche Ursache für Veränderungen betrachten die Autorinnen die kulturelle Identität.

Leszczynsky und Pink widmen sich schließlich der Frage, ob die Stärke der ethnischen Identifikation bei Heranwachsenden mit Zuwanderungshintergrund das Ausmaß ethnischer Segregation in Schulklassen beeinflusst. Dazu untersuchen sie Freundschaftsnetzwerke.

Das Symposium bringt Beiträge aus der Psychologie (Jansen et al.; Moffitt et al.), den Erziehungswissenschaften (Schotte) und der Soziologie (Leszczynsky & Pink) zusammen und wird aus bildungswissenschaftlicher (Stanat) und entwicklungspsychologischer (Juang) Perspektive diskutiert.

#### Beiträge des Symposiums

### Kulturelle Wertorientierungen von Schülerinnen und Schülern in Berlin: Faktorstruktur, Ausprägung und Bedingungsfaktoren

**Malte Jansen<sup>1</sup>, Jürgen Baumert<sup>2</sup>, Michael Becker<sup>3</sup>, Marko Neumann<sup>3</sup>, Olaf Köller<sup>4</sup>**

<sup>1</sup>Institut zur Qualitätsentwicklung im Bildungswesen (IQB), <sup>2</sup>Max-Planck-Institut für Bildungsforschung, <sup>3</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), <sup>4</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN)

Die von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund gehaltenen normativen Vorstellungen über das Zusammenleben in multiethnischen Gesellschaften werden als Akkulturations- oder Diversitätsnormen bezeichnet und gelten als zentrale Bedingungsfaktoren gelingender schulischer Integrationsprozesse (Hahn et al., 2010; Hachfeld et al., 2011; Schachner et al., 2015). Nach Hahn et al. (2010) werden in Bezug auf das Verhältnis von Mehrheitsgesellschaft und Zuwanderergruppen vier solcher normativen Orientierungen unterschieden. Diese unterscheiden sich bezüglich der Wahrnehmung von Gruppenunterschieden (wichtig vs. unwichtig) und bezüglich der Wertschätzung von Zuwanderergruppen und Mehrheitsgesellschaft (Präferenz vs. Gleichwertigkeit). Die Norm Egalitarismus ignoriert die Bedeutung von Unterschieden bei gleicher Wertschätzung der Gruppen. Im Multikulturalismus verbinden sich die Wahrnehmung und die Akzeptanz von Gruppenunterschieden mit der Wertschätzung der Gruppe der jeweilig Anderen. Assimilation zielt auf die Überwindung von Unterschieden bei einer Präferenz für die Kultur der Mehrheitsgesellschaft. Separierung schließlich betont die Salienz von Grenzziehungen bei Präferenz für die jeweilig eigene Gruppe, sei es die der Mehrheits- oder Herkunftsgesellschaft.

In Anbetracht ihrer angenommenen regulativen Funktion für interethnische Kontaktprozesse einerseits und für Akkulturationsstrategien von Jugendlichen mit Zuwanderungshintergrund andererseits (Baumert, 2017; Schachner et al., 2015, 2016), ist es überraschend, dass zur Prävalenz dieser Werthaltungen im deutschen Schulkontext bisher kaum Evidenz vorliegt.

Als Beitrag zur Schließung dieser Forschungslücke sollen zentrale Ergebnisse der BERLIN-Studie (Neumann et al., 2017) zu Akkulturationsnormen vorgestellt und um zusätzliche Analysen ergänzt werden. Dabei wird zunächst beschrieben, in welchem Maße Schülerinnen und Schüler an Berliner Sekundarschulen den vier Akkulturationsnormen zustimmen. In diesem Rahmen

wird auch auf Aspekte der Messung und faktorenanalytischen Trennbarkeit der Normen eingegangen. Auf dieser Basis werden Einflussfaktoren auf diese normativen Wertvorstellungen auf Individualebene (z. B. Zuwanderungshintergrund, Herkunftsgruppe, Generationenstatus, sozialer Hintergrund, kognitive Grundfähigkeiten) und auf Schulebene (Schulform, Komposition) analysiert.

Die im Beitrag vorgestellten Analysen basieren auf einer im Frühjahr 2014 durchgeführten Erhebungswelle der BERLIN-Studie, an der eine repräsentative Stichprobe von 2109 Neuntklässlerinnen und Neuntklässlern an Berliner Schulen teilnahm, wobei etwa die Hälfte der Schülerinnen und Schüler einen Zuwanderungshintergrund aufwies. Die Erfassung der Akkulturationsnormen erfolgte mit einem Instrument von Hachfeld et al. (2011, 2015), das je drei Items zu den vier erfassten Normen (Egalitarismus, Separierung, Assimilation, Multikulturalismus) beinhaltet. Zunächst wurde die Faktorstruktur der Akkulturationsnormen mit konfirmatorischen Faktorenanalysen geprüft. Anschließend wurden mit hierarchischen Regressionsmodellen individuelle und schulische Prädiktoren für die Zustimmung zu den vier Normen untersucht.

Die theoretisch postulierte vierfaktorielle Struktur ließ sich in einem konfirmatorischen Faktormodell gut an die Daten anpassen (CFI = .971, TLI = .960, RMSEA = 0.048, SRMR = 0.035). Die höchste Korrelation zeigte sich zwischen der Zustimmung zu den Normen Egalitarismus und Multikulturalismus ( $r = 0.73$ ), die die Vorstellung der Gleichwertigkeit verschiedener ethnischer Gruppen gemein haben. Dies waren auch die beiden Normen, die insgesamt die meiste Zustimmung bei den Jugendlichen erfuhren. Der Separierungsnorm wurde dagegen deutlich seltener zugestimmt. Dieser Befund zeigte sich in allen analysierten Subgruppen, also unabhängig von Migrations- und Generationenstatus, Herkunftsgruppe und Schulform.

Bei der multivariaten Analyse von Einflussfaktoren auf diese Normen zeigten sich eher geringe Unterschiede zwischen Jugendlichen mit und ohne Zuwanderungshintergrund. Allerdings stimmen Jugendliche an Gymnasien der Egalitarismus- und der Multikulturalisnorm eher zu als Jugendliche an Integrierten Sekundarschulen. Für Assimilation und Separierung zeigte sich ein gegenläufiges Ergebnis. Die Schulformunterschiede konnten für Egalitarismus vollständig, für die anderen Normen aber nur teilweise durch Unterschiede zwischen den Schülerschaften in Bezug auf Individualmerkmale wie elterlichen Bildungshintergrund und kognitive Fähigkeiten erklärt werden. In weiteren Analysen, die auf der Konferenz vorgestellt werden, sollen im nächsten Schritt auch verschiedene Kompositionscharakteristika (z. B. Anteil der Jugendlichen mit Zuwanderungshintergrund, Diversität der Herkunftsgruppen) ins Analysemodell einbezogen werden.

### **Kulturelle Sozialisation im Klassenraum: Zusammenhänge mit ethnischer und nationaler Identität, Motivation und Einstellungen zu Diversität bei Jugendlichen**

**Ursula Moffitt<sup>1</sup>, Linda P. Juang<sup>1</sup>, Maja K. Schachner<sup>1</sup>, Moin Syed<sup>2</sup>**

<sup>1</sup>Universität Potsdam, <sup>2</sup>University of Minnesota

Ein Literaturüberblick forderte kürzlich mehr Forschung zu ethnischer und kultureller Sozialisation im Schulkontext, wo Jugendliche einen Großteil ihrer Zeit verbringen (Aldana & Byrd, 2015). In der Schule werden Schüler\*innen sowohl durch das Kurrikulum als auch die Normen und Einstellungen der Pädagog\*innen formell und informell sozialisiert. In Deutschland dominiert unter Lehrerenden eine „farbenblinde“ bzw. egalitäre Perspektive. Das heißt Lehrkräfte tendieren dazu, kulturelle und ethnische Vielfalt zu ignorieren, während individuelle Merkmale, die unabhängig von der ethnischen und kulturellen Herkunft sind, für sie die Grundlage eines gleichberechtigten Miteinanders bilden. Mit dieser Position könnte eine fehlende Bereitschaft für die Betonung kultureller Vielfalt einhergehen (Hachfeld, Hahn, Schroeder, Anders, & Kunter, 2015). Allerdings ist bisher wenig darüber bekannt, ob und wie sich mehr oder weniger kultursensible Formen der Sozialisation im Klassenraum auf die Identität und das schulische Engagement und die Einstellungen von Jugendlichen zu Diversität auswirken. Positive Einstellungen zu Diversität unterscheiden sich insofern von allgemeiner Offenheit, als sie mit der Annahme positiver Auswirkungen von kultureller Vielfalt auf die Gesellschaft verknüpft sind (Kauff, Issmer, & Nau, 2013).

Die vorliegende Studie untersucht unterschiedliche Formen der Sozialisation im Klassenraum und möglich Zusammenhänge mit ethnischer und nationaler Identität, intrinsischer schulischer Motivation und Einstellungen zu Diversität bei Schülerinnen und Schülern. Wir gehen davon aus, dass es mindestens zwei Gruppen von Schüler\*innen gibt: solche, die ein niedriges Maß und solche, die ein moderates Maß an kultureller Sozialisation wahrnehmen. Wir vermuten weiterhin, dass Jugendliche, die höhere Werte von kultureller Sozialisation angeben, eine höhere ethnische und nationale Identität sowie eine höhere intrinsische Motivation aufweisen. Außerdem sollten sie positivere Einstellungen zu Diversität haben.

Die Stichprobe beinhaltet Schüler\*innen mit Migrationshintergrund der 9. Klassenstufe (N=976, 49% weiblich, 30,2% muslimisch) aus 17 kulturell heterogenen Schulen in Berlin. Alle Teilnehmer\*innen berichteten in Fragebögen Selbsteinschätzungen in Bezug auf kulturelle Sozialisation im Klassenraum, ethnische und nationale Identität, intrinsische Motivation und Einstellungen zu Diversität. Zur Erfassung der ethnischen Identität wurde eine für den deutschen Kontext adaptierte Variante des MIBI-teen (Scottham, Sellers, & Nguyễn, 2008) verwendet. Es wurde eine latente Profilanalyse in MPlus durchgeführt, die drei Subskalen einschloss, um die kulturelle Vielfalt im Klassenraum zu erfassen: Wissenserwerb über die Herkunftskultur der Schüler\*innen, kritisches Bewusstsein fördern und Diskussionen in der Klasse zu kultureller Vielfalt in Deutschland. Es wurden Effekte auf Klassenebene berücksichtigt, indem Intra-Klassen-Korrelationen in den Analysen kontrolliert wurden.

Basierend auf den Model-Fit Tests wurden drei Profile von kultureller Sozialisation identifiziert: (1) wenig kulturelle Sozialisation im Klassenraum; (2) wenig bis mittelmäßiger Wissenserwerb über die Herkunftskultur und kritisches Bewusstsein, mittelmäßige Diskussion über kulturelle Vielfalt in Deutschland; (3) mittelmäßig bis hoher Wissenserwerb über die Herkunftskultur und kritisches Bewusstsein, hohe Diskussion über kulturelle Vielfalt in Deutschland. ANOVAs wurden verwendet, um Unterschiede in den Profilen hinsichtlich ethnischer und nationaler Identität, intrinsischer Motivation und Einstellungen zu Diversität zu testen.

Im Allgemeinen zeigten sich signifikante Unterschiede zwischen jeweils zwei, teilweise auch zwischen allen drei Profilen: Diejenigen, die viel kulturelle Sozialisation im Klassenraum wahrnahmen berichteten eine höhere intrinsische Motivation, eine stärkere nationale Identität und positivere Einstellungen zu Diversität. Zusammengefasst deuten die Ergebnisse darauf hin, dass die Anerkennung von und der Wissenserwerb über die Herkunftskulturen der Schüler\*innen und Vielfalt in Deutschland im Rahmen der Sozialisation im Klassenraum deutlich mit Identität sowie akademischen und interpersonellen Outcomes zusammenhängt.

### **Welche Rolle spielt die kulturelle Identität für die Bildungsaspirationen Heranwachsender mit Zuwanderungshintergrund?**

**Kristin Schotte, Zerrin Salikutluk**

Berliner Institut für empirische Integrations- und Migrationsforschung (BIM) der Humboldt-Universität zu Berlin

Studien haben wiederholt gezeigt, dass einige Migrantengruppen durchschnittlich höhere Bildungsaspirationen aufweisen als ihre autochthonen Peers (Salikutluk, 2016; Brinbaum & Cebolla-Boado, 2007). Diese hohen Bildungsambitionen stehen in starkem Kontrast zum tatsächlichen Bildungserfolg dieser Gruppen (Olczyk et al., 2016). Als Erklärung für dieses paradoxe Verhältnis von Bildungsaspirationen und –erfolgen wird häufig das Phänomen des Immigrant Optimism aufgeführt (Kao & Tienda 1995), wonach zugewanderte Eltern eine hohe Aufstiegsmobilität von ihren Kindern erwarten, die sich in den Aspirationen der Schülerinnen und Schüler (SuS) niederschlägt. Empirisch haben sich Hinweise auf Immigrant Optimism verdichtet (Relikowski et al. 2012; Salikutluk 2016). Allerdings wird erwartet, dass sich die hohen Bildungsaspirationen in der Folge von Assimilationsprozessen langfristig an die Aspirationen der autochthonen SuS angleichen (Stanat et al., 2010).

Ein Faktor, der zur Angleichung der Aspirationen beitragen kann, ist die kulturelle Identität Heranwachsender mit Zuwanderungshintergrund, also ihr Zugehörigkeitsgefühl zu Mitgliedern des Mehrheitskontexts (nationale Identität) und des Herkunftskontexts (ethnische Identität) (Phinney et al., 2001). Theoretische Annahmen und empirische Befunde sprechen dafür, dass insbesondere die nationale Identität zum Bildungserfolg der Heranwachsenden beiträgt (Edele et al., 2013; Schotte et al., 2017), wohingegen sich für die ethnische Identität teilweise keine oder negative Zusammenhänge zeigen (Birman, 1998; Umana-Taylor et al., 2012). Außerdem wird argumentiert, dass eine hohe Identifikation mit beiden Kontexten besonders förderlich für den Bildungserfolg ist (Berry et al., 2006). Andere Arbeiten diskutieren, dass in assimilativen Mehrheitskontexten lediglich eine starke nationale Identität zum Bildungserfolg beiträgt (Makarova & Birman, 2015), während in diversitätsorientierten Mehrheitskontexten auch die ethnische Identität dazu beitragen kann. Inwiefern die kulturelle Identität zur Angleichung der Bildungsaspirationen von SuS mit und ohne Zuwanderungshintergrund beiträgt und ob sich der Zusammenhang je nach Diversitätsorientierung des Mehrheitskontexts unterscheidet, ist bislang ungeklärt. Die vorliegende Studie knüpft daran an und untersucht folgende Fragestellungen:

(1) Hängt eine ausgeprägte nationale Identität mit einer Angleichung des Aspirationsniveaus zwischen SuS mit und ohne Zuwanderungshintergrund zusammen?

Entgegen den Annahmen zum Bildungserfolg sind für SuS mit Zuwanderungshintergrund mit ausgeprägter ethnischer Identität höhere Bildungsaspirationen zu erwarten als für ihre autochthonen Peers und SuS mit Zuwanderungshintergrund mit stark ausgeprägter nationaler Identität, da die Aspirationen insbesondere aufgrund von Ressourcen des Herkunftskontexts hoch sein sollten. Für SuS mit einer lediglich stark ausgeprägten nationalen Identität ist eine Angleichung der Aspirationen an das Niveau ihrer autochthonen Peers zu erwarten (keine Unterschiede).

(2) Unterscheiden sich die Zusammenhänge zwischen Mehrheitskontexten mit unterschiedlicher Diversitätsorientierung?

In diversitätsorientierten Kontexten ist zu erwarten, dass die Bildungsaspirationen von einer ausgeprägten ethnischen Identität der SuS profitieren, wohingegen keine Vorteile für SuS mit einer ausgeprägten ethnischen Identität in assimilativen Kontexten zu erwarten sind.

Die Analysen basieren auf den Daten von 13,985 SuS mit und ohne Zuwanderungshintergrund (51% weiblich, MAlter= 14) aus Deutschland und den Niederlanden (assimilative Kontexte) sowie England und Schweden (diversitätsorientierte Kontexte), die im Schuljahr 2010/11 im Projekt Children of Immigrants Longitudinal Survey in four European Countries (CILS4EU; Kalter et al., 2016) erhoben wurden. Die dichotome abhängige Variable bildet die Aspiration, einen Universitätsabschluss zu erreichen. Als zentrale unabhängige Variablen wurden die nationale und ethnische Identität einbezogen. In den logistischen Regressionsanalysen wurde für den soziodemografischen, -kulturellen und -ökonomischen Hintergrund, kognitive Grundfähigkeiten, sprachliche Fähigkeiten sowie frühere Noten und die besuchte Schulform (Deutschland, Niederlande) kontrolliert.

Erste Ergebnisse zeigen, dass die Bildungsaspirationen von SuS mit Zuwanderungshintergrund mit ausgeprägter ethnischer Identität (ethnische Identität und integrierte Identität) unter Kontrolle der sozialen Herkunft und der Leistungen der SuS höher ausfallen als bei ihren autochthonen Peers und SuS mit einer lediglich stark ausgeprägten nationalen Identität. Dieses Muster lässt sich in allen vier Ländern beobachten, wenngleich die Unterschiede sowohl in ihrer Stärke als auch im Niveau ihrer statistischen Signifikanz variieren. Die Befunde werden vor dem Hintergrund der Diversitätsorientierungen und Durchlässigkeit der Bildungssysteme der Länder diskutiert.

## **Objektiv oder subjektiv? Was ist wichtiger für Freundschaft: das Herkunftsland oder die Stärke der ethnischen Identifikation?**

**Lars Leszczensky, Sebastian Pink**  
MZES, Universität Mannheim

Durch Migration bedingte ethnische Vielfalt ist ein wesentliches Merkmal von Schulen in Deutschland und anderen westeuropäischen Ländern. Diese zunehmende ethnische Diversität bietet verstärkte Möglichkeiten für Freundschaften zwischen Jugendlichen deutscher und nichtdeutscher Herkunft (Pettigrew und Tropp 2006; Davies et al. 2011). Allerdings wurde mehrfach gezeigt, dass interethnischer Kontakt allein für das Entstehen von interethnischen Freundschaften nicht ausreicht (Moody 2001; Wimmer und Lewis 2010). Auch in ethnisch gemischten Schulen werden Freundschaften demnach eher innerhalb als zwischen ethnischen Gruppen geschlossen (Leszczensky und Pink 2015; Smith et al. 2014). Erklärt wird dies zumeist durch das Vorhandensein ethnischer Homophilie im Sinne einer Präferenz für Freunde der gleichen ethnischen Herkunft. Ethnische Herkunft wird hierbei in der Regel mittels „objektiver“ Merkmale wie dem Geburtsland der Jugendlichen oder ihrer Eltern definiert. Nicht berücksichtigt wird hingegen, wie stark sich Jugendliche subjektiv überhaupt mit dem jeweiligen Land identifizieren.

Diese Kritik aufgreifend konnten Jugert et al. (2017) unlängst zeigen, dass Jugendliche mit Migrationshintergrund häufiger Freundschaften mit Jugendlichen ohne Migrationshintergrund schließen, wenn sie sich selbst ganz oder zum Teil als Deutsche fühlen. Nach wie vor offen ist jedoch, ob die Stärke der Identifikation mit der eigenen Herkunftsgruppe auch das Ausmaß ethnischer Homophilie beeinflusst. Einerseits könnten auch Jugendliche mit schwacher ethnischer Identifikation eine Vorliebe für Freunde der gleichen Herkunft haben, etwa aufgrund ähnlicher Erfahrungen oder Sprachkenntnisse. Andererseits sagt die Theorie der sozialen Identität (Tajfel und Turner 1986) voraus, dass die Bevorzugung von Mitgliedern der eigenen Gruppe mit zunehmender Gruppenidentifikation steigt. Demnach sollte die Präferenz für eigenethnische Freunde für Jugendliche mit starker ethnischer Identifikation ausgeprägter sein als für jene mit geringer Identifikation. Damit einhergehend stellt sich schließlich die empirische Frage, ob Jugendliche mit sehr schwacher ethnischer Identifikation überhaupt noch Tendenzen ethnischer Homophilie aufweisen, oder ob ihre subjektive (schwache) Identifikation das objektive Merkmal „sticht“.

Im vorliegenden Beitrag untersuchen wir die Bedeutsamkeit des Herkunftslands und der Stärke der ethnischen Identifikation für Freundschaftswahlen von Jugendlichen. Die empirische Untersuchung basiert auf Daten des Projektes „Freundschaft und Identität in der Schule“ (FIS). FIS ist eine Panelstudie für die mehr als 2,400 Schülerinnen und Schüler in 26 Jahrgängen der

fünften, sechsten und siebten Klassen an ethnisch diversen Schulen in Nordrhein-Westfalen befragt wurden. In der Analyse werden Daten von drei Befragungen verwendet, die im Abstand von ca. neun Monaten stattfanden.

Wir verwenden stochastische akteursorientierte Modelle (Snijders et al. 2010), die die Entstehung und Veränderung von Freundschaftsnetzwerken durch die explizite Modellierung des Zusammenspiels individueller Entscheidungen simulieren. Hierzu werden Entscheidungsregeln vorgegeben, gemäß derer Schülerinnen und Schüler Freundschaften schließen. Dies ermöglicht es unter Kontrolle gegebener Opportunitätsstrukturen wie der Gruppengröße auf das Vorhandensein entsprechender Präferenzen zu schließen, wie beispielsweise für Freunde mit einer bestimmten Herkunft oder Identität. Die Modelle werden jahrgangsweise geschätzt und anschließend in einer Meta-Analyse kombiniert.

Unsere vorläufigen Ergebnisse zeigen, dass auch Jugendliche mit schwacher ethnischer Identifikation häufiger mit Gleichaltrigen der eigenen ethnischen Gruppe befreundet sind. Folglich gibt es auch dann ein Grundlevel ethnischer Homophilie, wenn sich Jugendliche kaum mit dem Herkunftsland identifizieren. Diese allgemein vorhandene Tendenz ethnischer Homophilie wird jedoch von hoher ethnischer Identifikation verstärkt. Die Wahrscheinlichkeit, dass sich zwei Jugendliche der gleichen ethnischen Herkunft anfreunden, ist demnach am höchsten, wenn sich beide sehr stark mit dieser Gruppe identifizieren. Jugendliche mit weniger stark ausgeprägter ethnischer Identifikation werden von Gleichaltrigen der eigenen Gruppe zwar immer noch, aber in abnehmendem Maße, bevorzugt. Insgesamt deuten die Befunde darauf hin, dass es ein Mindestmaß „objektiver“ Homophilie gibt, das durch hohe ethnische Identifikation aber deutlich verstärkt wird.

## C17

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H114

### **Pädagogisch-psychologisches Wissen und Selbstwirksamkeitserwartungen zur Klassenführung**

Chair(s): **Christian Brühwiler** (Pädagogische Hochschule St.Gallen, Schweiz)

DiskutantIn(nen): **Thamar Voss** (Universität Freiburg)

Eine effektive Klassenführung gilt als evident für erfolgreichen Unterricht (W. Doyle, 2006; Helmke & Weinert, 1997) und als zentraler Bedingungsfaktor der Leistungen von Schülerinnen und Schülern (Hattie, 2010). Eine qualitätsvolle Klassenführung ist somit eine der Kernanforderungen an Lehrerinnen und Lehrer (Evertson & Weinstein, 2013) und das Wissen über Klassenführung als zentraler Teil des pädagogisch-psychologischen Wissens von Lehrpersonen zu betrachten (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Das Wissen zur Klassenführung soll zusammen mit anderen nicht-kognitiven Kompetenzaspekten, wie z.B. der Selbstwirksamkeitserwartung, dazu befähigen, Komplexität und vielfältige Ansprüche des Unterrichts zu bewältigen (Baumert & Kunter, 2006; Blömeke, Kaiser & Lehmann, 2008; Doyle, 2006).

Das Symposium setzt sich mit der Messung pädagogisch-psychologischen Wissens zur Klassenführung sowie der Erfassung der Selbstwirksamkeitserwartung zur Komplexität von Unterricht unter Berücksichtigung des Dreiländerkontexts Deutschland, Schweiz und Österreich auseinander.

Der erste Beitrag fokussiert den Einsatz des videobasierten Testinstruments Classroom Management Expertise (CME) in den unterschiedlichen Ausbildungskontexten der drei Länder Deutschland, Schweiz und Österreich. Es wird überprüft, ob sich für ein in Deutschland bewährtes valides Testinstrument Messinvarianz über die drei Länder belegen lässt. Der zweite Beitrag befasst sich mit einem nicht-kognitiven Aspekt professioneller Kompetenz von Lehrpersonen, der Selbstwirksamkeitserwartung in Zusammenhang mit der Klassenführung. Es wird der Frage nachgegangen, wie Selbstwirksamkeitserwartungen zur Komplexität von Unterricht auf der Grundlage der Komplexitätsmerkmale nach Doyle (1986; 2006) operationalisiert werden können. Der dritte Beitrag stellt ein neu entwickeltes, auf Textvignetten basierendes Testinstrument zur Erfassung pädagogisch-psychologischen Handlungswissens zur Klassenführung (PPHW-K) vor und untersucht Unterschiede zwischen Studierenden zu Beginn und am Ende der Lehrerausbildung. Der vierte Beitrag befasst sich mit dem Vergleich der beiden unterschiedlichen Testinstrumente (CME und PPHW-K) zur Erfassung von pädagogisch-psychologischem Wissen zur Klassenführung. Die Ergebnisse werden u.a. mit Blick auf die unterschiedlichen Länderkontexte diskutiert.

*Beiträge des Symposiums*

### **CME-Testung in unterschiedlichen Ausbildungskontexten: Ergebnisse aus Messinvarianzanalysen**

**Johannes König<sup>1</sup>, Charlotte Kramer<sup>1</sup>, Benita Affolter<sup>2</sup>, Andreas Bach<sup>3</sup>, Horst Biedermann<sup>2</sup>, Christian Brühwiler<sup>2</sup>, Ulrike Greiner<sup>3</sup>, Lena Hollenstein<sup>2</sup>, Michaela Katstaller<sup>3</sup>**

<sup>1</sup>Universität zu Köln, Deutschland, <sup>2</sup>Pädagogische Hochschule St.Gallen, Schweiz, <sup>3</sup>Paris Lodron Universität Salzburg, Österreich

#### Theoretischer Hintergrund

Qualitätsvolle Klassenführung gilt als eine Kernanforderung im Lehrer\*innenberuf (Evertson & Weinstein, 2013). In der empirischen Unterrichtsforschung ist die Bedeutung von Klassenführung als Prozessmerkmal für die Lernfortschritte von Schüler\*innen gut belegt (Hattie, 2012). Modellhafte Vorstellungen zur Unterrichtsqualität konzeptualisieren Klassenführung, neben kognitiver Aktivierung und konstruktiver Unterstützung, auch als eine Basisdimension, für die über unterrichtliche Domänen, Kontexte und Situationen hinweg Gültigkeit beansprucht wird (Klieme & Rakoczy, 2008). Damit Lehrer\*innen typische Herausforderungen der unterrichtlichen Klassenführung erfolgreich bewältigen können, wird schon lange auf die Notwendigkeit einer kognitiven Fundierung hingewiesen (Doyle, 1985, 2006). Der Frage, wie dieser „intellectual framework“ allerdings direkt gemessen werden kann, wird erst in jüngerer Zeit intensiver bearbeitet. Gewählt werden meist Ansätze, die eine situationspezifische Erfassung ermöglichen, um dem handlungsbezogenen Charakter von Anforderungen der Klassenführung gerecht zu werden (König, 2016).

#### Fragestellung

Die situationspezifische Erfassung von Wissen zur Klassenführung erfolgt derzeit in verschiedenen Verfahren mithilfe von videobasierten Stimuli. Die verwendeten Video-Clips haben z.B. gegenüber Textvignetten den Vorteil, komplexe Situationen mehrdimensional abzubilden. Nachteilhaft könnte jedoch sein, dass ihr Einsatz in unterschiedlichen kulturellen Kontexten nur bedingt möglich ist. Der Vortrag geht daher der Frage nach, ob sich für ein Messinstrument zur Erfassung von Pädagogisch-Psychologischem Wissen zur Klassenführung, das sich bislang bei angehenden Lehrkräften in Deutschland als valide erwiesen hat, Messinvarianz für die Ausbildungskontexte Deutschland, Schweiz und Österreich belegen lässt. Angesichts struktureller Unterschiede, jedoch sprachlich-kultureller Gemeinsamkeiten der Lehrerbildungssysteme in diesen drei Ländern (König & Rothland, 2015) wird konfigurale, metrische und skalare Messinvarianz angenommen.

#### Methode

Verwendet wird der videobasierte Test zur Erfassung von Classroom Management Expertise (CME; König, 2015; König & Kramer, 2016). Er besteht aus vier Video-Clips sowie 24 Testitems, die sich auf die gezeigten Klassenführungssituationen beziehen. Die Datenanalyse basiert auf Testdaten von 317 Lehramtsstudierenden der Universität zu Köln, rund 300 Lehramtsstudierenden der Pädagogischen Hochschule St. Gallen und rund 100 Lehramtsstudierenden der Universität Salzburg (die Datenerhebung ist zum Zeitpunkt der Abstract-Einreichung noch nicht abgeschlossen). Zur Überprüfung von Messinvarianz wird eine Serie von latenten konfirmatorischen Faktorenanalysen geprüft (Vandenberg & Lance, 2000; Van den Schoot et al., 2012). Zur Prüfung der konfiguralen Messinvarianz werden zuerst wenig restriktive, getrennte konfirmatorische Faktorenanalysen sowie eine einfache Multi-Gruppen-Analyse durchgeführt, bei denen Faktorladungen, Intercepts und Residualvarianzen frei geschätzt werden. Zur Prüfung der metrischen Invarianz wird ein konfirmatorisches Faktoranalysemodell spezifiziert, in dem die Faktorladungen über die Gruppen gleich sind, jedoch die Intercepts frei geschätzt werden. Dieser zweite Prüfschritt dient der Analyse, ob die Testpersonen über die Gruppen hinweg dem untersuchten Konstrukt gegenüber auch ein vergleichbares Verständnis entgegenbringen. Drittens werden Faktorladungen und Intercepts gleichgesetzt, um skalare Invarianz zu prüfen. Damit soll belegt werden, dass die Bedeutung des Konstrukts (Faktorladungen) und das jeweilige Niveau der dem Konstrukt zugrundeliegenden Items (Intercepts) über die Gruppen hinweg vergleichbar sind. Dies stellt üblicherweise die Voraussetzung dafür dar, um in weiterführenden Analysen Gruppen hinsichtlich ihrer erreichten Testmittelwerte vergleichen zu können (Vandenberg & Lance, 2000).

## Ergebnisse

Da die Datenerhebungen in Österreich und der Schweiz zum Zeitpunkt der Abstract-Einreichung noch nicht abgeschlossen sind, können an dieser Stelle noch keine finalen Ergebnisse zur Messinvarianz berichtet werden. Mithilfe der bereits vorliegenden Daten von Studierenden der Universität zu Köln ( $n = 317$ ) lässt sich jedoch bereits die eindimensionale Struktur in einer konfirmatorischen Faktorenanalyse für kategoriale Daten bestätigen ( $\chi^2 = 286.781$ ,  $df = 252$ ,  $p = .0652$ ,  $RMSEA = .021$ ,  $CFI = .922$ ). Auf der Tagung werden dann entsprechende Modelle präsentiert, die über die Messinvarianz des Konstrukts in den drei Ausbildungskontexten detailliert Einblick geben. Sollte sich Messinvarianz entsprechend belegen lassen, werden zudem deskriptive Befunde zum mittleren Abschneiden der Lehramtsstudierenden der drei Kontexte berichtet und zur Diskussion gestellt.

## Entwicklung einer Skala zur Erfassung von Selbstwirksamkeitserwartungen zur Komplexität von Unterricht

**Andreas Bach**

Paris Lodron Universität Salzburg, Österreich

### Theoretischer Hintergrund:

Trotz zahlreicher Studien zu Selbstwirksamkeitserwartungen von (angehenden) Lehrkräften (im Überblick Klassen, Tze, Betts & Gordon, 2011; Kleinsasser, 2014; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998; Zee & Koomen, 2016) sowie verschiedener Instrumente im englischsprachigen (u.a. Ashton and Webb, 1982; Gibson and Dembo, 1984; Tschannen-Moran and Woolfolk Hoy, 2001; Skaalvik & Skaalvik, 2007) und deutschsprachigen Raum (z.B. Schwarzer & Schmitz, 1999; Schulte, 2008) wird in der Lehrerbildungsforschung nach wie vor die fehlende konzeptuelle Klarheit des Konstrukts und dessen Operationalisierung problematisiert (Dellinger, Bobbett, Olivier & Ellett, 2008; Henson, 2002; Klassen, Tze, Betts & Gordon, 2011; Wyatt, 2014; Zee & Koomen, 2016). Die vorliegenden Instrumente erfassen häufig nicht Selbstwirksamkeit im Sinne der sozial-kognitiven Lerntheorie nach Bandura (1977, 1997). Zudem werden häufig die Aufgaben und Anforderungen von Lehrkräften im Unterricht nicht adäquat abgebildet. Bislang existiert auch kein Instrument, das den in der Unterrichtsforschung vielfach rezipierten ökologischen Ansatz zum Classroom Management nach Doyle, (1986, 2006) operationalisiert, obwohl die darin ausgearbeiteten Dimensionen der Komplexität konstitutive Strukturmerkmale des Unterrichts darstellen und somit zentrale Anforderungen an das Lehrerhandeln spezifizieren.

Die vorliegende Studie greift dieses Desiderat auf und fokussiert die Entwicklung und Validierung eines Instruments zur Erfassung von Lehrer-Selbstwirksamkeitserwartungen, welches theoretisch auf dem ökologischen Ansatz zum Classroom Management nach Doyle (1986, 2006) basiert.

### Fragestellung:

Die Studie richtet sich auf die konstruktausarbeitende und empirisch-instrumentelle Fragestellung, wie Selbstwirksamkeitserwartungen zur Komplexität von Unterricht auf Grundlage der Komplexitätsmerkmale nach Doyle (1986) operationalisiert werden können und über welche messtechnischen Eigenschaften die neu entwickelte Skala verfügt.

### Methode:

Die Datenbasis stammt aus einer Längsschnittbefragung von Lehramtsstudierenden im 1. BA-Semester an der Universität Salzburg ( $N = 234$ ; Alter:  $M = 20.6$  ( $SD = 4.0$ ); 68.7% weiblich). Der Itempool der neu entwickelten Likert-Skala umfasste 30 Items zu den sechs Subdimensionen nach Doyle (1986) (Multidimensionalität, Simultanität, Unmittelbarkeit, Unvorhersehbarkeit, Öffentlichkeit und Geschichtlichkeit). Die Items wurden basierend auf der Selbstwirksamkeitstheorie nach Bandura (1977, 1997) formuliert (Beispiel-Item: „Ich fühle mich in der Lage im Unterricht damit umzugehen, dass viele gleichzeitige Ereignisse koordiniert werden müssen.“).

Die Dimensionalität der Skala wurde in einer Vorstudie untersucht und das Instrument wurde skalenbezogen optimiert, d.h. auf drei bis vier Items pro Faktor reduziert. Die Prüfung der theoretisch postulierten Faktorenstruktur der Skala erfolgte mit Hilfe einer konfirmatorischen Faktorenanalyse mit dem Programm Mplus (Muthén & Muthén, 2009). Die Konstruktvalidität wurde untersucht, indem die Korrelationen zu bestehenden Instrumenten zur Selbstwirksamkeit (Tschannen-Moran und Woolfolk Hoy, 2001; Schwarzer und Jerusalem, 1995), zum Selbstkonzept (Pohlmann und Möller, 2010; Retelsdorf, Bauer, Gebauer, Kauper und Möller, 2014) sowie zur emotionalen Erschöpfung (Kerstin, 2008) berechnet wurden.

### Ergebnisse:

Die psychometrischen Ergebnisse für die Skala zeigen, dass sich die Merkmale komplexen Unterrichts nach Doyle (1986) als Selbstwirksamkeitserwartungen durch Items mit Likert-Skalierung operationalisieren lassen. Die postulierte sechsfaktorielle Struktur des Konstrukts konnte im Rahmen der konfirmatorischen Faktorenanalyse und des Vergleichs verschiedener Modelle (Generalfaktormodell, 6-Faktormodell, Second-Order-Modell) bestätigt werden. Für das spezifizierte Second-Order-Modell mit sechs latenten Faktoren sowie einem Faktor zweiter Ordnung konnten akzeptable Modell-Fit-Werte ermittelt werden ( $\chi^2 = 258.01$ ;  $df = 183$ ;  $p < .001$ ;  $CFI = .953$ ;  $RMSEA = .042$ ;  $SRMR = .054$ ). Die Cronbachs Alpha-Werte für die Subskalen sind zufriedenstellend bis gut ( $\alpha = .72$  bis  $\alpha = .82$ ), für die Gesamtskala exzellent ( $\alpha = .91$ ). Die Zusammenhangsanalysen ergaben erwartungskonform moderat positive Korrelationen zur Selbstwirksamkeit und zum Selbstkonzept sowie schwach negative Zusammenhänge zur emotionalen Erschöpfung.

Die deskriptiven Befunde zeigen, dass in der untersuchten Stichprobe der Mittelwert der Selbstwirksamkeitserwartungen knapp unter dem theoretischen Skalenmittelpunkt liegt. Darüber hinaus existieren Unterschiede zwischen den Dimensionen. Der höchste Mittelwert wird für die Dimension „Multidimensionalität“ erzielt, während die geringste Selbstwirksamkeitserwartung für die Dimension „Unmittelbarkeit“ angegeben wird.

## Unterschiede im pädagogisch-psychologischen Wissen zur Klassenführung zu Beginn und am Ende der Lehrerausbildung

**Benita Affolter, Lena Hollenstein, Christian Brühwiler**  
Pädagogische Hochschule St.Gallen, Schweiz

### Theoretischer Hintergrund

Ein zentrales Ziel der Lehrerausbildung ist das Herausbilden von Professionswissen (Bransford, Darling-Hammond & LePage, 2005) als ein Aspekt der professionellen Kompetenz von Lehrpersonen (Baumert & Kunter, 2011). Dieses professionelle Wissen, zu dem auch das fachunabhängige pädagogisch-psychologische Wissen gezählt wird, soll zusammen mit anderen



Kompetenzaspekten dazu befähigen, Komplexität und vielfältige Ansprüche des Lehrberufs zu bewältigen (Baumert & Kunter, 2011; Blömeke, Kaiser & Lehmann, 2008; Klieme & Leutner, 2006).

Es ist grundsätzlich davon auszugehen, dass sich pädagogisch-psychologisches Wissen während der Ausbildung aufgrund unterschiedlicher formaler und informeller Lerngelegenheiten aufbaut und weiterentwickelt (Affolter, Hollenstein & Brühwiler, in Vorbereitung). Ein Wissenszuwachs unter Berücksichtigung formaler Lerngelegenheiten konnte die Studie TEDS-FU (König, Blömeke, Klein, Suhl, Busse & Kaiser, 2014) für Primarlehrpersonen sowie die LEK-Studie (König, 2012) für Sekundarlehrpersonen nachweisen. Allerdings liegen erst wenige empirische Befunde zur Entwicklung pädagogisch-psychologischen Wissens vor.

Die in den bisherigen Studien eingesetzten Testinstrumente zur Erfassung pädagogisch-psychologischen Wissens sind zumeist inhaltlich breit konzeptualisiert (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Die empirische Validierung Tests zum pädagogisch-psychologischen Wissen fand bisher in der Forschung erst wenig Beachtung (z.B. Hohenstein, Köller & Möller, 2015). So bleibt teilweise fraglich, was mit den Tests wirklich gemessen wird.

Ein Teilbereich, welcher in nahezu allen Konzeptionen enthalten ist, ist die Klassenführung (Voss et al., 2015), die sich in drei Dimensionen unterteilen lässt: (1) Eine adäquate Unterrichtsgestaltung, (2) Beziehungsförderung in der Klasse sowie (3) Verhaltenskontrolle, welche ein konstruktiver Umgang mit Störungen sowie ein präventives und proaktives Lehrerhandeln beinhaltet (z.B. Hattie, 2010; Piwowar, 2014; Syring, 2017). Aktuell liegen erste Instrumente vor, welche sich nur mit der Klassenführung als ein Teilbereich des pädagogisch-psychologischen Wissens auseinandersetzen, beispielsweise das videobasierte Messinstrument classroom management expertise (CME) von König und Lebens (2012).

#### Fragestellung

In diesem Beitrag soll einerseits ein neu entwickeltes Messinstrument zur Klassenführung vorgestellt und validiert werden. Zudem sollen Hinweise zur Veränderung des pädagogisch-psychologischen Wissens zur Klassenführung während der Lehrerausbildung aufgezeigt werden. Folgende Fragestellungen sollen beantwortet werden: (1) Lässt sich die theoretisch postulierte Struktur des Wissenstests zur Klassenführung mit den Subdimensionen adäquate Unterrichtsgestaltung, Beziehungsförderung und Verhaltenskontrolle empirisch bestätigen? (2) Welche Unterschiede im pädagogisch-psychologischen Handlungswissen zur Klassenführung bestehen bei angehenden Primar und Sekundarlehrpersonen zwischen Beginn und Ende der Lehrerausbildung?

#### Methode

Das eingesetzte Testinstrument zur Erfassung pädagogisch-psychologischen Handlungswissens zur Klassenführung (PPHW-K) stellt eine Weiterentwicklung des Teils Klassenführung eines von Oser, Biedermann, Brühwiler, Krattenmacher und Steinmann (2008) entwickelten Testinstruments zur Erfassung pädagogisch-psychologischen Handlungswissens dar. Der PPHW-K ist ein Paper-Pencil-Test, bestehend aus neun Textvignetten zu hypothetischen Klassenraumsituationen, die den drei Klassenführungsaspekten Unterrichtsgestaltung, Beziehungsförderung und Verhaltenskontrolle zugeordnet werden können. Mit dieser kontextualisierten Erfassung soll handlungsnahes Wissen erhoben werden. Als Datengrundlage dienen rund 900 künftige Lehrpersonen der Primarstufe bzw. der Sekundarstufe I in der Deutschschweiz (je ca. 300 Studierende zu Beginn, in der Mitte und am Ende der Lehrerausbildung). Der Wissenstest wird einer Rasch-Skalierung unterzogen. Die Faktorenstruktur wird mittels konfirmatorischer Faktorenanalysen in Mplus (Muthén & Muthén, 1998-2006) überprüft.

#### Ergebnisse

Da die Datenerhebung bis zur Einreichungsfrist des Abstracts nicht abgeschlossen war, können an dieser Stelle nur theoretische Annahmen getroffen werden. So ist anzunehmen, dass die theoretisch postulierte Faktorenstruktur empirisch bestätigt werden kann. Weiter ist ausgehend von anderen Studien zur Entwicklung pädagogisch-psychologischen Wissens (z.B. König, 2012) anzunehmen, dass Studierende in höheren Semestern über ein höheres Wissen zur Klassenführung verfügen als Studierende zu Beginn des Studiums.

### **Die Erfassung von pädagogisch-psychologischem Wissen zur Klassenführung: Zwei Messinstrumente im Vergleich**

**Lena Hollenstein<sup>1</sup>, Benita Affolter<sup>1</sup>, Christian Brühwiler<sup>1</sup>, Charlotte Kramer<sup>2</sup>, Johannes König<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule St.Gallen, Schweiz, <sup>2</sup>Universität zu Köln, Deutschland

#### Theoretischer Hintergrund

Pädagogisch-psychologisches Wissen als Teil des Professionswissens von Lehrpersonen ist in den letzten Jahren vermehrt in den Fokus der erziehungswissenschaftlichen Forschung gerückt. Den im Rahmen verschiedener Forschungsprojekte (z.B. BiWiSS, Observe, ProWin, TEDS-M) entwickelten Messinstrumenten liegen unterschiedliche Konzeptualisierungen von pädagogisch-psychologischem Wissen zugrunde (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Gemeinsam ist den unterschiedlichen Instrumenten, dass sie zum Ziel haben, unterrichtsrelevantes Wissen zu erfassen, wie bspw. Strategien und Massnahmen zur Initiierung und Aufrechterhaltung lernförderlicher Bedingungen im Unterricht (Lenske, Thillmann, Wirth, Dicke & Leutner, 2015) oder Aspekte der Klassenführung wie sie bei nahezu allen Definitionen als Teilbereich des pädagogischen Wissens betrachtet wird (Voss et al., 2015).

Die unterschiedlichen Testinstrumente zur Erfassung pädagogisch-psychologischen Wissens unterscheiden sich nicht nur hinsichtlich ihrer Konzeptualisierung, sondern auch in Bezug auf die Testverfahren. Man findet Papier-Bleistift-Tests vs. computerbasierte Tests, Unterschiede hinsichtlich der Aufgabenformate (z.B. Testaufgaben vs. Vignetten), offene vs. geschlossene Antwortformate oder auch Unterschiede nach der Wissensart, z.B. eher deklaratives oder prozedurales Wissen. Dabei bezieht sich prozedurales Wissen auf konkrete Handlungsabläufe und gilt somit als besonders handlungsrelevant (Artelt & Wirth, 2014). Dieser Situationsbezug des prozeduralen Wissens wird in der aktuellen kompetenztheoretischen Diskussion aufgegriffen und bisherige Kompetenzmodelle um situations- und verhaltensnahe kognitive Fähigkeiten als vermittelnde Eigenschaften (Wahrnehmung, Interpretation, Entscheidung) zwischen Disposition und Performanz ergänzt (Blömeke, Gustafsson & Shavelson, 2015).

Bei diesem Transformationsprozess von der Kompetenz zur Performanz setzt die vorliegende Untersuchung an. Es werden zwei Instrumente, welche pädagogisch-psychologisches Wissen zur Klassenführung messen, miteinander verglichen. Dabei handelt es sich um das videobasierte Testinstrument zur Erfassung von Classroom Management Expertise (CME; König, 2016; König & Kramer, 2015) sowie ein neu entwickeltes textvignettenbasiertes Messinstrument zur Erfassung pädagogisch-psychologischen Handlungswissens zur Klassenführung (PPHW-K). Während der videobasierte Test CME vor allem die Bereiche Wahrnehmung und Interpretation testet, zielt der PPHW-K auf die Bereiche Interpretation und Entscheidung.

## Fragestellung

Ziel der geplanten Untersuchung ist es, die inhaltliche Validität der beiden Instrumente zur Erfassung pädagogisch-psychologischen Wissens zur Klassenführung zu untersuchen. Es soll geprüft werden, welche Zusammenhänge zwischen dem videobasierten Messinstrument zur Klassenführungsexpertise (CME) und dem textvignettenbasierten Test zum pädagogisch-psychologischen Handlungswissen zur Klassenführung (PPHW-K) bestehen.

## Methode

Die beiden Wissenstests zur Klassenführung werden bei Studierenden des 5. Semesters der Primar- und Sekundarlehrepersonenausbildung in der Deutschschweiz eingesetzt. Es stehen Daten von ca. 300 Studierenden zur Verfügung.

Das erste eingesetzte Testinstrument (CME) erfasst die Klassenführungsexpertise anhand von vier Video-Clips sowie 24 Testitems. Diese fokussieren auf die Wahrnehmung und Interpretation der einzelnen Klassenführungssituationen (König & Kramer, 2015).

Das zweite Instrument zur Erfassung pädagogisch-psychologischen Handlungswissens zur Klassenführung (PPHW-K) ist ein Papier-Bleistift-Test, der sich aus neun Textvignetten zu den drei Klassenführungsaspekten Unterrichtsgestaltung, Beziehungsförderung und Verhaltenskontrolle zusammensetzt. Es ist eine Weiterentwicklung des Instruments zur Erfassung pädagogisch-psychologischen Handlungswissens von Oser, Biedermann, Brühwiler, Krattenmacher und Steinmann (2008), dessen prädiktive Validität in Bezug auf die Unterrichtsqualität aufgezeigt werden konnte (Brühwiler, Hollenstein, Affolter, Biedermann & Oser, in Vorbereitung). Das Instrument zielt zum einen darauf ab, dass Lehrpersonen anhand von kritischen Klassenführungssituation entscheiden müssen, wie sie handeln würden. In einem zweiten Teil wird nach theoretischem Erklärungswissen gefragt, was die Interpretation der jeweiligen Situation voraussetzt.

## Ergebnisse

Da die Datenerhebungen zum Zeitpunkt der Abstract-Einreichung noch nicht abgeschlossen sind, können an dieser Stelle keine Ergebnisse berichtet werden. Aufgrund der theoretischen Annahmen ist zu erwarten, dass die beiden Testinstrumente mittlere Korrelationen aufweisen, da sie teilweise auf ähnliche theoretische Hintergründe rekurrieren (z.B. Übergänge gestalten oder Schülerrückmeldungen erteilen). Allerdings fokussieren die beiden Testinstrumente unterschiedliche situationspezifische Fähigkeiten (Wahrnehmen vs. Entscheiden), weshalb die Zusammenhänge nicht allzu hoch ausfallen dürften.

## C18

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H102

### Soziale und individuelle Einflussfaktoren zum Schulerfolg von Sekundarschüler/-innen

Chair(s): **Frances Hoferichter** (Universität Greifswald, Deutschland)

DiskutantIn(nen): **Clemens Hillenbrand** (Carl von Ossietzky Universität Oldenburg)

Schulische Lernprozesse sind in einen sozial-interaktiven Kontext eingebettet, der das Sozialverhalten (soziale Beziehungen), Schulengagement (aktive Einbindung in das Schulgeschehen) und akademische Selbstkonzept (Leistungsvergleich mit Mitschülern) maßgeblich beeinflusst (Wentzel & Ramani, 2016), die wichtige Einflussfaktoren für akademischen Erfolg und positive Entwicklungsverläufe von Kindern und Jugendlichen darstellen (Algozzine, Wang, & Violette, 2011; Valentine, DuBois, & Cooper, 2004; Wonglorsachon, Wongwanich, & Wiratchai, 2014). Im Rahmen des Symposiums gilt es, diese Einflussfaktoren interdisziplinär und empirisch unter Sekundarschüler/-innen näher zu beleuchten, mit dem Ziel theoretische Modelle zu erweitern und schulpraktische Implikationen abzuleiten. In diesem Sinne möchte das Symposium evidenzbasierte Handlungsempfehlungen geben, um einen Beitrag zum professionellen Handeln im Bildungskontext zu leisten.

Der erste Beitrag beschäftigt sich mit der Wirkung eines kognitiv-behavioral verankerten Trainings von Schüler/-innen zur Förderung der sozial-kognitiven Informationsverarbeitung. Bei diesem Training liegt der Fokus auf Schüler/-innen der Ober-/Haupt- und Förderschule mit dem Förderschwerpunkt Lernen. Die Evaluation des Trainings zeigt verschiedene Wirkmechanismen für unterschiedliche Risikogruppen auf. Der zweite Beitrag geht der Frage nach, inwiefern behaviorales und emotionales Schulengagement den Zusammenhang von sozialen schulischen Beziehungen und Verhaltensauffälligkeiten während der Adoleszenz mediiert. Die Ergebnisse zeigen, dass die Eingebundenheit in den Schulalltag als präventiver Faktor zum Abbau von Verhaltensstörungen fungieren kann. Im dritten Beitrag wird mittels Messung von Sozialverhalten von Schüler/-innen aus einer Multi-Informanten-Perspektive untersucht, welche Mechanismen dazu führen, dass sich die Selbst- und Fremdeinschätzung von Schülern/-innen und deren Eltern signifikant unterscheiden. Die Ergebnisse zeigen, dass das Geschlecht bei Schüler/-innen und die soziale Herkunft der Eltern einen Einfluss auf die jeweilige Einschätzung des Sozialverhaltens darstellen. Der vierte Beitrag widmet sich der Frage, ob der Big-Fish-Little-Pond-Effekt (BFLPE) auch für die vier Facetten des schulischen Selbstkonzepts (domain-unabhängig) in unterschiedlich leistungsstarken Klassen an Gymnasien bestätigt werden kann. Die Ergebnisse der Mehrebenenanalyse zeigen, dass der BFLPE für alle vier Facetten des schulischen Selbstkonzepts (sozial, kriterial, individuell und absolut) besteht.

Das interdisziplinäre Symposium leistet einen empirisch-fundierte Beitrag für das Verständnis und die Wirkungsweisen von sozialen und individuellen Einflussfaktoren auf den Schulerfolg von Sekundarschüler/-innen. Die Ergebnisse dienen der Weiterentwicklung des professionellen Handelns aller Bildungsakteure und begünstigen somit Entwicklungsverläufe und Bildungserfolge von Jugendlichen.

*Beiträge des Symposiums*

### Prävention von Schulabbruch - Evaluation eines Trainings der sozial-kognitiven Informationsverarbeitung für Jugendliche mit hohem Entwicklungsrisiko

**Marie-Christine Vierbuchen**

Universität Vechta

Theoretischer Hintergrund

Die Partizipation von Schülerinnen und Schülern am Schulalltag und an der Klassengemeinschaft spielt eine wichtige Rolle für den Schulerfolg (Hagen, Hillenbrand & Vierbuchen, 2012). Dafür ist eine gelingende sozial-kognitive Informationsverarbeitung (Crick & Dodge, 1994; Lemerise & Arsenio, 2000) eine wesentliche unterstützende Ressource. In ungünstiger Ausprägung ist die sozial-kognitive Informationsverarbeitung eng mit aggressivem Verhalten und weiteren Entwicklungsschwierigkeiten verbunden (Fontaine & Dodge, 2009; Lansford, Malone, Dodge, Pettit & Bates, 2010). Schülerinnen und Schüler mit geringer sozialer Kompetenz, schwacher Lesekompetenz, hoher psychosozialer Belastung, geringem Integrationsstatus in der Klasse oder auffälligem Schulbesuchsverhalten tragen spezifische Risikofaktoren für ungünstige Entwicklungsverläufe ihrer Schullaufbahn bis zum Schulabbruch (Hammond, Linton, Smink & Drew, 2007; Hennemann, Hagen & Hillenbrand, 2010). Im Sinne eines transaktionalen Entwicklungsverständnisses (Sameroff & Fiese, 2000; Lösel & Bender, 2003; Beelmann & Raabe, 2007) kann eine unterstützende Maßnahme im Jugendalter relevante Risikofaktoren direkt fokussieren und so präventiv wirksam werden. Solche oft kognitiv-behavioral verankerten Trainings sind nach Hammond et al. (2007) unter den individuumzentrierten Maßnahmen zu fassen und zählen zu den wirksamsten Präventionsmaßnahmen in Bezug auf Schulabbruch (Wilson, Lipsey & Derzon, 2003; Cobb, Sample, Alwell & Johns, 2005).

Fragestellung

Wirkt sich ein Training der sozial-kognitiven Informationsverarbeitung auf spezifische Risikofaktoren für Schulabbruch kurzfristig (direkt nach Abschluss des Trainings) und langfristig (sechs Monate nach Abschluss des Trainings) aus? Und wie wirkt sich dieses Training für Jugendliche mit besonders ausgeprägtem Risiko kurz- und langfristig aus?

Methode

Die vorliegende Untersuchung konzipiert und evaluiert ein Trainingsprogramm zur Unterstützung der sozial-kognitiven Informationsverarbeitung und überprüft dessen Wirksamkeit auf spezifische Risikofaktoren für schulischen Dropout in Klassen der siebten bis neunten Jahrgangsstufe an Haupt- und Oberschulen sowie Förderschulen mit dem Schwerpunkt Lernen in Niedersachsen (N=373, Alter M=14;6). Die Erhebung wurde in einem Pre-Post-Follow Up-Design mit Kontroll- und Experimentalgruppe durchgeführt und anhand von zweifaktoriellen Varianzanalysen mit Messwiederholung analysiert. Neben der Wirksamkeit für alle Schülerinnen und Schüler werden im Vortrag Subgruppen mit besonders hoher Risikoausprägung in den Fokus der Analysen genommen.

Ergebnisse

Für die Gesamtstichprobe kann eine positive Wirksamkeit im Bereich des Sozialverhaltens (Lehrereinschätzliste für Sozial- und Lernverhalten, Petermann & Petermann, 2013) mit einer kleinen Effektstärke (pre-post:  $dkorr=.23$ ,  $p=.000$ ; pre-post-follow up:  $dkorr=.27$ ,  $p=.004$ ) und im Bereich der psychosozialen Belastung (Strength and Difficulties Questionnaire, Goodman, 1997) kurzfristig mit kleiner (pre-post:  $dkorr=-.23$ ,  $p=.011$ ), langfristig mit mittlerer korrigierter Effektstärke (pre-post-follow up:  $dkorr=-.54$ ,  $p=.000$ ) aus Lehrkraftperspektive festgestellt werden.

Die Subgruppenanalyse zeigt ein komplexes Bild; zusammenfassend stellt sich eine kurzfristige positive Wirksamkeit auf die Dimension des Sozialverhaltens mit kleiner bis mittlerer Effektstärke dar (z.B. Subgruppe unterdurchschnittliches Sozialverhalten  $dkorr=.24$ ,  $p=.045$ ) und langfristig eine signifikante Senkung der psychosozialen Belastung mit mittlerer Effektstärke (z.B. Subgruppe auffälliges Schulbesuchsverhalten  $dkorr=-.56$ ,  $p=.026$ ).

Die Ergebnisse zeigen eine deutliche Wirksamkeit der Präventionsmaßnahme (Hagen, 2014; Vierbuchen, 2015; Hagen, Vierbuchen, Hillenbrand & Hennemann, 2016), machen jedoch auch deutlich, dass für die heterogene Zielgruppe mit starker Belastung und hohem Entwicklungsrisiko weitere flankierende und spezifische Maßnahmen zur Prävention und Intervention im Sinne eines mehrdimensionalen Bausteinkonzepts wie beispielweise nach Ricking (2009) notwendig sind.

## **Der Einfluss von sozialen Beziehungen auf Verhaltensstörungen bei Jugendlichen – spielt schulisches Engagement eine Rolle?**

**Alexander Lätsch, Diana Raufelder**  
Universität Greifswald

### Theoretischer Hintergrund

In den vergangenen Jahrzehnten war ein Anstieg der Prävalenz von Verhaltensstörungen im Jugendalter zu verzeichnen (Dryfoos, 1990; Feldman & Elliot, 1990; Simons- Morton et al., 1999), was dramatisch ist, insofern Studienbefunde zeigten, dass diese mit diversen negativen Folgen einhergehen – beispielsweise mit schlechteren Schulleistungen, Schulabbruch, aber auch selbstverletzenden Verhaltensweisen oder Suizid (Dryfoos, 1990; Feldmann & Elliott, 1990; Jessor & Jessor, 1977; Simons-Morton et al., 1999). Verhaltensstörungen im Schulkontext werden in Anlehnung an Goodman (1997) als Schummeln, Wut, impulsives Verhalten, oder Ungehorsam verstanden, welche oftmals mit delinquentem Verhalten im späteren Verlauf der Adoleszenz einhergehen (Loeber, 1988; Loeber & Dishion, 1983). Aus Studien ist bekannt, dass Verhaltensstörungen und Delinquenz negativ mit sozialen Beziehungen (social bonds) in der Schule (mit Lehrern und/oder Peers) in Zusammenhang stehen (Cernkovich & Giordano, 1992; Liska & Reed, 1985). Umgekehrt, stehen positive schulische Beziehungen und schulisches Engagement (emotionales und behaviorales Schulengagement) positiv miteinander in Beziehung, wobei schulisches Engagement als eine weitere Form von social bonds (Bindung an die Institution Schule) verstanden werden kann, die Verhaltensstörungen im Schulalltag entgegenwirken kann (Hawkins et al., 1992; Murray & Greenberg, 2001; Werner & Smith, 1989).

Als theoretische Grundlage der vorliegenden Studie dient die Social Bond Theory (Hirschi, 1969), die soziale Beziehungen und deviantes Verhalten in ihrem Theorem vereint. Nach Hirschi (1969) sollen soziale Beziehungen (attachment, commitment, involvement und belief) ein Individuum davon abhalten, deviant oder delinquent zu handeln.

Attachment manifestiert sich nach Hirschi (1969) in positiven sozialen Beziehungen mit Lehrern und Peers, was gleichsam dafür sorgt, dass Schüler/-innen sich in der sozialen Gruppe und der Institution Schule engagieren. Außerdem wird attachment als Sensibilität eines Individuums gegenüber der Einstellung anderer verstanden (Eifler, 2002; Hirschi, 1969). Commitment beschreibt das Ausmaß, in dem ein Individuum in konventionelle Ziele investiert – Individuen, die viel in konventionelle Ziele investieren haben eine geringere Wahrscheinlichkeit Handlungen auszuführen, die den Erfolg gefährden (Eifler 2002; Hirschi 1969). Das Konzept Involvement umfasst das Ausmaß, in dem ein Individuum an konventionelle Aktivitäten gebunden ist – je höher dieses Ausmaß, desto weniger Zeit bleibt deviante Verhaltensweisen auszuführen (Eifler 2002; Hirschi 1969). Das Konzept belief hingegen beschreibt den Glauben eines Individuums an die Legitimität des Systems einer Gesellschaft, sowie deren Normen und Werte. Wie oben angeführt stehen soziale Beziehungen (social bonds), Schulengagement und Verhaltensstörungen miteinander in Beziehung, wobei Schulengagement möglicherweise diese Beziehung im Verlauf der Adoleszenz mediiert – was in bisherigen Studien bislang empirisch nicht untersucht wurde.

### Fragestellung

Um diese Forschungslücke zu schließen wurden – aufbauend auf der Social Bond Theory– folgende Hypothesen in dieser Studie überprüft:

(1) Emotionales Schulengagement (ESE) mediiert die Beziehung zwischen (negativen) Lehrer-Schüler und Schüler-Schüler Beziehungen in der frühen Adoleszenz und Verhaltensstörungen in der mittleren Adoleszenz.

(2) Behaviorales Schulengagement (BSE) mediiert die Beziehung zwischen (negativen)

Lehrer-Schüler und Schüler-Schüler Beziehungen in der frühen Adoleszenz und Verhaltensstörungen in der mittleren Adoleszenz.

### Methode

Die vorliegende Studie basiert auf Daten einer längsschnittlichen Untersuchung des von der Volkswagen-Stiftung finanzierten SELF-Projekts. Als Grundlage dienen hierbei Selbstauskunftsdaten von insgesamt 1088 Schüler/-innen (Time 1: Mage = 13.70, SD = 0.53; Time 2: N= 845, Mage= 15.32, SD = .49) aus 23 Sekundarschulen in Brandenburg, die anhand von latenten Strukturgleichungsmodellen mit indirekten Effekten (Mediation) in Mplus 7.2 (Muthén & Muthén, 1998–2012) analysiert wurden.

### Ergebnisse

Die erste Hypothese konnte bestätigt werden, insofern ESE die Beziehung vollständig mediiert. D.h., sind Schüler/-innen emotional engagiert, kann das den Zusammenhang zwischen negativen sozialen Beziehungen und Verhaltensstörungen aufheben. BSE mediiert die Beziehung zwischen sozialen Beziehung in der frühen und Verhaltensstörungen in der mittleren Adoleszenz nur teilweise, insofern sie den Zusammenhang abschwächt. Insgesamt lässt sich festhalten, dass schulisches Engagement ein wichtiger Präventions- und Interventionsansatz für mögliche Verhaltensstörungen in der Adoleszenz darstellt.

## **Die soziale Kompetenz von Zwölfjährigen. Ein Vergleich der Fremd- und Selbsteinschätzung**

**Carina Schönmoser**  
Leibniz-Institut für Bildungsverläufe e. V.

### Theoretischer Hintergrund

Die objektive Messung von Persönlichkeits- und Verhaltenscharakteristika ist kostenintensiv und zeitaufwändig, da das Sozialverhalten in verschiedenen sozialen Kontexten und Gruppen variieren kann (Achenbach, McConaughy & Howell, 1987). Deshalb erfolgt die Erhebung sozialen Verhaltens - anstelle einer Beobachtung der Zielpersonen in den verschiedenen sozialen Situationen - meist anhand von Selbst- und Fremdeinschätzungen mittels eines Fragebogens (Multi-Informanten-Perspektive). Um ein ganzheitliches Verhaltensbild der Zielperson zu erlangen, welches wiederum als Grundlage für

Interventionen dient, greifen die ForscherInnen auf die Einschätzungen von Eltern, Geschwistern, LehrerInnen, Peers und der Zielpersonen selbst zurück. Diese Einschätzungen korrelieren jedoch nur geringfügig miteinander (Wirtz & Caspar, 2002).

Das prosoziale Verhalten und das Problemverhalten sind zwei Indikatoren der sozialen Kompetenz, welche Kanning (2009) als „die Gesamtheit des Wissens, der Fähigkeiten und Fertigkeiten einer Person, welche die Qualität des eigenen Sozialverhaltens - im Sinne sozial kompetenten Verhaltens - fördert“ versteht. Die Fähigkeit zu sozialem Verhalten stellt einen wichtigen Aspekt menschlicher Entwicklung dar und gilt als (Langzeit-)Prädiktor für akademische Leistungsfähigkeit oder beruflichen Erfolg (Denham, Wyatt, Bassett, Echeverria, & Knox, 2009). In diesem Sinne dient die valide Messung sozialer Kompetenzen der Beobachtung und Erforschung einer gesunden Entwicklung der Kinder und Jugendlichen. Voraussetzung dafür ist jedoch das Verständnis für die Daten, welche wir über die soziale Kompetenz unserer Zielpersonen mittels der Multi-Informanten-Perspektive erheben.

Es existieren bereits einige Forschungsarbeiten, die eine geringe Übereinstimmung der Einschätzungen des Sozialverhaltens im Zuge der Multi-Informanten-Perspektive belegen (u.a. Satake, Yoshida, Yamashita, Kinukawa, & Takagishi, 2003; Koskelainen 2008; Elberling, Linneberg, Olsen, Goodman, & Skovgaard, 2010; Stone, Otten, Engels, Vermulst, & Janssens, 2010). Jedoch nur wenige Studien haben sich bisher intensiver mit den zugrundeliegenden Mechanismen auseinandergesetzt, welche die Einschätzung von Personen beeinflussen. So könnten auch andere Faktoren wie das Alter, das Geschlecht und die soziale Herkunft aber auch die mathematischen und Lesekompetenzen der Kinder zu unterschiedlichen Einschätzungen führen. Einen ersten Hinweis darauf liefern Korsch und Petermann (2013), die zeigen konnten, dass die Elterneinschätzungen zu positivem als auch zu negativem Kindsverhalten systematisch höher ausfallen als die ErzieherInnen-Einschätzungen.

#### Fragestellung

Es ergeben sich für die vorliegende Arbeit folgende Forschungsfragen:

1. Inwiefern stimmen die Einschätzungen der Eltern und der Zielpersonen miteinander überein?
2. Sind Kind- und Familienmerkmale (Alter, Geschlecht, soziale Herkunft und kognitive Kompetenzen) der Grund für die unterschiedlichen Einschätzungen der Eltern und Zielpersonen zum sozialen Verhalten der Zwölfjährigen?

#### Methode

Mit den Daten der zweiten Welle der Startkohorte 3 des Nationalen Bildungspanels (NEPS, Blossfeld, Roßbach, & von Maurice, 2011) nimmt vorliegender Beitrag einen Vergleich der Eltern- und Selbsteinschätzung der sozialen Kompetenz von Zwölfjährigen vor (N=2942). Die soziale Kompetenz wird über zwei Subskalen des Fragebogens zu Stärken und Schwächen - der Skala für prosoziales Verhalten und der Skala für Problemverhalten (Strengths and Difficulties Questionnaire - SDQ, Goodman, 1997) - operationalisiert. Neben dem Alter und dem Geschlecht werden die mathematischen und die Lesekompetenzen der Zielpersonen sowie der soziale Hintergrund der Eltern in die Analysen aufgenommen. Zur Beantwortung der Forschungsfragen wird eine Mehrebenenanalyse durchgeführt, um der genesteten Datenstruktur gerecht zu werden.

#### Ergebnisse

Die Einschätzungen der Eltern und der Zielpersonen zum Sozialverhalten von Zwölfjährigen stimmen – wie erwartet - nur gering überein ( $r = 0.23/0.41$  für prosoziales/Problemverhalten). Es zeigt sich weiterhin, dass sich die Einschätzungen der Eltern und die der Jugendlichen hinsichtlich des Zusammenhangs mit den betrachteten Faktoren unterscheiden. Für die Einschätzungen der Jugendlichen zeigen sich statistisch signifikante Korrelationen mit dem eigenen Geschlecht und der mathematischen Kompetenz der Jugendlichen, bei den Elterneinschätzungen spielt hingegen der eigene soziale Hintergrund eine bedeutende Rolle. Die Fremd- und Selbsteinschätzungen zum sozialen Verhalten von Zwölfjährigen werden somit systematisch von unterschiedlichen Faktoren moderiert. Die Ergebnisse werden anhand anschaulicher Grafiken dargestellt und im Anschluss diskutiert.

### **Die Rolle des BFLPE für die vier Selbstkonzept-Facetten unter Gymnasiasten**

**Frances Hoferichter<sup>1</sup>, Alexander Lätsch<sup>1</sup>, Rebecca Lazarides<sup>2</sup>, Diana Raufelder<sup>1</sup>**

<sup>1</sup>Universität Greifswald, <sup>2</sup>Universität Potsdam

#### Theoretischer Hintergrund

Das (schulische) Selbstkonzept ist eines der wichtigsten Kompetenzkonstrukte für akademischen Erfolg (Valentine, DuBois, & Cooper, 2004). Nach Marsh (1990) besagt der Big-Fish-Little-Pond-Effect (BFLPE), dass vergleichbar begabte Schüler in einer leistungsstarken Klasse ein geringeres Selbstkonzept (SK) ausbilden als Schüler, die eine Klasse mit geringem Leistungsniveau besuchen. Dieser Effekt, auch negativer BFLPE genannt, konnte insbesondere in homogenen Klassen an Gymnasien nachgewiesen werden, in denen z.B. die Matheklassenleistung relativ hoch war und das individuelle Mathe-Selbstkonzept folglich gering (Marsh, Köller, & Baumert, 2001). Zahlreiche Studien haben den BFLPE länderübergreifend untersucht, wobei zumeist Aussagen über das fachspezifische Selbstkonzept in Verbindung mit der Fächernote getroffen wurden (e.g., Chiu, 2012; Marsh & Hau, 2003; Preckel & Brüll, 2010; Trautwein, Gerlach, & Lüdkte, 2008; Wouters, De Fraine, Colpin, Van Damme, & Verschueren, 2012). Bislang gibt es jedoch nur wenige Studien, die alle vier Facetten des schulischen SK (SESSKO) nach Schöne, Dickhäuser, Spinath und Pelster (2002) im Zusammenhang mit dem Notendurchschnitt der Klasse für Gymnasiasten untersuchen. SESSKO beinhaltet das soziale SK (basierend auf dem Vergleich mit Mitschülern), das individuelle SK (basierend auf dem Vergleich mit den eigenen Fähigkeiten in der Vergangenheit), das kriteriale SK (basierend auf der Einstufung der eigenen Leistung orientiert an einem sachlichen Kriterium) und das absolute SK (schließt keine soziale oder individuelle Bezugsnormen mit ein) (Schöne et al., 2002).

#### Fragestellung

Basierend auf der oben spezifizierten theoretischen Annahme und den empirischen Studien stellten wir folgende Hypothesen auf:

- 1) Je besser der individuelle Notendurchschnitt ist, desto höher sind die Facetten des SK (absolut, individuell, kriterial, sozial) ausgeprägt.
- 2) Im Vergleich zweier Schüler aus unterschiedlichen Klassen miteinander, zeigen Schüler in leistungsstarken Klassen im Vergleich zu Schülern in leistungsschwächeren Klassen ein geringeres absolutes, individuelles, kriteriales und soziales SK.

#### Methode

Insgesamt nahmen 779 Gymnasiasten (MAlter= 13.09; SD= .50, 57% Mädchen) aus 47 achten Klassen an der Fragebogenstudie im Herbst 2011 teil. Um die Hypothesen zu überprüfen wurden vier Multilevel-Strukturgleichungsmodelle

(MLSEM) mit den jeweiligen Facetten des schulischen Selbstkonzepts und deren Zusammenhang mit den Noten (rekodiert und gemittelt aus der Deutsch-, Mathe-, Englisch- und Biologienote,  $M= 4.02$ ,  $SD= .64$ ) erstellt.

#### Ergebnisse

Die Ergebnisse zeigen, dass der Zusammenhang zwischen den Noten und allen Facetten des Selbstkonzepts positiv ist: absolutes SK ( $\beta= .47$ ,  $SE= .05$ ,  $p < .001$ ), individuelles SK ( $\beta= .30$ ,  $SE= .08$ ,  $p < .001$ ), kriteriales SK ( $\beta= .51$ ,  $SE= .05$ ,  $p < .001$ ) und soziales SK ( $\beta= .54$ ,  $SE= .06$ ,  $p < .001$ ). Somit konnte die erste Hypothese bestätigt werden, insofern Schüler mit durchschnittlich besseren Noten ein höheres SK (absolut, individuell, kriterial, sozial) aufweisen. Ebenso konnte die zweite Hypothese bestätigt werden, insofern der Kontexteffekt zeigt, dass im Vergleich zweier Schüler mit gleichen Noten aber aus unterschiedlich leistungsstarken Klassen die Schüler aus leistungsstärkeren Klassen ein niedrigeres absolutes ( $\beta= -.27$ ,  $SE= .09$ ,  $p < .05$ ), individuelles ( $\beta= -.25$ ,  $SE= .12$ ,  $p < .05$ ), kriteriales ( $\beta= -.51$ ,  $SE= .19$ ,  $p < .05$ ) und soziales SK ( $\beta= -.47$ ,  $SE= .09$ ,  $p < .001$ ) aufweisen. Die zentrierten Kontexteffekte zeigen darüber hinaus, dass mit jeder Verbesserung der Klassennote um eine Standardabweichung das absolute SK um  $-.20$  SD, das individuelle SK um  $-.13$  SD, das kriteriale SK um  $-.51$  SD und das soziale SK um  $-.99$  SD geringer ausfällt.

Die Ergebnisse bestätigen den negativen BFLPE für alle Facetten des SK bei Gymnasiasten und deuten insbesondere darauf hin, dass das soziale SK am stärksten durch die Leistungen von Mitschülern beeinflusst wird.

## C19

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H001

### Forschungskompetenzen im Lehramtsstudium

Chair(s): **Franziska Böttcher** (Freie Universität Berlin, Deutschland), **Felicita Thiel** (Freie Universität Berlin)

DiskutantIn(nen): **Johannes Bauer** (Universität Erfurt)

Neben der Vermittlung von fachlichem Wissen ist die Vermittlung von Forschungskompetenz eine definierte Zielkompetenz von Universitätsabschlüssen (Wissenschaftsrat, 2006). Auch für pädagogische, evidenzbasierte Expertenberufe wie das Lehramt gilt der Anspruch, in der Praxis wissenschaftliche Evidenz bei der Begründung von Entscheidungen und bei der Lösung konkreter Probleme nutzen zu können, sowie methodisch gestützt eigene Erkenntnisse, z.B. zur Evaluation der eigenen Unterrichtspraxis, generieren zu können (Kultusministerkonferenz, 2004). Daher wird das Thema „Forschungskompetenzen im Lehramtsstudium“ immer stärker fokussiert.

Dieses Thema kann aus unterschiedlichen Perspektiven behandelt werden: aus der Perspektive der Entwicklung von institutionellen Tools, die es Lehrkräften erleichtern, empirische Befunde für ihre Praxis zu nutzen – sowie aus der Perspektive des capacity building von Akteuren, die basale Forschungskompetenzen benötigen, um diese Tools effektiv für die Entwicklung ihrer Unterrichtspraxis verwenden zu können

In dem geplanten Symposium sollen beide Perspektiven thematisiert werden. Zunächst wird gefragt, wie Tools beschaffen sein müssen, die Forschungsbefunde für die Lehrerbildung niedrigschwellig und praxisbezogen aufbereiten. Komplementär wird thematisiert, wie Forschungskompetenzen im Lehramtsstudium vermittelt werden (können) und welche Rolle Überzeugungen beim Erwerb von Forschungskompetenzen spielen.

Der erste Beitrag widmet sich am Beispiel des „Clearing House Unterricht“ der Technischen Universität München konzeptionell der Frage, wie Forschungsbefunde qualitativ und adressatengerecht aufbereitet werden können, um eine evidenzbasierte Lehrerbildung bestmöglich zu unterstützen.

Der zweite Beitrag widmet sich der Frage, inwiefern sich Studierende in Abhängigkeit ihrer Überzeugung zum Nutzen von Forschung in Gruppen einteilen lassen und ob sich diese Gruppe hinsichtlich ihrer Forschungskompetenzen und ihrer Einschätzung der Wichtigkeit von Forschung im Unterrichtskontext unterscheiden.

Die Beiträge drei und vier widmen sich der Entwicklung und Förderung von Forschungskompetenzen im universitären Kontext. Zum einen wird die Veränderung von Forschungskompetenzen während eines Semesters untersucht (Beitrag 3). Zum anderen wird eine Intervention zur gezielten Förderung von Forschungskompetenzen durch zwei unterschiedliche Treatments vorgestellt (Beitrag 4).

#### Beiträge des Symposiums

### Das Clearing House Unterricht: Analyse und adressatengerechte Kommunikation wissenschaftlicher Evidenz für die Lehrerbildung

**Maximilian Knogler, Tina Seidel, Andreas Hetmanek, Sog Yee Mok, Anne Birte Wiesbeck**

Technische Universität München

#### Die Herausforderung

Der Lehrberuf wird in der nationalen wie internationalen Diskussion um Standards von Lehrerprofessionalität zunehmend als eine forschungsbasierte Profession begriffen (Bauer & Prenzel, 2012; Darling-Hammond & Bransford, 2005). Lehrerbildende unterschiedlicher Phasen haben daher den Auftrag aktuelle empirische Evidenz in die Lehrerbildung zu integrieren. Sie müssen sich mit dem rasant entwickelnden Forschungsstand kontinuierlich auseinandersetzen und Entscheidungen darüber treffen, welche Evidenz für die ihnen anvertrauten zukünftigen Lehrerinnen und Lehrer qualitativ und relevant ist, und wie sie diese Evidenz in Lehre und Ausbildung vermitteln wollen (Seidel et al., 2016). Aktuelle Entwicklungen im Bereich der empirischen Bildungsforschung zeigen, dass sich diese hinsichtlich Spezialisierungsgrad und wissenschaftliche Produktivität (z.B. in Form von Publikationen) deutlich gesteigert hat (vgl. Beelmann und Bliesener, 1994). Auf der Basis verschiedener Forschungszugänge entsteht kontinuierlich eine Fülle an Ergebnissen, die unterschiedlich interpretierbar sind und in ihrer Gesamtheit für die Lehrerbildung relevante empirische Evidenz bereitstellen. In verschiedenen Disziplinen, wie etwa der Medizin, haben sich in diesem Zusammenhang sogenannte „Clearing Houses“ etabliert (Slavin, 2008). Ein Clearing House übernimmt die Aufgabe, den aktuellen Forschungsstand themenbezogen zu ordnen, zusammenzufassen und Kriterien zu deren Beurteilung festzulegen. Damit soll es Praktikerinnen und Praktikern erleichtert werden, sich in Bezug auf aktuelle Forschungsevidenz fortzubilden (Stark, 2017). Dieser Beitrag stellt einen Ansatz vor, mit dem durch eine adressatengerechte Bereitstellung aktueller Forschungsbefunde eine Brücke zwischen Forschung und Lehrerbildungspraxis gebaut werden soll: Das Clearing House Unterricht.

#### Die Evidenzgrundlage

Das Clearing House Unterricht zielt darauf ab den aktuellen Forschungsstand zu wichtigen Themen des MINT-Unterrichts zusammenzufassen und für die Lehrerbildung aufzubereiten. Dabei werden als Evidenzgrundlage Meta-Analysen herangezogen. Diese verfolgen einen quantifizierenden Ansatz, der es ermöglicht, große Mengen an Einzelergebnissen zu Gesamteffekten zu aggregieren. Meta-Analysen erlauben es weiterhin im Rahmen von Moderationsanalysen die Heterogenität der Befunde (u.a. in der Forschungsmethodik) ausdrücklich zu berücksichtigen und gesicherte Aussagen darüber zu treffen, inwiefern die Effekte in Abhängigkeit von weiteren Bedingungen variieren. Somit können Meta-Analysen gerade in der Bildungsforschung als ein geeignetes methodisches Instrument zur Erzeugung von Evidenz gesehen werden, welches zudem in der Lage ist, auch vielfach entscheidende Kontextbedingungen zu modellieren (Pant, 2014). Eine aktuelle Analyse des Forschungsstands zu effektivem MINT-Unterricht ergab 66 relevante Meta-Analysen. Diese beziehen sich größtenteils auf pädagogisch-psychologisch relevante Themenstellungen und weisen eine gute bis sehr gute Qualität auf (Knogler et al., in Vorbereitung). Sie bieten daher eine tragfähige Grundlage an validen und relevanten Befunden.

#### Adressatengerechte Kommunikation

Neben der Bestimmung und Analyse der Evidenzgrundlage betreffen weitere wichtige Entscheidungen zur Konzeption eines Clearing Houses in den Bildungswissenschaften die Formate, in welcher die Evidenz zielgruppengerecht vermittelt wird. Ein zentraler Anspruch bei der Entwicklung aller Formate ist, dass die Materialien in zweierlei Hinsicht genutzt werden können: Lehrerbildner sollten das Wissen sowohl für die Weiterentwicklung ihres eigenen Kenntnisstandes als auch direkt für

Vermittlung in Ausbildung und Lehre einsetzen können. Das aktuelle Kernprodukt des Clearing House Unterricht sind sogenannte Kurzreviews über relevante Meta-Analysen zum MINT-Unterricht. In einem Kurzreview werden auf wenigen Seiten wichtige Befunde einer Meta-Analyse in einen ersten Teil zusammengefasst. In einem zweiten Teil werden die Ergebnisse vor dem Hintergrund anerkannter Kriterien (vgl. Abelson, 1995) eingeordnet, um Nutzer bei der Einschätzung der Befunde weitergehend zu unterstützen. Die Erstellung der Kurzreviews folgt einen standardisierten und intersubjektiven Prozess, bei dem sowohl Wissenschaftlerinnen und Wissenschaftler des Clearing House Unterricht sowie Vertreterinnen und Vertreter der Lehrerbildung verschiedener Phasen beteiligt sind. Durch ein iteratives Verfahren der Produktoptimierung durch Einholen von Nutzerfeedback wird sichergestellt, dass das Produkt eine hohe Verständlichkeit, Interessanz und Relevanz für die Nutzergruppen erreicht. Aktuelle Befunde aus der Evaluation deuten darauf hin, dass dieser Anspruch auch erfüllt wird. Kurzreviews sowie alle zukünftigen Formate können über eine Online-Plattform ([www.clearinghouse-unterricht.de](http://www.clearinghouse-unterricht.de)) abgerufen und in der Lehrerbildung eingesetzt werden.

## **Die Einschätzung des Nutzens von Forschung als Teil professioneller Lehrerkompetenz**

**Christina Haberfellner**

Pädagogische Hochschule Salzburg

### Theoretischer Hintergrund

Wissenschaftlich fundierte Informationen zielgerichtet erschließen, durchdringen, kritisch reflektieren und auf konkrete Problemstellungen anwenden zu können, stellt im Sinne einer modernen evidenzbasierten Unterrichtspraxis einen wesentlichen Anspruch an (zukünftige) Lehrer/-innen dar (Altrichter & Mayr, 2004). Dementsprechend hat der Anspruch, Forschungskompetenz (i.S. von Educational Research Literacy, u.a. Shank & Brown, 2007) in der Lehrerbildung zu fördern, auch in Österreich Berücksichtigung gefunden (BMUKK & BMWF, 2010). Wird dieser Nutzen von wissenschaftlichem Wissen von den Lehramtsstudierenden erkannt, werden sie gemäß aktueller Lerntheorien motivierter sein, sich mit den geforderten Inhalten auseinanderzusetzen, was in weiterer Folge einen höheren Lerneffekt erwarten lässt (Eccles & Wigfield, 2002; Vetter & Ingrisani, 2013; Wigfield & Eccles, 2000) und dazu beitragen kann bekannte Ängste und Abwehrhaltungen (Fried, 2003; Murtonen, 2005) in diesem Zusammenhang abzubauen. Wie Vetter und Ingrisani (2013) feststellten, erkennen Lehramtsstudierende durchaus einen Nutzen in der Auseinandersetzung mit Forschungsmethodik bezogen auf ihr zukünftiges Berufsfeld. Werden diese jedoch allgemein nach ihrer Einschätzung des Nutzens von Forschung befragt, so konnte Haberfellner (2016) zeigen, dass darüber hinaus auch Aspekte wie die allgemeine Professionalisierung der Lehramtsausbildung, persönliche Karriereüberlegungen oder Forschung als Werkzeug zur Erstellung der akademischen Abschlussarbeit im Studium relevant sind (Haberfellner, 2016; Haberfellner & Fenzl, 2017).

### Fragestellung

Bei getrennter Betrachtung der verschiedenen Nutzenaspekte konnten entgegen den Erwartungen (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000) in der Untersuchung von Haberfellner (2016) keine Zusammenhänge zur objektiv gemessenen Forschungskompetenz der Studierende festgestellt werden. Im geplanten Beitrag soll nun untersucht werden, inwieweit sich Lehramtsstudierende in Abhängigkeit von ihrer Einschätzung des Nutzens von Forschung für die akademische Abschlussarbeit, ihren weiteren Karriere- und Berufsweg und ihre zukünftige Tätigkeit als Lehrer/innen in homogene Gruppen einteilen lassen und ob sich diese Gruppen in Bezug auf (1) ihre Fähigkeit, Forschungswissen zu verstehen und anzuwenden und (2) die Einschätzung der Wichtigkeit verschiedener forschungsorientierter Handlungen zur professionellen Berufsausübung unterscheiden.

### Methodik

Zur Beantwortung der Fragestellungen wird auf Fragebogen und Kompetenztestdaten zurückgegriffen, welche von 295 Studierenden aus zwei Pädagogischen Hochschulen in Österreich vorliegen. Hierbei wurde zunächst eine Clusteranalyse berechnet, um zu überprüfen, inwieweit die Studierenden in Abhängigkeit von ihren Angaben zu den drei Nutzenaspekten von Forschung in homogene Subgruppen eingeteilt werden können. Varianzanalysen verdeutlichen mögliche Unterschiede zwischen den Clustern im Hinblick auf die Fähigkeit, Forschungswissen zu verstehen und anzuwenden und ihrer Einschätzung der Wichtigkeit verschiedener forschungsorientierter Handlungen zur professionellen Berufsausübung. Alle eingesetzten Skalen und Items wurden bereits in vorhergehenden Studien validiert (Groß Ophoff, Wolf, Schladitz & Wirtz, 2017; Haberfellner, 2016; Vetter & Ingrisani, 2013) und weisen gute Skalenkennwerte auf.

### Ergebnisse

Die Ergebnisse legen nahe, dass die Studierenden in sechs homogene Gruppen aufgeteilt werden können. Hinsichtlich möglicher Gruppenunterschiede zeigt sich, dass im Hinblick auf die objektiv gemessenen Kompetenz Forschungswissen zu verstehen und anzuwenden auch bei simultaner Berücksichtigung der Nutzenaspekte keine signifikanten Differenzen feststellbar sind. Die Einstellungen der Studierenden gegenüber forschungsorientierten Handlungen im Beruf unterscheiden sich jedoch teilweise signifikant. So sehen Studierende, die man als „forschungsablehnend“ oder „bachelorarbeitsorientiert“ bezeichnen kann, im Gegensatz zu „Forschungsaffinen“ keine Qualitätsverbesserung von schulischer Arbeit, wenn Lehrer/innen über Forschungskompetenz verfügen ( $F[5, 248] = 4.902, p < 0.001$ ). Sie weisen auch der (Selbst-)Evaluation des Unterrichts ( $F[5, 208] = 5.901, p < 0.001$ ) bzw. dem Rezipieren von Forschungsliteratur zur Weiterentwicklung von Unterricht ( $F[5, 208] = 4.204, p < 0.001$ ) eine signifikant geringere Bedeutung zu.

## **Die (Un)möglichkeit die Entwicklung bildungswissenschaftlicher Forschungskompetenz im Verlauf eines Semesters empirisch abzubilden**

**Jana Groß Ophoff<sup>1</sup>, Sandra Schladitz<sup>2</sup>, Markus Wirtz<sup>1</sup>**

<sup>1</sup>Pädagogische Hochschule Freiburg, <sup>2</sup>Philipps-Universität Marburg

### Theorie

Wissenschaftlich fundierte Informationen erschließen, reflektieren und auf Problemstellungen anwenden zu können (Bildungswissenschaftliche Forschungskompetenz = BFK, z.B. Shank & Brown, 2007), ist grundlegend für die Teilhabe an der sich durch wissenschaftlichen Fortschritt ständig wandelnden Wissensgesellschaft. Da bildungswissenschaftliche Studierende (u.a. Lehramt, Erziehungswissenschaft) im späteren Berufsleben für Bildung und Erziehung künftiger Generationen verantwortlich sein werden, ist BFK grundlegend für eine kontinuierliche Professionalisierung (Jindal-Snape, Hannah, Smith, Barrow & Kerr, 2009). Im Zuge der sogenannten empirischen Wende im deutschen Bildungswesen ist daher die Forderung laut geworden, dass pädagogische Berufe dafür qualifiziert werden müssen, Forschungswissen für die Praxis zu nutzen (z.B.



Lehrerbildung, Kultusministerkonferenz, 2004). Hochschuleinrichtungen werden dafür als besonders geeignet angesehen, da sie Forschung und Lehre miteinander verbinden (Healey, 2005).

Groß Ophoff, Wolf, Schladitz und Wirtz (2017) beschreiben BFK als mehrdimensionales Konstrukt mit einem Hauptfaktor, der generische Aspekte von BFK repräsentiert, und den drei Sekundärfaktoren Informationskompetenz (IL), Statistische Kompetenz (SL) und Evidenzbasiertes Schlussfolgern (ES). Letztere spiegeln spezifische Anforderungen im Forschungszyklus wieder: IL bezieht sich auf die Fähigkeit, angemessene Forschungsfragen zu formulieren und relevante Informationen zu recherchieren (Blixrud, 2003). SL beschreibt die Fähigkeit, numerische Informationen verstehen und organisieren sowie mit verschiedenen Repräsentationen umgehen zu können (Watson & Callingham, 2003), und ES die Fähigkeit, Schlussfolgerungen untermauern und unter Bezug auf wissenschaftliche Gütekriterien kritisch prüfen zu können (Dunn, Halonen & Smith, 2008).

#### Fragestellung

Obwohl typischerweise in curricularen Modellen die Anforderung formuliert ist, dass sich BFK im Studium von der reinen Rezeption hin zur Anwendung von Forschungsinformationen und schließlich dem eigenständigen Forschen entwickelt (Willison & O'Regan, 2007), gibt es bislang nur wenig Befunde zur tatsächlichen Entwicklung. Dieser Mangel an Evidenz steht vermutlich in Zusammenhang damit, dass BFK nicht als Fachwissen, sondern als generisch-prozedurales Wissen aufgefasst wird (Wecker, Hetmanek & Fischer, 2014), das gewissermaßen als Nebeneffekt eines Universitätsstudiums erworben wird. Daher widmet sich der Vortrag der Frage, inwiefern ein Zuwachs in BFK im Verlauf eines Semesters erwartet und empirisch abgebildet werden kann.

#### Methode

Die vorgestellte Studie wurde im Sommersemester 2014 an drei deutschen Universitäten durchgeführt (t1: April/Mai 2014; t2: Juli 2014). Die Testhefte setzten sich aus Aufgaben zusammen, die im Rahmen einer umfangreichen Normierungsstudie entwickelt und empirisch überprüft wurden (Groß Ophoff, Schladitz, Lohrmann & Wirtz, 2014; Groß Ophoff et al., 2017). So konnten für die Studienteilnahme zu t1 753 Studierende der Bildungswissenschaften gewonnen werden, von denen jedoch nur 104 Personen zu t2 wieder erreicht wurden.

Die Daten wurden mittels mehrdimensionaler Item-Response-Modell ausgewertet (R-Paket TAM, Kiefer, Robitzsch & Wu, 2016). Um querschnittliche Unterschiede zwischen den Messzeitpunkten zu untersuchen, wurden zunächst mehrdimensionale latente Regressionsmodelle analysiert. Anschließend wurden pro Person und BFK-Dimension Plausible Values geschätzt und über t-Tests für abhängige Stichproben analysiert.

#### Ergebnisse

Weder in den Längs- noch Querschnittsanalysen ergaben sich signifikante Unterschiede. Einzig in SL zeigten die Studierenden zu t2 eine um .20 logits signifikant niedrigere Leistung. Dieser leichte Rückgang ( $d = -.32$ ) steht vermutlich in Zusammenhang mit einem systematischen Wegbleiben leistungsstärkerer Studierender im Semesterverlauf (vgl. Groß Ophoff et al., 2015). Diese Annahme wird unterstützt durch die signifikanten negativen Korrelationen zwischen der BFK und der selbst berichteten Teilnahmehäufigkeit an forschungsbezogenen Seminaren (z.B. Tutorat für Forschungsmethoden, B.A. Erziehungswissenschaft:  $r_{min} = -.33$  to  $r_{max} = -.44$ ). Demnach scheinen schwächere Studierende in den Lehrveranstaltungen zu verbleiben, aber nicht davon zu profitieren. Während laut Healey (2005) in der Hochschullehre nach wie vor dozentenorientierte Lehrangebote dominieren, konnte in forschungsbasierten Seminarangeboten oder Studienprogrammen eine Entwicklung von BFK empirisch aufgezeigt werden (Groß Ophoff et al., 2015; Kittel, Rollett & Groß Ophoff, in Druck). Dies verdeutlicht einerseits, dass es nur bedingt aussagekräftig ist, die Entwicklung von BFK zu modellieren, ohne die curriculare Umsetzung zu berücksichtigen, und dass andererseits Bedarf an systematischen Interventionsstudien besteht.

## **Die Förderung studentischer Forschungskompetenzen im Lehramtsstudium - Evaluation zweier Formate zur Einbindung von Forschung in die Lehre**

**Franziska Böttcher, Felicitas Thiel**

Freie Universität Berlin

#### Theoretischer Hintergrund

Evidenzbasierte Entscheidungen in Unterrichtsentwicklung erfordern grundlegende Forschungskompetenzen auf Seiten von Lehrerinnen und Lehrern (KMK, 2004). Die Förderung studentischer Forschungskompetenzen im Lehramtsstudium erfährt daher eine steigende Aufmerksamkeit.

Forschungskompetenzen können im Rahmen forschungsorientierter Lehre gezielt gefördert werden. Forschungsorientierte Lehre bietet die Möglichkeit, über die Einbindung von Forschung in die Lehre den Research-Teaching-Nexus zu stärken (Brew, 2006; Griffith, 2004; Healey, 2005). Diese Einbindung kann entweder über die Integration aktueller Forschungsbefunde in die Lehre geschehen (research-led teaching; Healey, 2005) oder durch die Bearbeitung von Forschungsfragen und die aktive Anwendung von Forschungsmethoden (research-oriented teaching; Healey, 2005).

#### Fragestellung

In zwei Seminartypen wurden diese zwei unterschiedlichen Ansätze zur Einbindung von Forschung in die Lehre im Lehramtsmasterstudium umgesetzt. In einem Seminartyp wurde Forschung im Sinne des research-led teaching in die Lehre integriert. Dabei haben die Studierenden aktuelle Forschungsbefunde zum Thema Lehren und Lernen rezipiert, diskutiert und interpretiert. Im zweiten Seminartyp wurde Forschung im Sinne des research-oriented teaching in einem Lernforschungsprojekt eingebunden. Studierende haben in diesem Seminartyp im Laufe des Praxissemesters selbst zum Thema Lehren und Lernen eine eigene Forschungsfrage bearbeitet. Die Lernforschungsprojekte wurden in einem Seminar vorbereitet und begleitet. Im Rahmen des begleiteten Lernforschungsprojekts haben die Studierenden sowohl Wissen über den Forschungsprozess erworben als auch alle Schritte des Forschungsprozesses bei der Bearbeitung ihre Fragestellung selbst durchlaufen. Ziel der Studie ist es, die Wirksamkeit der Einbindung von Forschung im Rahmen der beiden Seminarformate im Hinblick auf den studentischen Kompetenzerwerb zu evaluieren.

#### Methode

In einem quasi-experimentellen Design wurden dazu Daten zu zwei Messzeitpunkten erhoben: zu Beginn und zum Ende des Semesters. Eingesetzt wurde der Fragebogen zur Erfassung studentischer Forschungskompetenzen (F-Komp; Böttcher & Thiel, 2017), der auf dem RMRK-W-Modell basiert (Thiel & Böttcher, 2014). In den Dimensionen Recherche-, Methoden-, Reflexions- und Kommunikationskompetenzen sowie Fachliches Wissen schätzten die Studierenden ihre Forschungskompetenzen selbst ein. Zum zweiten Messzeitpunkt wurde der F-Komp zusätzlich in einer retrospektiven Version

(then-test) eingesetzt, sodass die Studierenden ihre Kompetenzen vor der Lehrveranstaltung sowie zum Ende der Lehrveranstaltung einschätzten. Um die subjektive Einschätzung der Studierenden durch ein objektives Maß zu ergänzen, wurde als weiteres Instrument ein Kompetenztest eingesetzt, der bildungswissenschaftliche Forschungskompetenz in den Dimensionen Information Literacy, Statistical Literacy und Evidence-based Reasoning erfasst (Groß Opfhoff et al, 2014).

Es liegen Daten von insgesamt 95 Studierenden vor, die zu beiden Messzeitpunkten an der Erhebung teilgenommen haben. Davon befanden sich  $n = 74$  Studierende im Lernforschungsprojekt und  $n = 21$  im research-led-Seminar. Den Kompetenztest haben die Studierenden der research-led-Seminare sowie eine Teilgruppe der Studierenden in den Lernforschungsprojekten ( $n = 36$ ) ausgefüllt.

#### Ergebnisse

Erste Analysen deuten darauf hin, dass die Studierenden der Lernforschungsprojekte in der Selbsteinschätzung von Kompetenzen mit dem F-Komp sowohl im Längsschnitt als auch in der Retrospektive einen signifikanten Kompetenzzuwachs auf fast allen Dimensionen (mit Ausnahme der Reflexionskompetenz im Längsschnitt) wahrgenommen haben. Die Studierenden der research-led-Seminare haben in der Selbsteinschätzung von Kompetenzen mit dem F-Komp retrospektiv auf allen Dimensionen einen signifikanten Kompetenzzuwachs wahrgenommen. Im Längsschnitt konnte dagegen auf keiner Dimension ein Kompetenzzuwachs erzielt werden.

Beim Kompetenztest hatten die Studierenden des Lernforschungsprojektes einen signifikanten Anstieg der Lösungshäufigkeit in den Bereichen Information Literacy und Statistical Literacy. Bei den Studierenden der research-led-Seminare zeigte sich ein signifikanter Anstieg lediglich im Bereich der Information Literacy.

## C20

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Kollegienhaus H120

### Innovations for Educational Assessment by Natural Language Processing

Chair(s): **Fabian Zehner** (Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland), **Frank Goldhammer** (Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Zentrum für internationale Vergleichsstudien (ZIB), Deutschland)

DiskutantIn(nen): **Dirk Ifenthaler** (Universität Mannheim, Curtin University)

Natural language plays a crucial part in educational assessment. Both for practitioners and researchers, language is the most common medium for transferring contents that are to be instructed, learned, discussed, or tested. Thus recently, several innovative educational applications of natural language processing (NLP) techniques have been proposed. These are capable of improving professionalization in educational research and practice in various ways. For example in educational research, the volume of unstructured text data can constitute a large obstacle when thousands or even millions of text responses need to be scored. Modern technologies can help in such cases as the manual scoring can be supported or replaced by automatic scoring systems (cf. Smarter Balanced Assessment Consortium, 2014, who analyzed the performance of various automatic scoring systems). In educational practice, technologies employing NLP can open up new possibilities just as well. For example, classroom engagement systems can help teachers to receive responses to open-ended questions from all students instead of being urged to select only one student. The system groups the students' responses, so that the teacher can review the class's (mis)conceptions and intervene in real time (e.g., Paiva et al., 2014). These examples show the promising potentials of applying NLP techniques in education.

Therefore, the interdisciplinary symposium brings together computer linguists and educational researchers in four contributions, all of them presenting novel educational applications of NLP techniques for (semi-)automatically coding student texts as well as automatic item generation of language assessment items or reporting on the use of such. (I) Oliver Meyer, Maleika Krüger, and Stephan Keller present a study in which they used e-rater for automatically analyzing student essays. They discuss the agreement between the human and computer scores as well as borderline cases that are not properly coded by the automatic system. (II) Fabian Zehner, Christine Sälzer, and Frank Goldhammer present a first insight into an international cooperation project on automatically coding students' short text responses from PISA across multiple languages. The project tests the feasibility and performance of automatic scoring across different test languages and develops a methodology to directly compare text responses from different countries. (III) Magdalena Wolska introduces a system that supports human scorers of text responses by grouping responses into sets of similar responses. Finally, (IV) Torsten Zesch proposes a new response format and an automatic item generator for the assessment of language learning. Similar to C-tests, the respondents are presented with a gap in a sentence that needs to be filled in. But for disambiguating the gaps, an item consists of four sentences with the same gap. The system automatically selects these sentences.

The symposium attempts to provide a look behind the doors that technologies can open for educational research and practice. It will also critically discuss their limitations and potential risks. For this, the well-recognized educational technologist Prof. Dr. Dirk Ifenthaler will discuss the four contributions.

#### Beiträge des Symposiums

### Automated Essay Evaluation of Complex Essay Writing Skills at Upper Secondary Level – Results from the MEWS Study

**Oliver Meyer, Maleika Krüger, Stefan Keller**

PH FHNW, Institut für Bildungswissenschaften, University of Basel

The MEWS project (Measuring English Writing at Secondary Level; SNF-DFG Lead Agency project) is one of the first large-scale studies for English essay writing in Germany and Switzerland at upper secondary level, and one of the first studies in Europe to use „Automated Essay Evaluation“ (AEE).

The ability to write longer, coherent and persuasive English texts is increasingly important in tertiary education, the labor market and everyday life (Keller, 2013). Therefore, essay writing is an important part of the German and Swiss upper secondary English curriculum. Despite its increasing importance for learners' educational success (Koeller, Knigge & Tesch, 2010), there is still a lack of empirical research on advanced students' English essay writing skills. One reason is the considerable organizational and monetary challenge researchers face as "writing competence" is a highly complex construct to operationalize. The assessment of complex L2 essays requires intensive and expensive rater training. Moreover, even if raters are trained thoroughly, rater effects (e.g. lenience, fatigue) and text effects (Porsch & Koeller, 2010) might still influence the rating process.

Automated essay evaluation (AEE) is one way to minimize rater and text effects (Ramineni, Trapani, Williamson, Davey & Bridgeman, 2012). The engine uses Natural Language Processing (NLP) and linear regression models to calculate a holistic score from 40 features such as grammar, language, mechanics, style, organization, use of pronouns, word collocations and lexical complexity (Ramineni et al., 2012).

By employing AEE in addition to human scoring, MEWS breaks new ground concerning scale and scoring methods. The study recruited  $n \approx 1700$  students in Switzerland and  $n \approx 800$  students in Germany from upper secondary schools willing to participate in the study. Students were tested twice (t1 and t2) during their penultimate school year and responded to two types of writing tasks at each time point. The study also tested students' reading and listening skills, as well as their general cognitive skills. In addition, students provided answers to extensive questionnaires concerning family background, personal characteristics (e.g. self-concept), classroom activities and out-of-school contact with English.

The AEE-based analyses were carried out in co-operation with Educational Testing Service (ETS) in Princeton, USA, who provided the scoring engine and a selection of independent writing tasks (argumentative essays) and integrated writing tasks (juxtaposing and summarizing listening and reading passages). Essays were scored by 36 trained human raters (2 raters per essay) on a 6-level holistic scale and an AEE-model was calibrated and run for the specific sample of learners and writing prompts used in the study.

First results show satisfying inter-rater agreement between human and machine rating for the independent essay by using a general calculation model (Quadratic weighted Kappa (QWK)= .67). However, the use of a generic model resulted in less

satisfying inter-rater agreement for the integrated task (QWK = .48), which called for further investigation. The use of prompt-specific models ultimately resulted in satisfying agreement rates (QWK = .60 and QWK = .74) for the integrated tasks as well.

However, the use of prompt-specific models poses a challenge when using AEE in high stake situations. Students asked to complete tasks with inferior performing models (e.g. lower human-machine agreement) would be in a disadvantaged position by getting lower-quality machine scores than others.

What is more, first analysis of individual cases with a high discrepancy between human and machine rating reveal AEE's weaknesses in capturing edge cases (e.g. students who misunderstood the prompt) and responses with inappropriate undertones.

In our presentation we will show in-depth analyses of the human and machine scores and talk about advantages and disadvantages of AEE in large-scale and high-stakes assessments.

### **ReCo-Multi: Automatically Coding PISA Short Text Responses across Multiple Languages**

**Fabian Zehner<sup>1</sup>, Christine Sälzer<sup>2</sup>, Frank Goldhammer<sup>3</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland, <sup>2</sup>Technische Universität München, Zentrum für internationale Vergleichsstudien (ZIB), <sup>3</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Zentrum für internationale Vergleichsstudien (ZIB), Deutschland

The presentation gives first insights into the project ReCo-Multi (Automatic Response Coding across Multiple Languages). The project centers around the automatic coding of short text responses to tests in different languages. In automatic coding, computers automatically categorize responses, for example as correct or incorrect.

The automatic coding of short text responses overcomes some well-known problems that are associated with human coders. Human coders (cf. Bejar, 2012) bring in subjective perspectives, their abilities and personality traits (e.g., experience, stamina, lenience), the need for consensus building measures, and, thus, cost quite some time and money. When trained software supports or complements these coding processes, the data quality can be improved with respect to multiple aspects such as validity, reliability, objectivity, and fairness. While several research groups have shown that computers can attain sufficient performance levels in scoring (cf. Burrows, Gurevych, & Stein, 2014), the next step is to make software scalable to multiple languages. Particularly in international large-scale assessments, such as the Programme for International Student Assessment (PISA), large volumes of text responses require manual scoring. Hitherto, software is most often constrained to English (e.g., Leacock & Chodorow, 2003) or one other language (e.g., Zehner, Sälzer, & Goldhammer, 2016). Making automatic coding accessible to such studies by supporting multiple languages at a time would, thus, be particularly beneficial, also because they usually contain recurring items across assessment cycles.

ReCo-Multi pursues two main goals. (I) The scope of the computer program ReCo (Zehner et al., 2016), which was only evaluated on German data, will be extended to multiple languages (i.a., English, Portuguese). It uses Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) and clustering for identifying semantically homogeneous response groups, so-called response types. For this, a text corpus (parts of the Wikipedia) is processed for computing something like a dictionary on semantic similarity of words, called semantic vector space model. The semantic information in the responses is then analyzed by machine learning techniques for building a model on which responses need to be assigned to which code. (II) Since this procedure is constrained to language-specific semantic spaces, the project will develop a methodology for comparing response types across languages, that is, across semantic spaces. The semantic spaces will be mapped to each other using spatial warping based on document and term equivalence provided in the text corpus.

In the project, eighteen PISA National Centers have joined a cooperation sharing their text response data to reading questions from PISA 2015. This way, a unique data pool evolves containing about 640.000 responses in fourteen test languages. All responses have already been scored by human coders.

The presentation will illustrate the methodology and elucidate first substantive findings. As the data transfers are currently coming in, no results can be provided in this abstract. But the presentation will compare text responses (i) in German from Germany (n = 6,504 students), Austria (n = 7,007), and Luxembourg (n = 3,608), (ii) in English from Ireland (n expected to be about 5,000) and Luxembourg (n = 209), as well as (iii) in Portuguese from Portugal (n = 7,325). This way, the responses can be compared, partly disentangling language and the cultural context. The presentation will show first performance comparisons across languages and countries (human-computer agreement:  $\kappa^2$ ) as well as exploratory findings in surface level comparisons (e.g., word counts) differentiated by relevant variables (e.g., gender). These analyses will also serve as indications whether a test language needs to be approached differently in different countries. Finally, the presentation concludes with critically discussing the limitations of the automated approach and how it might be employed in educational (large-scale) assessments.

### **Computer-assisted scoring of short constructed responses in low-stakes assessment: Towards supporting scoring efficiency and consistency**

**Magdalena Wolska**

Eberhard Karls Universität Tübingen

With large-scale international assessments, such as PISA or PIRLS, gradually, but systematically evolving towards computer-based administration, innovative computational technologies that would facilitate efficient and reliable coding of large amounts of test-takers' responses are becoming naturally called for. Computer-based support for coding constructed response (CR) items, which require human manual coding due to their open-ended nature, could especially contribute to streamlining the turnaround of results.

In this presentation we discuss our recent experiments on using Natural Language Processing (NLP) techniques in a web-based system for computer-based scoring of short constructed responses, that is responses to items which elicit brief textual production, rather than structured discourse. Our goal in this work is not to build systems to score such responses automatically, but rather only to facilitate the manual scoring task: make it more efficient and more consistent, while relying on human interaction with the entire response set.

Our data stems from several cohorts of learners of German as a Foreign Language taking a placement test administered as part of registration for German language courses at a university. Short CRs have been used in a listening comprehension section of the placement test and have been scored by language teachers administering the course placement. Both the test taking and the scoring have been conducted via a web-based platform. Data samples for experiments were drawn from 23 items with at least 100 distinct responses.

The experiments are based on a computer-assisted scoring system which uses language processing techniques (text normalization and vector-based response representation, among others) and data mining methods (efficient sampling of the data space, iterative classifier training, and clustering) to present sets of "similar" responses for scoring and to pre-score responses in case of highly predictive scores; similarity is operationalized in terms of vector distance. Our overarching research questions concern the effect of the mode of response presentation on the scoring time and scoring consistency: Does NLP-based grouping of responses for scoring reduce the time on task? Does it improve scoring reliability? For the purpose of present experiments, the former was operationalized as total response set scoring time per item whereas the latter as intrarater agreement over three repeat rounds of scoring sessions on the same sample.

We show that gains are achieved on both variables, however, not of the same size for all items: For some items, statistical difference is found between time-on-task in the grouped responses condition vs. non-grouped mode and between percentage agreement and intraclass correlation in the grouped vs. non-grouped condition. We discuss qualitative differences between the response sets on which larger and smaller differences have been observed and suggest improvements to the response presentation modality. While the present experiments have been conducted in the low-stakes setting of placement testing, the methods we use are scalable to larger tests that elicit textual answers of analogous character.

## **Gap-fill Bundles – A New Paradigm for Automatically Generated Language Assessment Items**

**Torsten Zesch**

Language Technology Lab, University of Duisburg-Essen

Automatically generating high-quality language assessment items is a challenging problem. Although a wide range of methods for creating gap-fill items (Horsmann & Zesch, 2014) or multiple-choice items (Sakaguchi, Arase, & Komachi, 2013; Sumita, Sugaya, & Yamamoto, 2005) has been explored, the resulting items are not of the same quality as manually created ones. For example, automatically constructed multiple-choice items are usually either too easy (when the algorithm generates nonsensical distractors) or invalid (when the algorithm uses alternative, correct solutions as distractors). Consequently, generating high-quality language assessment items is still a labor-intensive manual task.

We present gap-fill bundles (Wojatzki, Melamud, & Zesch, 2016), a new paradigm for language assessment items that—with the help of Natural Language Processing—can be automatically generated and at the same time automatically scored. In contrast to classical multiple-choice tests, a gap-fill bundle consists of four sentences where the same word is removed instead of one sentence with a blanked word and four alternative solutions. A gap-fill bundle is created by taking a source sentence with a blanked word and then searching a large space of candidate sentences containing the same word. By selecting sentences where the word appears in a context that rules out some of the alternative solutions of the first sentence, the bundle in its entirety has only one plausible solution that can be easily checked automatically. Alternative solutions provided by test takers are highly unlikely to be correct. The difficulty of the bundles can be adapted by varying the difference in probability between the correct solution and the next best alternative solution.

A major advantage of gap-fill bundles is that it is a production task while multiple-choice is a recognition task. As gap-fill bundles are a relatively new development, there are many open questions regarding among others the nature of the competence being tested, whether they complement or replace other test formats, or how adaptable the paradigm actually is.

In a recent study, Meyer, Wojatzki, and Zesch (2016) assessed the language skills of university students ( $n = 75$ , 52 female, 1 non specified gender) using both a C-test and a set of generated bundles. They found that participants' C-test scores are only moderately correlated ( $r = .48$ ) with bundle scores.

In summary, gap-fill bundles are an interesting new test paradigm, which might be utilized in empirical educational research as well as in language training and testing. Before gap-fill bundles can be used productively, more interdisciplinary research is required. However, we see a huge potential especially for areas where the current test formats are considered insufficient or where automatic scoring should be used. The presentation will discuss further potential applications in and their implications for educational assessment and practice.

## C21

Zeit: Freitag, 16.02.2018: 9:00 - 10:45 · Ort: Bernoullianum 30, H148

### Professionelles Handeln im Bereich Mathematik als Herausforderung für frühpädagogische Fachkräfte

Chair(s): **Lars Jenßen** (Freie Universität Berlin, Deutschland)

DiskutantIn(nen): **Simone Dunekacke** (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)

Die OECD (2011) beschreibt für Deutschland, dass Kindertagesstätten zunehmend eine Bildungsfunktion erfüllen müssen. Diese Entwicklung ist zum einen auf den sogenannten „PISA-Schock“ und zum anderen auf vermehrte Forschungsergebnisse zur Bedeutsamkeit einer frühen Bildung zurückzuführen (Drieschner, 2010). Neben konzeptioneller Weiterentwicklung früher Bildung steht die Professionalisierung des frühpädagogischen Personals seitdem im Fokus (vgl. von Balluseck, 2008). Trotz einiger Forschungsprojekte in dem Bereich (z.B. Kucharz et al., 2014), muss jedoch festgestellt werden, dass nach wie vor drängender Forschungsbedarf zur Professionalisierung in der Frühpädagogik besteht (Faas, 2013; Köller, 2016). Blömeke, Gustafsson und Shavelson (2015) beschreiben in ihrem Modell verschiedene Facetten professioneller Kompetenz, die eine Person haben muss, um kompetent in einer Profession tätig zu sein. Neben kognitiven und affektiv-motivationalen Facetten sind demnach auch situationsspezifische Fertigkeiten nötig (z.B. Beobachtung oder Handlungsplanung) entscheidend. In diesen verdichten sich das Wissen und affektiv-motivationale Dispositionen und sie bilden die Grundlage für ein beobachtbares kompetentes Verhalten (Handeln).

International besteht mittlerweile eine Vielzahl an empirischen Untersuchungen, die die Bedeutung der frühpädagogischen Fachkraft im Kontext früher mathematischer Bildung stärken (z.B. Pianta et al., 2014; Clements & Sarama, 2011; Bogard et al., 2008). Im europäischen Raum, und vor allem in Deutschland, muss aber festgestellt werden, dass die Erforschung von Kompetenzen von frühpädagogischen Fachkräften im Bereich Mathematik ein wesentliches Forschungsdesiderat darstellt (European Commission, EACEA, Eurydice und Eurostat, 2014; Köller, 2016).

Die Beiträge des Symposiums stammen aus dem Projekt Pro-KomMa (Jenßen et al., 2016). Das Projekt untersucht vor dem theoretischen Hintergrund die Multidimensionalität professioneller Kompetenz von frühpädagogischen Fachkräften im Bereich Mathematik, indem es das Zusammenspiel der einzelnen Facetten fokussiert. Ziel ist dabei ein besseres Verständnis von professionellem Handeln von Fachkräften in der Kindertagesstätte in einer Domäne zu bekommen, die von den Fachkräften immer wieder als Herausforderung wahrgenommen wird (Levin, Meyer-Siever & Gläser, 2015). Erkenntnisse der Lehrerbildungsforschung können in diesem Fall nicht übertragen werden, sondern lediglich Ideen liefern, da sich Schule und Kindertagesstätte formal und strukturell deutlich unterscheiden (Schuler, Pelzer, Wittkowski & Wittmann, 2015).

Das Symposium fokussiert entsprechend dem Thema der GEBF-Tagung vor allem performanznahe Konstrukte wie die Wahrnehmung und Interpretation mathematischer Inhalte, die Beobachtung kindlichen Verhaltens im Kontext mathematischer Bildung und das professionelle Handeln der frühpädagogischen Fachkraft im Bereich Mathematik.

Die Beiträge nähern sich dabei der Frage, welchen Anforderungen und Herausforderungen frühpädagogische Fachkräfte in ihrem Handeln im Bereich Mathematik begegnen müssen und inwieweit sie dazu in der Lage sind, diese zu bewältigen. In einem Großteil der Beiträge wird dabei auch das outcome, die mathematische Leistung der Kinder, in die Modellierungen einbezogen. Ziel ist dabei eine empirische Überprüfung grundlegender Annahmen der frühen mathematischen Bildung.

Im ersten Beitrag wird der Einfluss von Mathe-Angst auf zentrale situationsspezifische Fertigkeiten wie die Wahrnehmung und die Interpretation mathematischen Handelns von Kindern untersucht. Im zweiten Beitrag wird die Urteilsgenauigkeit frühpädagogischer Fachkräfte bezüglich der Leistung der Kinder fokussiert. In dem dritten Beitrag werden mathematikspezifische Interaktionsmodi zwischen Fachkraft und Kind aus der Kita-Praxis qualitativ beleuchtet. Der vierte Beitrag rundet das Symposium ab, indem ein standardisiertes Verfahren zur Erfassung zentraler Verhaltensindikatoren zur Anrechnungsqualität im Bereich Mathematik vorgestellt wird und hinsichtlich seiner Nützlichkeit diskutiert wird.

#### Beiträge des Symposiums

### Der Einfluss von Mathe-Angst auf situationsspezifische Fertigkeiten im Bereich Mathematik bei frühpädagogischen Fachkräften

**Lars Jenßen<sup>1</sup>, Julia Bruns<sup>2</sup>, Lars Eichen<sup>3</sup>, Katja Eilerts<sup>4</sup>, Michael Eid<sup>1</sup>, Sigrid Blömeke<sup>5</sup>**

<sup>1</sup>Freie Universität Berlin, <sup>2</sup>Universität Osnabrück, <sup>3</sup>Universität Graz, <sup>4</sup>Humboldt-Universität Berlin, <sup>5</sup>Centre for Educational Measurement at the University of Oslo (CEMO)

#### Theoretischer Hintergrund

Situationsspezifische Fertigkeiten wie die Wahrnehmung und Interpretation der Leistungen der Kinder gehören zu den wesentlichen Aufgaben einer frühpädagogischen Fachkraft (Walter-Laager, Pfiffner & Schwarz, 2011) und können demnach als eine alltägliche Herausforderung für frühpädagogische Fachkräfte verstanden werden. Sowohl national als auch international wird der Erforschung dieser performanznahen Konstrukte in der Elementarpädagogik, insbesondere im Bereich Mathematik, eine zunehmend größere Bedeutung beigemessen (z.B. Dunekacke, Jenßen, Eilerts & Blömeke, 2015; Kilday, Kinzie, Mashburn & Whittaker, 2012). Es wird angenommen, dass situationsspezifischen Fertigkeiten eine Mediatorrolle in der Vermittlung zwischen professionellem Wissen und kompetentem Handeln zukommt (Blömeke, Gustafsson & Shavelson, 2015). Des Weiteren werden vor allem affektiv-motivationale Facetten professioneller Kompetenz in ihrer Bedeutung neben Wissensbeständen diskutiert (Lee & Ginsburg, 2007; Thiel, 2010; Dunekacke et al., 2015). Studien zeigen dabei, dass Mathe-Angst ein wesentliches Phänomen bei frühpädagogischen Fachkräften darstellt (Gresham, 2007; Jenßen, Dunekacke, Eid & Blömeke, 2015). Forschungsergebnisse deuten darauf hin, dass ein bedeutsamer negativer Zusammenhang zwischen Mathe-Angst und dem mathematischen Fachwissen und auch dem mathematikdidaktischen Wissen bei angehenden frühpädagogischen Fachkräften besteht (Jenßen et al., 2015; Thiel & Jenßen, eingereicht). Zudem gibt es Hinweise, dass das mathematische Fachwissen und das mathematikdidaktische Wissen situationsspezifische Fertigkeiten wie die mathematikbezogene Situationswahrnehmung vorhersagt (Dunekacke et al., 2015). Ein direkter Zusammenhang von Mathe-Angst und situationsspezifischen Fertigkeiten bei frühpädagogischen Fachkräften in der Praxis wurde bisher jedoch noch nicht untersucht.

#### Fragestellung

Der vorliegende Beitrag fokussiert die Frage, ob Mathe-Angst situationsspezifische Fertigkeiten wie die mathematikbezogene Situationswahrnehmung und die Interpretation von mathematischen Äußerungen von Kindern vorhersagt.

## Methode

Um die Mathe-Angst zu erfassen, wurde die Mathematics Anxiety Scale in der revidierten Fassung (Bai et al., 2009) eingesetzt, welche anhand von 14 Items Aussagen über die affektive und die kognitive Facette der Mathe-Angst erlaubt. Hinweise für Reliabilität (Cronbach's Alpha=0,91) und Validität des Fragebogens liegen vor (Bai, 2011; Jenßen et al., 2015). Um die mathematikbezogene Situationswahrnehmung und die Interpretation zu erfassen, wurden zwei videobasierte Tests eingesetzt (VimaS\_num, Eichen & Bruns, 2015). Die Tests bestehen aus vier bzw. fünf Items im MC-Format und zeigen eine gute Reliabilität (Cronbach's Alpha=0,97 für beide Tests). Die Inhaltsvalidität der Tests wurde anhand eines Expertenreviews sichergestellt. In den Videoclips werden den Fachkräften Ausschnitte aus der alltäglichen Praxis einer Kindertagesstätte mit mathematischem Bezug präsentiert. Im ersten Test werden die Fachkräfte aufgefordert einzuschätzen, welche mathematischen Aspekte in dem Verhalten der Kinder zu erkennen sind. Im zweiten Test sollen die Fachkräfte die Handlungen und Äußerungen der Kinder in Bezug zu dem mathematischen Lernprozess des Kindes einschätzen. Bisher wurden n=101 Fachkräfte online befragt und getestet. Das vollständige Datenset mit etwa 200 Fachkräften wird im Dezember 2017 vorliegen.

## Ergebnisse

Die bisher vorliegenden Daten wurden mittels Strukturgleichungsmodellen überprüft. Im komplexen Modell wurde für jedes Konstrukt jeweils ein latenter Faktor modelliert. Zwischen Mathe-Angst und den situationsspezifischen Fertigkeiten wurde eine regressive Beziehung angenommen, wobei zwischen den Fertigkeiten eine korrelative Beziehung zugelassen wurde. Der Model Fit kann als sehr gut eingeschätzt werden ( $\chi^2(52) = 499.51$ ,  $p = .713$ , RMSEA=0.00 [0.00; 0.04], CFI = 1.00). Situationswahrnehmung und Interpretation zeigten einen starken positiven korrelativen Zusammenhang ( $r=.97$ ), wobei Mathe-Angst erwartungskonform einen negativen Effekt auf beide Faktoren zeigte ( $\beta_1=-.54$  für Wahrnehmung und  $\beta_2=-.51$  für Interpretation). Der Effekt ist dabei deutlich stärker ausgeprägt als er in vergleichbaren wissenschaftlichen Testungen zu finden ist (vgl. Ma, 1999). Konkurrierende Modelle, beispielsweise mit einem gemeinsamen Faktor für die situationsspezifischen Faktoren, führten zu vergleichbaren Ergebnissen. Die Ergebnisse sprechen erstmals für den Einfluss von Mathe-Angst auf performanznahe Konstrukte und belegen somit auch die praktische Bedeutsamkeit von Mathe-Angst bei frühpädagogischen Fachkräften. Die Ergebnisse stützen zudem die These, dass frühpädagogische Fachkräfte Mathe-Angst bewältigen müssen (Jenßen et al., 2015), um den Herausforderungen früher mathematischer Bildung in der Kindertagesstätte effektiv begegnen zu können.

## **Können frühpädagogische Fachkräfte die mathematischen Fähigkeiten von Kindergartenkindern angemessen einschätzen? – Anhaltspunkte und Folgen für die individuelle Leistungsentwicklung**

**Markus Szczesny<sup>1</sup>, Lars Jenßen<sup>2</sup>, Katja Eilerts<sup>1</sup>, Thomas Koinzer<sup>1</sup>, Sigrid Blömeke<sup>3</sup>**

<sup>1</sup>Humboldt-Universität zu Berlin, <sup>2</sup>Freie Universität Berlin, <sup>3</sup>Centre for Educational Measurement at the University of Oslo (CEMO)

## Theoretischer Hintergrund

Die Einschätzung und Beurteilung der Fähigkeiten und des Vermögens des Lernenden durch den Lehrenden ist der (wenigstens implizite) Ausgangspunkt jedweder pädagogischen Interaktion oder Intervention. Dabei zeigt sich auf internationaler Ebene insbesondere für den Bereich der Mathematik, dass frühpädagogische Fachkräfte sich nicht nur vergleichsweise selten und ungern mathematischen Themen und Aktivitäten widmen (National Research Council, 2009; Clements & Sarama, 2007). Darüber hinaus unterschätzen frühpädagogische Fachkräfte die mathematischen Fähigkeiten der von ihnen betreuten Kinder systematisch (Heuvel-Panhuizen, 1990; Kilday et al., 2012). Auch wenn die zitierten Autoren und Autorinnen davon ausgehen, dass die Unangemessenheit der Einschätzung mathematischer Fähigkeiten einen suboptimalen Einfluss auf die individuelle Leistungsentwicklung von Kindergartenkindern im Bereich Mathematik ausübt, fehlt es bislang an Studien, die diesen Zusammenhang direkt untersuchen.

## Fragestellung

Untersucht wird zum einen, ob frühpädagogische Fachkräfte die mathematischen Fähigkeiten der von ihnen betreuten Kinder angemessen einschätzen. Zum anderen wird der Frage nachgegangen, welchen Einfluss die Einschätzung der mathematischen Fähigkeiten von Kindergartenkindern auf die Leistungsentwicklung von Kindergartenkindern im Bereich Mathematik ausübt.

## Methode

Im Rahmen der Pro-KomMa-Studie werden bis Anfang 2018 N≈50 frühpädagogische Fachkräfte und N>300 von den Fachkräften betreute Kinder in über 40 Kindertagesstätten in Berlin und Brandenburg untersucht. Zur Erfassung der Beurteilung der mathematischen Fähigkeiten der von frühpädagogischen Fachkräften betreuten Kinder wird das Kinderdiagnostetool KiDiT© verwendet (Walter-Laager et al., 2012), in dem die mathematischen Fähigkeiten von Kindergartenkindern durch 25 Items abgebildet werden. Das in KiDiT© eingesetzte Instrument hat sich in mehreren Studien als reliabel (Reliabilitäten zwischen .81 und .98) und valide erwiesen (Walter-Laager et al., 2012). In einem 2-monatigen Turnus werden die mathematischen Fähigkeiten der durch die frühpädagogischen Fachkräfte betreuten Kinder an fünf Messzeitpunkten von den betreuenden Fachkräften eingeschätzt. Die mathematischen Fähigkeiten der Kindergartenkinder werden mithilfe des MBK-0 (Mathematische Basiskompetenzen im Kindergartenalter; Seeger et al., 2014) erfasst. Der MBK-0 weist eine interne Konsistenz von Cronbach's Alpha=.87 und eine Retest-Reliabilität von  $r_{tt}=.88$  (drei Monate) sowie eine hohe prognostische Validität für die Mathematikleistung am Ende der dritten Klasse auf. Die Testung der mathematischen Fähigkeiten der Kindergartenkinder wird in einem Abstand von etwa sechs Monaten an zwei Messzeitpunkten durchgeführt. Zur Kinder-Testung werden geschulte Studierende des Grundschullehramts eingesetzt.

## Ergebnisse

Da ein belastbarer Datensatz erst Januar 2018 vorliegen wird, konnten bisher nur Varianzanalysen der Einschätzung der mathematischen Fähigkeiten der Kindergartenkinder zu den ersten beiden Messzeitpunkten durchgeführt werden. Dabei zeigt sich, dass sich etwa 40% der Varianz in den Fähigkeitseinschätzungen auf der Ebene der frühpädagogischen Fachkräfte manifestiert. Ein solches Ergebnis weist darauf hin, dass die untersuchten frühpädagogischen Fachkräfte subjektive (und das heißt: möglicherweise unangemessene) Maßstäbe zur Bewertung der mathematischen Fähigkeiten der von ihnen betreuten Kinder ansetzen.

Beabsichtigt ist, die Stabilitäten der Einschätzungen der mathematischen Fähigkeiten durch die frühpädagogischen Fachkräfte und der mathematischen Fähigkeiten der betreuten Kinder sowie die Zusammenhänge zwischen beiden Instrumenten in Cross-Lagged-Panel-Modellen zu analysieren.

Erwartet wird zum einen, dass die Korrelationen zwischen den Einschätzungen der mathematischen Fähigkeiten durch die frühpädagogischen Fachkräfte und den mathematischen Fähigkeiten der betreuten Kinder über die Messzeitpunkte zunehmen. D.h., dass sich die Angemessenheit der Einschätzungen der frühpädagogischen Fachkräfte über die Zeit verbessert. Zum anderen wird angenommen, dass erst die späteren Messzeitpunkte der Einschätzungen der mathematischen Fähigkeiten die (tatsächlichen) mathematischen Fähigkeiten der betreuten Kinder zum zweiten Messzeitpunkt signifikant beeinflussen. Ein solches Ergebnis würde darauf hindeuten, dass die Angemessenheit der Einschätzungen der frühpädagogischen Fachkräfte tatsächlich einen positiven Einfluss auf die mathematische Leistungsentwicklung ausübt.

### **Professionalität in der Gestaltung mathemathikhaltiger Interaktionen in der Kindertagesstätte und die Entwicklung mathematischer Kompetenzen bei Kindern**

**Aljoscha Jegodtka<sup>1</sup>, Corinna Schmude<sup>1</sup>, Katja Eilerts<sup>2</sup>, Sigrid Blömeke<sup>3</sup>**

<sup>1</sup>Alice Salomon Hochschule Berlin, <sup>2</sup>Humboldt-Universität zu Berlin, <sup>3</sup>Centre for Educational Measurement at the University of Oslo (CEMO)

#### Theoretischer Hintergrund

Dass die Interaktion zwischen Fachkraft und Kind in der Kindertagesstätte von besonderer Bedeutung für die kognitive Entwicklung von Kindern ist, ist in vielfältigen Studien untersucht worden, wobei insbesondere Formen langanhaltender Interaktionen bzw. sustained shared thinking als Kriterien für qualitativ hochwertige Interaktionen betrachtet werden können (z.B. König, 2009, Weltzien, 2013, 2014, Siraj-Blatchford et al., 2002, Pianta, 2017). Auch wenn aktuell durch vergleichende Untersuchungen zu den Effekten von lehrgangs- und spielgestützten Lernformen die Relevanz mathemathikhaltiger Interaktionen kritisch diskutiert wird (Hauser, 2017), kann basierend auf den Resultaten von Klibanoff et al. (2006) davon ausgegangen werden, dass die Entwicklung mathematischer Kompetenzen von Kindern in einem engen und signifikanten Zusammenhang zur Häufigkeit der durch frühpädagogische Fachkräfte genutzten mathemathikhaltigen Sprache steht. Nach Klibanoff et al. (2006) gibt es zwischen verschiedenen frühpädagogischen Fachkräften dramatische Differenzen in Bezug auf die Häufigkeit genutzter mathemathikhaltiger Sprache; und Rudd et al. (2008) haben gezeigt, dass es in Bezug auf die verschiedenen Facetten früher Mathematik deutliche Unterschiede gibt: die Bereiche Zahlen und Raum-Lage stehen deutlich im Mittelpunkt, Aspekte wie Musterfolgen oder dreidimensionale Körper tauchen hingegen deutlich weniger oft auf. Es ist insgesamt zu vermuten, dass eine größere Anzahl und verschiedene mathematische Facetten umfassende mathemathikhaltige Interaktionen als Kriterium für Professionalität frühpädagogischer Fachkräfte in Kindertagesstätten betrachtet werden kann.

Im Projekt Pro-KomMa: Professionalisierung des frühpädagogischen Studiums im Bereich Mathematik (Jenßen et al., 2016) wird ausgehend vom Pro-KomMa-Kompetenzmodell (Jegodtka et al., eingereicht) die Wirkkette zwischen mathematikbezogener Dispositionen, situationspezifischen Fertigkeiten, Performanz und mathematikbezogenem outcome bei den Kindern in der Population von frühpädagogischen Fachkräften in den ersten Jahren ihres Berufslebens untersucht.

#### Fragestellung

Der vorliegende Beitrag fokussiert die Frage, welcher Zusammenhang zwischen mathematikbezogener Performanz frühpädagogischer Fachkräfte und mathematikbezogenem outcome bei Kindern besteht.

#### Methode

Als ein Aspekt pädagogischer Professionalität wird der Zusammenhang zwischen mathematikbezogener Sprache und mathematischen Kompetenzen von Kindern fokussiert. Hierfür sind typische Alltagssituationen in Kindertagesstätten (Tietze, 1998) – Morgenkreis, gemeinsames Mittagessen und Freispielbegleitung – bei 12 frühpädagogischen Fachkräften videographiert worden und werden nach den Kriterien der qualitativen Inhaltsanalyse (Mayring, 1983, 2015) mit MAXQDA ausgewertet.

Unter einer mathemathikhaltigen Interaktion wird dabei eine sprachliche Äußerung oder Handlung verstanden, auf die mindestens einmal mathematikbezogen sprachlich oder durch Handlung reagiert wird; von einer langanhaltenden Interaktion wird dann ausgegangen, wenn sie mindestens über 30 Sekunden hinweg stattfinden. Die Qualitätsdimensionen langanhaltender mathemathikhaltiger Interaktionen sind dabei in Anlehnung an Siraj-Blatchford (2005) durch offene Fragen, modellierendes lautes Denken, das Einbringen von Vorschlägen, Aufforderungen und alternativer Sichtweisen, die auf sorgfältigem Beobachten, Respektieren und Begleiten der kindlichen Zugangsweisen basieren gekennzeichnet. Gleichzeitig wird in allen Frageaspekten mit untersucht, im Bereich welcher mathematischen Facetten (z.B. Musterfolgen, Raumorientierung, Zählen etc.) sich die Interaktion abspielt.

Die Resultate dieser Untersuchung werden trianguliert (Flick, 2011) in einen Zusammenhang gestellt zur mathematischen Kompetenz der Kinder (N= ca. 110-140), die mit dem MBK-0 (Seeger et al., 2014) erhoben wird. Hierdurch soll nachgezeichnet werden, welchen Einfluss mathemathikhaltige Interaktionen auf die Entwicklung mathematischer Kompetenzen von Kindern haben und inwiefern somit eine höhere Zahl und qualitativ hochwertigere mathemathikhaltiger Interaktionen als Kriterium pädagogischer Professionalität frühpädagogischer Fachkräfte gelten kann.

#### Ergebnisse

Die Datenauswertung findet aktuell statt und wird bis Ende 2017 abgeschlossen sein.

### **MiA-Num: Ein domänenspezifisches Beobachtungsinstrument zur Messung der Anregungsqualität von frühpädagogischen Fachkräften**

**Lara Pohle<sup>1</sup>, Lars Jenßen<sup>2</sup>, Catharina Ma<sup>3</sup>, Katja Eilerts<sup>3</sup>, Sigrid Blömeke<sup>4</sup>**

<sup>1</sup>Humboldt-Universität zu Berlin, <sup>2</sup>Freie Universität Berlin, <sup>3</sup>Humboldt-Universität Berlin, <sup>4</sup>Centre for Educational Measurement at the University of Oslo (CEMO)

#### Theoretischer Hintergrund

Performanz bildet eine Facette der professionellen Kompetenz von Fachkräften (Blömeke, Gustafsson, & Shavelson, 2015), die mit steigendem Interesse betrachtet wird. Die kindliche Entwicklung ist unter anderem vom professionellen Handeln abhängig (Mashburn et al., 2008), sodass von einem Zusammenhang zwischen dem Verhalten der Fachkräfte und der Leistung von Kindern ausgegangen werden kann (z.B. Terhart, 2012). Groß angelegte Studien der Schulforschung haben in diesem Zusammenhang spezifische Merkmale aufgezeigt, die einen (in)direkten Effekt auf den Lernzuwachs von Kindern haben (z.B. Kunter & Voss, 2013). Gleichzeitig wird in diesem Kontext auf das Fehlen fachspezifischer Instrumente verwiesen, mithilfe derer die Unterrichtsqualität im Fach Mathematik beurteilt werden kann (Schlesinger & Jentsch, 2016).



Auch die in der Frühpädagogik eingesetzten und validierten Instrumente (z.B. CLASS: Pianta, La Paro & Hamre, 2008; ECERS: Harms, Clifford & Cryer, 1980) unterstreichen das Fehlen von Beobachtungsverfahren, welche die Qualitätsmerkmale domänenspezifischer mathematischer Angebote erfassen und Effekte auf die numerische Kompetenzentwicklung von Kindern erklären. Da numerische Basiskompetenzen den wichtigsten Prädiktor für den späteren Schulerfolg von Kindern im Bereich Mathematik darstellen (Nguyen et al., 2016), wird ein Instrument benötigt, welches die domänenspezifische Untersuchung der Anregungsqualität von Fachkräften misst, psychometrischen Standards genügt und Schlussfolgerungen hinsichtlich der mathematischen Entwicklung von Kindern zulässt.

#### Fragestellung

Diese Präsentation stellt ein Instrument vor, welches den oben genannten Bedingungen entspricht, um die mathematische Anregungsqualität von ErzieherInnen im Bereich numerischer Basiskompetenzen zu beurteilen.

#### Methoden

Auf Basis eines für den nicht-deutschsprachigen Raum entwickelten Beobachtungsinstruments (BEMBEL: Thiel, 2009) ist ein domänenspezifisches Instrument entwickelt worden, welches die Beurteilung realen Handelns frühpädagogischer Fachkräfte im Bereich numerischer Basiskompetenzen erlaubt. Zur Sicherung der Anschlussfähigkeit sind in Anlehnung an Studien der Schulpädagogik (TIMSS: Wendt et al., 2016; COACTIV: Kunter et al., 2011) sowie auf Grundlage einer systematischen Recherche frühpädagogischer Fachliteratur Qualitätsmerkmale professionellen mathematischen Handelns identifiziert worden, die positive Effekte auf die mathematische Entwicklung von Kindern im Bereich numerischer Basiskompetenzen erwarten lassen. Die theoretisch angenommenen Qualitätsindikatoren sind auf einer dreistufigen Skala operationalisiert worden und repräsentieren die Konstrukte kognitive Aktivierung, konstruktive Lernunterstützung und mathematisch-inhaltliche Klarheit, die die theoretisch hergeleitete Anregungsqualität von Fachkräften auszeichnen.

Um die Aussagekraft des Instruments zu kontrollieren, werden rund 50 frühpädagogische Fachkräfte von geschulten BeobachterInnen mithilfe des Instruments MiA-Num eingeschätzt. Weiterhin werden die numerischen Basiskompetenzen von jeweils zehn Kindern pro Fachkraft (n=500) mithilfe eines standardisierten Instruments (MBK-0: Seeger et al., 2014; Cronbach's Alpha=0.87) gemessen. Die Kinder werden über einen Zeitraum von sechs Monaten zweimal getestet, wobei die zweite Testung nach der beobachteten Intervention stattfindet, sodass Effekte der Anregungsqualität auf die Entwicklung der numerischen Basisfähigkeiten der Kinder ermittelt werden können.

#### Ergebnisse

Erste vorläufige Ergebnisse der Validierung des Instruments weisen auf eine akzeptable Interrater-Reliabilität hin (ICC<sub>unjust</sub> = .63\*\*) und auch die Test-Retest-Reliabilität zeigt, dass die durchgeführten Ratings über einen Zeitraum von sechs Wochen stabil sind (r = .77\*\*). Nach Abschluss einer systematischen und standardisierten Inhaltsvalidierung durch neun externe Experten sind drei Items eliminiert worden, da sowohl Indikator als auch Item keine zufriedenstellende Repräsentation des Konstrukts darstellten. Sieben Items mussten überarbeitet werden, 14 Items konnten ohne Anpassung übernommen werden (Jenßen, Dunekacke, & Blömeke, 2015). Das Instrument umfasst somit Indikatoren, die das Konstrukt der Anregungsqualität in der Kita im Bereich Mathematik adäquat beschreibt.

Die vollständige Datenerhebung wird im Dezember 2017 und die Analyse der Daten im Januar 2018 abgeschlossen sein. Ergebnisse der ersten Kindertestung liegen bereits vor. Die Daten der Beobachtungen und Kindertestungen werden mittels Regressionen analysiert. Auf der Basis des theoretischen Hintergrunds wird erwartet, dass sich die numerischen Basiskompetenzen der Kinder in Abhängigkeit der erhobenen Anregungsqualität der Fachkräfte entwickeln. Die Diskussion erfolgt auf Grundlage verschiedener Konzepte hinsichtlich der Beziehung zwischen den numerischen Basiskompetenzen der Kinder und der Performanz der Fachkräfte (z.B. Sarama, Clements & Wolfe, 2016; Mashburn et al., 2008).

## D01

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus S103

### **Das Modell der Didaktischen Rekonstruktion im inklusiven Fachunterricht – Eine Pilotstudie in der (Fach-)Didaktik Bautechnik**

**Volker Rexing, Christina Keimes**

RWTH Aachen University, Deutschland; [rexing@fdb.rwth-aachen.de](mailto:rexing@fdb.rwth-aachen.de), [keimes@fdb.rwth-aachen.de](mailto:keimes@fdb.rwth-aachen.de)

#### a) Theoretischer Hintergrund

Die Forderung nach inklusiver Bildung ist keineswegs mehr neu und wird gleichermaßen international wie auch deutschlandweit diskutiert. Sie betrifft auch die berufliche Bildung im Allgemeinen und die beruflichen Schulen im Besonderen. Ein Blick auf den Forschungsstand offenbart allerdings, dass es im Hinblick auf eine inklusive Berufsbildung immer noch mehr Fragen als Antworten zu geben scheint (vgl. Niethammer/Friese 2017). Ein zentrales Desiderat in diesem Zusammenhang ist die Entwicklung von Leitlinien für einzelne Unterrichtsfächer und somit auch für die Didaktiken der beruflichen Fachrichtungen respektive die (Fach-)Didaktik Bautechnik (vgl. HRK/KMK 2015). Bislang fehlt es noch weitgehend an – insbesondere empirischen – Evidenzen für eine spezielle inklusive (Fach-)Didaktik (vgl. z. B. ebd.). Dies kann umso mehr für die berufliche Bildung und insbesondere die (Fach-)Didaktik der beruflichen Fachrichtungen konstatiert werden. Zwei originär fachdidaktische Kernthemen zeichnen sich allerdings ab, die in diesem Kontext von besonderer Relevanz sein dürften: (1) die Identifikation kognitiver Lernbarrieren (vgl. Nickolaus 2016) sowie (2) die Gestaltung differenzierter und individueller Lernwege (vgl. Niethammer/Langner 2017). Ausgangspunkt unserer Überlegungen ist die Annahme, dass sich die Heterogenität Lernender (und damit auch deren individuelle Lernbarrieren) stets in Abhängigkeit von einer konkreten Handlungssituation bzw. in Beziehung zu einem konkreten Lerngegenstand zeigt (vgl. Amrhein/Reich 2014; Feuser 2011). Es müssen also einerseits die individuellen (Lern-)Voraussetzungen der Lernenden berücksichtigt und andererseits die sachlogischen Zusammenhänge der Aneignungsgegenstände rekonstruiert werden, um mögliche Lernhürden (aber auch Lernpotentiale) offenzulegen und diese zum Ausgangspunkt für einen inklusionsorientierten Unterricht zu wählen (vgl. Niethammer/Friese 2017).

#### b) Fragestellung

In den naturwissenschaftlichen Fachdidaktiken ist das Modell der Didaktischen Rekonstruktion (vgl. z. B. Kattmann 2005) bereits als Forschungsparadigma etabliert und wird als theoretischer und methodischer Rahmen sowohl für die Unterrichtsplanung als auch für die fachdidaktische Lehr-Lernforschung verwendet (vgl. Reinfried et al. 2009). Die entscheidende Leistung des Modells besteht darin, systematisch Bezüge zwischen fachlichen Konzepten und den Vorverständnissen der Lernenden herzustellen (vgl. Gropengießer 2008). Dabei werden insbesondere Schülervorstellungen als subjektive gedankliche Prozesse fokussiert, die das Lernen und damit den Wissenserwerb wesentlich beeinflussen (vgl. ebd.). Nach unserem Kenntnisstand liegen zur Übertragbarkeit des Modells auf die Didaktiken beruflicher Fachrichtungen bisher keine Studien vor. Insofern wurde für eine Pilotstudie im Rahmen der (Fach-)Didaktik Bautechnik folgende leitende Fragestellung formuliert: Inwiefern können im Kontext einer inklusiven (Fach-)Didaktik Bautechnik über das Modell der Didaktischen Rekonstruktion (sub-)domänenspezifische Lernbarrieren identifiziert werden?

#### c) Methode

Die fachwissenschaftlichen Konzepte wurden auf Basis einer systematischen Auswahl von Quellentexten (vgl. Gropengießer 2008) und einer entsprechenden Dokumentenanalyse in Anlehnung an die Qualitative Inhaltsanalyse herauspräpariert (vgl. Mayring 2015). Die Erhebung der schülerseitigen Vorstellungen erfolgte mittels leitfadengestützter, halbstrukturierter (Gruppen-)Interviews mit Lernenden (n=35) in einer Fachklasse des Dualen Systems im Berufsfeld Bautechnik. Ausgewertet wurden die transkribierten Interviews mit Hilfe der Qualitativen Inhaltsanalyse (vgl. ebd.) sowie der Metapheranalyse (vgl. z.B. Schmitt 2011). Die Verknüpfung von wissenschaftlichen und Schülerkonzepten erfolgte auf Basis der Methode des wechselseitigen Vergleichs (vgl. Gropengießer 1997).

#### d) Ergebnisse

Auf Basis der bisherigen Erkenntnisse zeigt das Modell der Didaktischen Rekonstruktion auch im hier gegebenen Feld der (Fach-)Didaktik der beruflichen Fachrichtung Bautechnik Potential, (sub-)domänenspezifische Lernbarrieren, aber auch Lernpotentiale zu erfassen. Letztere erscheinen insbesondere vor dem zentralen Paradigma wertschätzenden Umgangs mit Heterogenität im Kontext einer inklusiven Fachdidaktik von besonderer Relevanz. Allerdings zeigt die beschränkte Reichweite durch den Fokus auf Schülervorstellungen die Notwendigkeit einer ergänzenden Diagnostik (im Hinblick auf weitere Facetten kognitiver Dispositionen). Außerdem sind direkte didaktisch-methodische Konsequenzen für inklusiven (Fach-)Unterricht nur bedingt ableitbar. Die Durchführung weiterer Studien zur Ausdifferenzierung der Schülervorstellungen bzw. zur Validierung der Übertragbarkeit von den naturwissenschaftlichen auf die (Fach-)Didaktiken der beruflichen Fachrichtungen zeigt sich als sinnvoll und notwendig.

### **Englischkompetenzentwicklung während der dualen Berufsausbildung in Abhängigkeit von dem Ausbildungsberuf und dem Stundenumfang des Englischunterrichts**

**Carolin Rätzer, Thomas Canz, Kathrin Jonkmann**

FernUniversität in Hagen, Deutschland; [carolin.raetzer@fernuni-hagen.de](mailto:carolin.raetzer@fernuni-hagen.de)

#### Theoretischer Hintergrund

Sichere Englischkenntnisse in Wort und Schrift werden für eine erfolgreiche Berufskarriere immer häufiger vorausgesetzt. Die Kompetenzentwicklung in der Domäne Englisch erfährt in einer globalisierten Welt ein großes öffentliches und bildungspolitisches, aber auch zunehmend wissenschaftliches Interesse. Da Englisch als Unterrichtsfach in Berufsschulen in Baden-Württemberg bisher nur eine Nebenrolle spielte, empfahl die Enquetekommission die Einführung des Fachs als einstündigen Pflichtunterricht, um das Missverhältnis zwischen Anspruch an die Fremdsprache und ihrem praktischen Stellenwert im Berufsschulcurriculum auszugleichen. Somit liegen derzeit drei Bedingungen für den Berufsschulenglischunterricht vor: kein Unterricht, eine Stunde Pflicht- und zwei Stunden Wahl(pflicht)unterricht. Welche Effekte eine fremdsprachliche Beschulung in Abhängigkeit vom Beschulungsumfang hat, ist bisher kaum untersucht, da das Forschungsfeld wenig Gelegenheit bietet, Variationen im Stundenumfang zu analysieren. Eine wichtige Ausnahme stellen die Befunde der ULME-Studie II (Lehmann et al., 2006) dar, die im teilqualifizierenden Berufsfachschulbereich eine Tendenz andeuten, dass eine höhere Anzahl an Englischunterrichtsstunden mit größeren Leistungszuwächsen einhergehen. Es ist jedoch bekannt, dass eine höhere Stundenanzahl nur bei gleichzeitiger Erhöhung der time on task, also der aktiven Zeit zur

Auseinandersetzung mit den Lerninhalten, zu besseren Lernerfolgen führt (Carroll, 1989). Insgesamt ist festzuhalten, dass bislang im Bereich der Entwicklung schulischer Kompetenzen nach der Schulzeit und der Kompetenzentwicklung während der dualen Berufsausbildung nur wenige Kenntnisse vorliegen. Die bestehenden Forschungslücken in dieser wichtigen Säule des deutschen Bildungssystems gilt es sowohl für den berufs- als auch für den allgemeinbildenden Teil zu schließen.

#### Fragestellung

Ziel der Studie ist die Untersuchung der Effekte des Englischunterrichts auf die Kompetenzentwicklung im Hinblick auf den Stundenumfang und das Berufsfeld. Es soll analysiert werden, wie sich die Englischkompetenzen der Schüler/innen in den Berufen Industriemechaniker/in, Industriekaufleute und Kaufleute für Büromanagement im Laufe der Ausbildung entwickeln und ob sich mögliche Kompetenzunterschiede auf die Berufsgruppen und den variierenden Englischstundenumfang zurückführen lassen.

#### Methode

Die vorliegende Untersuchung basiert auf den im Rahmen des Forschungsprojekts „Englischlerngelegenheiten in der dualen Berufsausbildung“ (EiBa) erhobenen Daten. Insgesamt nahmen 1.404 Schüler/innen der Berufe Industriemechaniker/in, Industriekaufleute und Kaufleute für Büromanagement aus 66 Berufsschulklassen zu drei Messzeitpunkten (1. Ausbildungsjahr: Dezember 2014 und Juli 2015, 2. Ausbildungsjahr: Juli 2016) teil. Zur Messung der Kompetenzen in Englisch wurden Textrekonstruktionen (C-Tests) und Hörverstehenstests (HV-Tests) eingesetzt. Fehlende Werte wurden mithilfe multipler Imputation ersetzt. Die Datenauswertung erfolgt mittels Wachstumskurvenmodellen unter Berücksichtigung der Mehrebenenstruktur der Stichprobe. Kriterien sind die IRT – skalierten Schätzungen der Personenfähigkeiten basierend auf Testergebnissen der Kompetenztests zum zweiten und dritten Messzeitpunkt. Als Prädiktoren gehen die Schülermerkmale Geschlecht, Schulabschluss, Migrationshintergrund, Englischabschlussnote, Schulabschluss, Bildungsabschluss der Eltern und ein Indikator der kognitiven Fähigkeiten (figurale Analogien des KFT) sowie die Klassenvariablen Ausbildungsberuf und Anzahl der Englischunterrichtsstunden ein.

#### Ergebnisse

Die Analysen zeigen für beide Kriterien signifikante Effekte auf die Kompetenzsteigerung über die Zeit. Für die Schülermerkmale Englischabschlussnote, Schulabschluss und KFT zeigen sich statistisch bedeutsame Einflüsse auf die C-Test-Leistung sowie auf die HV-Leistung. Nach Kontrolle der Schülervariablen ergeben sich sowohl signifikante Effekte der Berufsgruppe als auch der Unterrichtszeit. Es zeigen sich höhere Englischkompetenzen im Hörverstehen für Industriekaufleute mit einer und zwei Englischunterrichtsstunden sowie für Kaufleute für Büromanagement mit zwei Englischstunden im Vergleich zur Referenzgruppe Industriemechaniker/innen mit einer Unterrichtsstunde. Bei den Textrekonstruktionen wird nur der Kompetenzunterschied zwischen Industriekaufleuten mit einer Unterrichtsstunde und der Referenzgruppe signifikant. Bezüglich der Veränderungen der Englischkompetenzen über die Zeit zeigen sich bisher keine statistisch bedeutsamen Effekte der Berufsgruppe und der variierenden Stundenanzahl.

#### Ausblick

Die vorliegenden Ergebnisse liefern erste Eindrücke über die Englischkompetenzentwicklung während der dualen Berufsausbildung innerhalb eines Zeitraums von 18 Monaten. Diese Befunde werden in nächster Zeit mit Mehrebenen-Wachstumskurvenmodellen unter Einbezug motivationaler Konstrukte auf Grundlage des Erwartungs-mal-Wert Modells auf Schülerenebene und Variablen bezüglich Unterrichtsqualität auf Klassenebene zur Varianzaufklärung detaillierter überprüft.

## **Schülerfehler im Rechnungswesenunterricht – Eine Mixed-Methods-Analyse des Lernpotenzials**

**Patricia Köpfer**

Goethe-Universität Frankfurt, Deutschland; [koepfer@econ.uni-frankfurt.de](mailto:koepfer@econ.uni-frankfurt.de)

#### Theoretischer Hintergrund

Es wird angenommen, dass SchülerInnen aus ihren Fehlern im Unterricht lernen können (Yerushalmi & Polinger, 2006). Damit SchülerInnen aus Fehlverständnissen lernen und Fehler in Zukunft vermeiden können, benötigen sie die Unterstützung der Lehrkraft (Spychiger, Mahler, Hascher & Oser, 1998). Hierfür müssen Lehrkräfte über Professionelle Fehlerkompetenz (PFK) mit folgenden (domänen-spezifischen) Facetten verfügen: (1) Wissen über mögliche Fehlerarten, (2) Wissen über Handlungsstrategien und (3) das Vorhandensein von zielführenden Sichtweisen (Seifried & Wuttke, 2010).

Ergebnisse aus früheren Studien weisen auf Defizite von Lehrkräften hauptsächlich bezüglich Facette 1 und 2 hin, weshalb wir auf diese fokussieren: um Lernen aus Fehlern zu unterstützen, ist es erforderlich, dass Lehrkräfte (1) Schülerfehler erkennen, der Fehlerursache auf den Grund gehen und (2) hilfreiches Feedback geben, das den Lernprozess unterstützt.

In dieser Studie sollen die Wahrnehmungen der Lehrkraft und der SchülerInnen bezüglich Schülerfehler und des Umgangs mit Fehlern kombiniert werden.

Ziel dieser Studie ist es, das Lernpotenzial von Schülerfehlern aus unterschiedlichen Blickwinkeln zu betrachten.

Hierfür videographierten wir erfahrene Lehrkräfte in realen Unterrichtssituationen, führten Interviews mit den beteiligten Lehrkräften und SchülerInnen durch und testeten das Wissen der SchülerInnen in Rechnungswesen.

Folgende Fragestellungen sollen untersucht werden:

(1) Wie reagieren die Lehrkräfte auf auftretende Schülerfehler (Facette 1 und 2 der PFK)?

(2) Welches Lernpotenzial weisen die auftretenden Fehlersituationen – aus theoretischer Sicht und aus der Perspektive der SchülerInnen?

#### Methode

Wir videographierten Rechnungswesenunterricht an kaufmännisch-verwaltenden Vollzeitschulen von fünf Lehrkräften (nur männlich; Alter:  $M = 44.25$  ( $SD = 11.62$ ,  $Min = 35$ ,  $Max = 60$ ); Lehrerfahrung:  $M = 10.25$  ( $SD = 7.93$ ,  $Min = 2$ ,  $Max = 20$ )). Insgesamt nahmen 81 SchülerInnen teil (47% weiblich, 53% männlich; Alter:  $M = 17.03$  ( $SD = .96$ ,  $Min = 15$ ,  $Max = 20$ )). Das Unterrichtsthema war jeweils eines der zwei fehleranfälligsten Buchführungsthemen (Türling, 2014). Insgesamt ergaben sich 916 Minuten Videoaufnahmen. Zunächst identifizierten wir mittels qualitativer Inhaltsanalyse (Mayring, 2014) Fehlersituationen (d.h. ein Schülerfehler tritt in einer öffentlichen Unterrichtsinteraktion auf und wird als Fehler erkannt). Die identifizierten Fehlersituationen untersuchten wir auf ihr Lernpotenzial. Seifried und Wuttke (2010) folgend, kategorisierten wir hierfür das Lehrerfeedback anhand der Dimensionen 'Nachfassen' (dem Schülerfehler auf den Grund gehen) und 'Feedbackqualität'

(geringe Elaboration (Kurzantwort) versus hohe Elaboration (z.B. Erklärung, was falsch ist und wie es richtig gemacht wird)). Abhängig davon, ob die Lehrkraft ‚nachfasst‘ und dem Elaborationsgrad des Feedbacks, definieren Seifried und Wuttke (2010) vier Fehlersituationstypen. Es wird angenommen, dass Typ 1 (‘Nachfassen‘; hoher Elaborationsgrad) das höchste Lernpotenzial aufweist. Typ 2 (‘Nachfassen‘; geringer Elaborationsgrad) und Typ 3 (kein ‘Nachfassen‘; hoher Elaborationsgrad) scheinen weniger Lernpotenzial zu haben, während für Typ 4 (kein ‘Nachfassen‘ und geringer Elaborationsgrad) kein Lernpotenzial angenommen werden kann (Seifried & Wuttke, 2010). Die Intercoderreliabilität (Cohens Kappa) ist zufriedenstellend. Mit den fünf Lehrkräften sowie je sechs SchülerInnen pro Klasse führten wir Interviews durch. Wir zeigten ihnen im Rahmen von stimulated Recalls verschiedene Fehlersituationen aus den aufgezeichneten Unterrichtsstunden und fragten die SchülerInnen, ob und wie sie aus Fehlern lernen konnten und welche Feedbackarten sie präferieren. Die Lehrkräfte fragten wir, warum sie derart handelten.

#### Ergebnisse

Insgesamt identifizierten wir 98 Fehlersituationen, variierend von 12 in einer Klasse bis 38 in einer anderen. Empirische Ergebnisse und theoretische Überlegungen deuten auf das höchste Lernpotenzial bei Typ 1 hin. Jedoch fanden wir diesen Fehlersituationstyp lediglich fünf Mal in den beobachteten Unterrichtssituationen. Typ 2 trat einmal auf. Typ 3 (57 Mal) und Typ 4 (35 Mal) hingegen treten häufig auf, obwohl diesen Typen wenig und nahezu kein Lernpotenzial zugeschrieben werden. Die Interviewdaten zeigen, dass die Lehrkräfte glauben, die Fehlerursachen zu kennen und deshalb selten nachfassen. Die SchülerInnen aber bestätigen unter anderem die Wichtigkeit des Nachfassens, um aus Fehlern lernen zu können und wünschen sich diese Reaktion der Lehrkraft vermehrt.

## D02

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus S104

### Die Rolle der regionalen Bildungsinfrastruktur in Bezug auf Stadt-Land Disparitäten beim Gymnasialbesuch

**Michaela Sixt<sup>1</sup>, Christian Aßmann<sup>1,2</sup>**

<sup>1</sup>Leibniz-Institut für Bildungsverläufe, Deutschland; <sup>2</sup>Otto-Friedrich-Universität Bamberg; [michaela.sixt@lifbi.de](mailto:michaela.sixt@lifbi.de)

Vor dem Hintergrund (noch) sinkender Schülerzahlen und erfolgter Schulschließungen (Weishaupt, 2006) gewinnt die für die 1960er Jahre in der Literatur dokumentierte Bildungsbenachteiligung von Kindern auf dem Land im Vergleich zu den Stadtkindern aktuell an politischer Brisanz. So zeigt eine Studie für die Geburtskohorten 1980-1996, dass die Stadt-Land Disparitäten beim Gymnasialbesuch immer noch existent sind und durch die Unterschiede im regionalen Bildungsangebot erklärt werden können (Sixt, 2013).

In Anlehnung an Erikson und Jonsson (1996) lassen sich die erwarteten Effekte des regionalen Bildungsangebots und dessen Erreichbarkeit in einem erweiterten Nutzenkalkül für die Modellierung von Entscheidungen zwischen Bildungsalternativen theoretisch auch fundieren: Je besser die Bildungsinfrastruktur mit Blick auf eine bestimmte Alternative ausgebaut und erreichbar ist, desto geringer sind auch die damit einhergehenden Kosten. Der Nutzen und damit die Wahrscheinlichkeit der Wahl dieser Bildungsalternative steigen.

Die empirische Überprüfung und Quantifizierung der erwarteten Effekte erfolgen mittels Daten der Startkohorte 3 – Klasse 5 des Nationalen Bildungspanels (NEPS). Sie enthalten neben reichhaltigen Informationen zur tatsächlichen Bildungsentscheidung, auch die in der Literatur diskutierten Kontrollmerkmale auf Individualebene, wie z. B. Kompetenzen, Noten, Gymnasialempfehlung, soziale Herkunft und Migrationshintergrund. Zum anderen lassen sich die NEPS-Daten auf Kreisebene mit Daten zur Bildungsinfrastruktur des Sekundarbereichs als auch anderen strukturellen Merkmalen der Kreise verknüpfen.

Der Beitrag fokussiert die Bildungsentscheidung an der Schwelle vom Primar- zum Sekundarbereich des Bildungssystems, insbesondere die Wahl des Gymnasiums. Die aus dem theoretischen Kalkül abgeleitete empirische Spezifikation lässt sich in Form von hierarchischen logistischen Regressionen operationalisieren. Fehlende Werte in den Befragungswerten werden durch hierarchische Imputationen in der statistischen Analyse berücksichtigt. In ersten Analysen bestätigt sich der theoretisch vermutete positive Zusammenhang zwischen Gymnasialbesuch und Anteil der Gymnasien im Kreis des Wohnorts.

### Die Rolle von sprachlichen Fähigkeiten in der frühen Kindheit zur Erklärung von zuwanderungsbezogener Ungleichheit beim Übergang auf das Gymnasium

**Birgit Becker<sup>1,2</sup>, Oliver Klein<sup>3</sup>**

<sup>1</sup>Goethe Universität Frankfurt, Deutschland; <sup>2</sup>IdEA Zentrum, Frankfurt; <sup>3</sup>Mannheimer Zentrum für Europäische Sozialforschung (MZES); [Bi.Becker@em.uni-frankfurt.de](mailto:Bi.Becker@em.uni-frankfurt.de)

Einleitung und Fragestellung:

Im deutschen Bildungssystem schneiden Kinder und Jugendliche mit türkischem Migrationshintergrund unterdurchschnittlich ab (Kalter, Granato & Kristen 2007; Söhn & Özcan 2006). Disparitäten in verschiedenen Fähigkeiten, insbesondere im sprachlichen Bereich, werden dabei bereits in der frühen Kindheit beobachtet (Olczyk et al. 2016). Es stellt sich die Frage, ob solche Unterschiede in frühkindlichen Fähigkeiten für spätere Bildungsergebnisse eine Rolle spielen und die zuwanderungsbezogenen Nachteile (teilweise) erklären können. Konkret werden in diesem Beitrag die folgenden Forschungsfragen untersucht: Welchen Beitrag leisten kognitive, sprachliche und mathematische Fähigkeiten im Alter von drei Jahren zur Erklärung der ungleichen Übergangsraten auf das Gymnasium zwischen Kindern mit türkischem Migrationshintergrund und Kindern ohne Migrationshintergrund? Spielen frühkindliche sprachliche Fähigkeiten hierbei eine besondere Rolle?

Theoretischer Hintergrund:

In theoretischer Hinsicht beziehen wir uns auf das Modell der kumulativen Entwicklung von Fähigkeiten von Cunha und Heckman (2007). Heckman (2006: 1900) argumentiert, dass die kumulative Entwicklung von Fähigkeiten an der „Natur des Lernens“ liege, wobei frühe Fähigkeiten einerseits zu einer sich selbst verstärkenden Motivation für weiteres Lernen führen und andererseits das Lernen zu einem späteren Zeitpunkt effizienter und einfacher machen. Dieses Modell verbinden wir mit der theoretischen wie empirischen Erkenntnis aus der Migrationsforschung zur Schlüsselposition der Sprache für die Integration von Migranten und deren Nachfahren im Bildungssystem (Bleakley & Chin 2008; Demie & Strand 2006; Dollmann & Kristen 2010; Esser 2006; Kieffer 2008).

Daten und Methode:

Die empirischen Analysen werden mit den Daten aus dem Projekt „Erwerb von sprachlichen und kulturellen Kompetenzen von Migrant\*innenkindern und der Übergang nach der vierten Klasse“ durchgeführt. In diesem Längsschnittprojekt wurden türkischstämmige Kinder sowie Kinder ohne Migrationshintergrund ab dem dritten Lebensjahr bis zur vierten Klasse verfolgt, wobei regelmäßige Elterninterviews und standardisierte Leistungstests mit den Kindern durchgeführt wurden. Das Analysesample besteht aus 946 Kindern, wobei etwa die Hälfte einen türkischen Migrationshintergrund aufweist (mindestens ein Elternteil in der Türkei geboren).

Als abhängige Variable wird untersucht, ob das Kind in der vierten Klasse in einem Gymnasium angemeldet wurde. Die Fähigkeiten der Kinder in den Bereichen Kognition, Mathematik und deutscher Wortschatz im Alter von drei und sechs Jahren wurden mittels verschiedener Untertests der „Kaufman Assessment Battery for Children“ erhoben. In der dritten Klasse wurden die Schulleistungen in Lesen und Mathematik mittels der Tests DEMAT 3+ und WLLP erfasst. Zudem wurden die Noten in Deutsch und Mathematik in der vierten Klasse sowie die Übergangsempfehlung einbezogen.

Für die Auswertungen verwenden wir Linear Probability Models, geclustert nach Schulen. Um zu untersuchen, welche unabhängigen Variablen den Nachteil der türkischstämmigen Kinder bei der Gymnasialanmeldung erklären können, führen wir eine Dekomposition des indirekten Effekts des türkischen Migrationshintergrundes auf die Gymnasialanmeldung mittels Strukturgleichungsmodelle durch.

Ergebnisse:

Es zeigt sich, dass türkischstämmige Kinder eine wesentlich geringere Wahrscheinlichkeit für eine Gymnasialanmeldung aufweisen als Kinder ohne Migrationshintergrund. Dieser Nachteil reduziert sich sehr stark, wenn die soziale Herkunft (Bildung und Beruf der Eltern) statistisch kontrolliert wird. Werden zusätzlich die frühkindlichen Fähigkeiten einbezogen, so verschwindet dieser Nachteil komplett. Die mathematischen und sprachlichen Kompetenzen im Alter von drei Jahren sind dabei signifikant mit der Gymnasialanmeldung assoziiert. In weiteren Modellschritten zeigt sich, dass dieser Effekt der frühkindlichen Fähigkeiten durch spätere Fähigkeiten und Schulleistungen mediiert wird. Im finalen Modell, das alle Leistungsindikatoren umfasst, zeigt sich ein signifikant positiver Effekt des türkischen Migrationshintergrundes: Bei gleichem sozialen Hintergrund und gleichen Leistungen haben türkischstämmige Kinder eine höhere Wahrscheinlichkeit für eine Gymnasialanmeldung als Kinder ohne Migrationshintergrund.

Die Dekomposition des indirekten Effekts des türkischen Migrationshintergrundes auf die Gymnasialanmeldung ergibt, dass die soziale Herkunft (50%) und die sprachlichen Fähigkeiten im Alter von drei Jahren (30%) hierbei den größten Erklärungsbeitrag leisten. Die Nachteile türkischstämmiger Kinder beim Übergang auf das Gymnasium lassen sich somit vor allem auf diese Merkmale zurückführen.

## **Herkunftsspezifische Unterschiede bei der Schulwahl? Eine multivariate Untersuchung von Schulwahlmotiven bei der Wahl der weiterführenden Einzelschule.**

**Nicky Zunker, Marko Neumann, Kai Maaz**

Deutsches Institut für Internationale Pädagogische Forschung; [zunker@dipf.de](mailto:zunker@dipf.de)

### Theoretischer Hintergrund

Die Wahl der weiterführenden Schule wurde in Deutschland lange Zeit in erster Linie mit der Wahl der weiterführenden Schulform assoziiert. Empirische Untersuchungen haben dabei wiederholt herausgestellt, dass mit der Wahl der Schulform herkunftsbedingte Ungleichheiten einhergehen und führten dies unter anderem auf schichtspezifische Motivlagen und Kosten-Nutzen-Abwägungen zurück (z.B. Maaz et al., 2010). Infolge der Umstellung auf zweigliedrige Schulsysteme, in denen die zweite Säule neben dem Gymnasium oftmals die Möglichkeit zum Erwerb aller Abschlüsse bietet, richtet sich der Blick in der jüngeren Zeit zunehmend auf die Wahl der weiterführenden Einzelschule. Dies gilt im Besonderen für großstädtische Regionen, in denen es auf relativ engem Raum ein qualitativ wie quantitativ breites Angebot an Einzelschulen gibt. Gestützt auf internationale Forschungsbefunde wird dabei gemeinhin davon ausgegangen, dass auch bei der Einzelschulwahl von herkunftsspezifischem Wahlverhalten auszugehen ist, welches unter anderem als Resultat herkunftsspezifischer Motivlagen interpretiert wird. So wird unter anderem angenommen, dass für statushöhere Familien in stärkerem Maße schulische Qualitäts-, Angebots- und Leistungsmerkmale im Vordergrund stehen, während statusniedrigere Familien eher leistungsdistalen Aspekten wie der Wohnortnähe oder dem Besuch von Schulen, die auch von Geschwistern und Freunden besucht werden, größere Bedeutung beimessen würden (vgl. Hamilton & Guin, 2005). Empirische Befunde zu diesen Annahmen liegen jedoch für den deutschen Sprachraum bislang kaum vor.

### Fragestellung

An dieser Stelle setzt der vorliegende Beitrag an und geht der Frage nach, welche Aspekte Eltern bei der Wahl der konkreten weiterführenden Einzelschule wichtig sind. Im Fokus steht die Frage, ob und inwieweit sich herkunftsspezifische Unterschiede in der Bedeutsamkeit der Motive für die Schulwahl zeigen.

### Methode

Die Datengrundlage stellt eine Elternbefragung im Rahmen der BERLIN-Studie zur Evaluation der Berliner Schulstrukturreform dar (Maaz et al., 2013) dar, in der eine repräsentative Stichprobe von Berliner Eltern (N = 2658) mit Hilfe eines standardisierten Fragebogens zu ihren Schulwahlmotiven befragt wurde. Die Befragung fand zu Beginn der sechsten Klassenstufe und damit im letzten Schulbesuchsjahr der Grundschule statt. Die Erfassung der Schulwahlmotive orientierte sich an den Arbeiten von Clausen (2006) sowie Maaz et al. (2010), wobei den Eltern insgesamt 16 Items zu verschiedenen schulwahlrelevanten Aspekten zur Bewertung vorgelegt wurden. Die Untersuchung der herkunftsspezifischen Unterschiede in den Schulwahlmotiven erfolgte sowohl bivariat (Korrelationsanalysen und einfaktorielle Varianzanalysen) als auch multivariat (multiple lineare Regressionsanalysen).

### Ergebnisse

Mittels einer explorativen Faktorenanalyse konnten zunächst elf der 16 Einzelitems zu vier Motیفaktoren (Wohnortnähe, Bildungsorientierung, Leistungsorientierung, Angebotsorientierung) zusammengefasst werden. Die übrigen fünf Items wurden als Einzelitems in den Analysen berücksichtigt. Über alle Eltern hinweg stellten sich die Bildungsorientierung, ein guter Ruf und eine gute Ausstattung der Schule als bedeutendste Motive heraus. Die Betonung praxisorientierten Lernens sowie die Leistungs- und Angebotsorientierung der Eltern rangierten im mittleren Bereich. Als weniger bedeutsam stellten sich die Wohnortnähe, der gemeinsame Schulbesuch mit Geschwisterkindern und das Vorhandensein eines Ganztagsangebotes heraus.

Mit Blick auf herkunftsspezifische Motivlagen ließen sich kaum bedeutsame Unterschiede in Abhängigkeit der sozialen Herkunft und des Migrationshintergrundes ausmachen. Es deutete sich eher ein gegenläufiger Trend an, wonach die untersuchten Motive – mit Ausnahme der Bildungsorientierung – mit steigendem sozialem Status tendenziell weniger stark ausgeprägt waren. Ähnliches zeigte sich auch in Bezug auf den Zuwanderungshintergrund: Mit Ausnahme der Bildungsorientierung messen Familien mit Zuwanderungshintergrund allen Motiven eine höhere Bedeutung bei, als Familien ohne Zuwanderungshintergrund. Insgesamt fielen die gefundenen Unterschiede jedoch sehr gering aus. In den multivariaten Analysen ließen sich nur zwischen 1 und 8 Prozent der Varianz in den Schulwahlmotiven auf Herkunftsmerkmale zurückführen. Die Ergebnisse liefern Hinweise darauf, dass schichtspezifische Motivlagen bei der Wahl der weiterführenden Einzelschule geringer ausfallen als zu erwarten gewesen wäre und sollen in zukünftigen Analysen als Ausgangspunkt für die vertiefende Untersuchung der Herausbildung konkreter Schulpräferenzen und tatsächlich umgesetzter Schulwahlen dienen.

## **State- und Trait-Mathematikangst in der Übergangsphase von Primar- in die Sekundarstufe**

**Lars Orbach<sup>1</sup>, Moritz Herzog<sup>1,2</sup>, Annemarie Fritz<sup>1,2</sup>**

<sup>1</sup>Lehr- und Forschungsambulanz FoBu, Institut für Psychologie, Universität Duisburg-Essen, Germany; <sup>2</sup>Institute for Childhood Education, University of Johannesburg, South Africa; [lars.orbach@uni-due.de](mailto:lars.orbach@uni-due.de)

Bei der Bearbeitung einer Mathematikaufgabe hat der affektive Zustand einen bedeutenden Einfluss auf den Problemlösungsprozess. SuS antizipieren nach dem Kontroll-Wert-Ansatz (Pekrun, 1992) die subjektiv empfundene Kontrolle

und die Valenz angesichts einer zukünftigen Mathematikaufgabe. Demnach wird eine Emotion durch die subjektive Bewertung einer Leistungssituation erzeugt (Frenzel, Götz, & Pekrun, 2008). Wenn ein Kind dem Misserfolg in einem Mathematiktest eine hohe Valenz zuschreibt und kaum Möglichkeiten wahrnimmt, diesen Misserfolg zu verhindern (Kontrolle), sind die Bedingungen zur Entstehung einer Mathematikangst gegeben. Mathematikangst beeinträchtigt den Problemlösungsprozess, da angstbesetzte Gedanken Kapazitäten im Arbeitsgedächtnis blockieren, die für die Aufgabenbearbeitung notwendig sind (Hopko et al., 1998). Auch kann Mathematikangst negative Auswirkungen auf die Lernmotivation haben und Vermeidungsverhalten auslösen (Goldin, 2000; Dowker, 2005). Obwohl die negativen Auswirkungen der Mathematikangst unbestritten sind, liegen kaum Forschungserkenntnisse und Messinstrumente für die Grundschule und den Anfang der Sekundarstufe vor. So ist der Forschungsstand hinsichtlich der Erstmanifestation des Zusammenhangs zwischen Mathematikangst/Mathematikleistung inkonsistent. Viele Studien konnten Zusammenhänge in der Sekundarstufe feststellen, während die wenigen Studienergebnisse zur Primarstufe widersprüchlich sind. Eine Unterscheidung zwischen Zustandsängsten (State) und überdauernden Eigenschaftsängsten (Trait) wurde für diese Fragestellung bisher nicht vorgenommen. Ferner ist ungeklärt, ob Mathematikängste schulformspezifische Effekte aufweisen. Dies ist von Interesse, da Ängste einen moderaten Erblichkeitskoeffizienten im Vergleich zu anderen psychischen Auffälligkeiten besitzen und Umweltfaktoren einen wesentlichen Einfluss auf die Entwicklung von Ängsten nehmen.

Die empirische Studie untersucht die State- und Trait-Mathematikangst bei SuS der 4. und 5. Klasse. Insgesamt nahmen ca. 1300 SuS aus dem Ruhrgebiet und dessen Umland an der Studie teil. Zur Erfassung der Trait-Mathematikangst wurde eine adaptierte Version des Fragebogens zur Rechenangst (FRA, Krinzinger et al., 2007) und zur Erfassung der State-Mathematikangst wurde der Kinder-Angst-Test III (KAT-III) in einem Gruppentest durchgeführt. Zudem wurden die mathematischen Fertigkeiten mit den curricular orientierten Rechentests DEMAT 4/5+ und einem Test arithmetischer Basisfertigkeiten (Ehler et al., in press) erfasst. Der Fragebogen zur Rechenangst ist eine deutschsprachige Version des Math Anxiety Questionnaire (MAQ) von Thomas und Dowker (2000), der für 6.- bis 9.-jährige Kinder im Einzelsetting durch Krinzinger et al. (2007) normiert wurde. Im Rahmen der Studie wurde der FRA als Gruppentest adaptiert und für Kinder der 4. und 5. Klasse hinsichtlich seiner psychometrischen Güteeigenschaften analysiert und normiert. Neben der Erhebung deskriptiver Informationen zur Auftrittshäufigkeit und Verteilung der Mathematikangst in der Übergangsphase von Primar- in die Sekundarstufe befasst sich die Untersuchung mit den Zusammenhängen zwischen State- und Trait-Mathematikangst sowie den mathematischen Fertigkeiten. Ein besonderes Augenmerk liegt auf der Analyse von geschlechts- und schulformspezifischen Effekten.

Die Ergebnisse verdeutlichen eindeutige geschlechts- und schulformspezifische Ausprägungen der State- und Trait-Mathematikangst. Am Übergang zwischen Primar- und Sekundarstufe zeigten Mädchen aller Schulformen eine deutlich höhere Ausprägung der Trait-Mathematikangst. Erste Ergebnisse zeigten bei SuS des Gymnasiums im Vergleich zu den anderen Schulformen eine höhere Ausprägung der Trait-Mathematikangst, Geschlechtsunterschiede in der State-Mathematikangst variierten in Abhängigkeit von der Schulform: Hier zeigten nur Gesamt- und Realschülerinnen signifikant höhere Angstwerte. Insgesamt war die State-Mathematikangst auf der Grundschule höher ausgeprägt als in der Sekundarstufe. Insbesondere Gymnasiast\*innen berichteten eine geringere Ausprägung der State-Mathematikangst.

In der Grundschule wurden keine Geschlechtsunterschiede der mathematischen Fertigkeiten festgestellt, in der Sekundarstufe erzielten jedoch Jungen deutlich bessere Ergebnisse im DEMAT sowie in den Basisfertigkeiten. Es lag kein Unterschied zwischen den Basistestergebnissen der Primar- und Sekundarstufe vor.

Der Zusammenhang zwischen Mathematikangst und Rechenleistung wird vom Geschlecht moderiert. Während in der Grundschule keine Zusammenhänge zwischen State- oder Trait-Mathematikangst und Rechenleistung bestanden, waren mittlere Zusammenhänge zwischen State-Mathematikangst und Rechenleistung in der Sekundarstufe zu beobachten. Insbesondere Mädchen zeigten schlechtere Rechenleistung bei ausgeprägten State-Mathematikängsten. Für Jungen der Sekundarstufe wurden ähnliche Zusammenhänge für die Basisfertigkeiten ermittelt. In der Sekundarstufe zeigte sich bei den Jungen zudem ein leichter Zusammenhang zwischen Trait-Mathematikangst und schwächeren Rechenleistungen.

## D03

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus S105

### **“KEEP IT GOING, GIRL!” An Empirical Analysis of Gendered Career Chances and Career Aspirations among German Graduates in Computer Sciences**

**Silvia Maria Förtsch<sup>1</sup>, Anja Gärtig-Daug<sup>2</sup>, Ute Schmid<sup>3</sup>**

<sup>1</sup>Otto-Friedrich-Universität Bamberg, Deutschland; <sup>2</sup>Otto-Friedrich-Universität Bamberg, Deutschland; <sup>3</sup>Otto-Friedrich-Universität Bamberg, Deutschland; [silvia.foertsch@uni-bamberg.de](mailto:silvia.foertsch@uni-bamberg.de), [anja.gaertig-daug@uni-bamberg.de](mailto:anja.gaertig-daug@uni-bamberg.de), [ute.schmid@uni-bamberg.de](mailto:ute.schmid@uni-bamberg.de)

#### Theoretical framework:

The proportion of women in leading positions in the IT sector is small. Empirical analyses of the research project “Alumnae Tracking” (2012-2015) provide results on different career paths of men and women in computer science.

We discuss the theoretical model underlying our empirical analysis of gendered professional self-efficacy belief (Bandura 1986, 1997), career chances, and career aspirations (Lent 1986, 1994, 2000; Abele 2002, 2003, 2009, 2012) among German graduates in computer sciences. Of course, the factors mentioned above are the result of a complex interplay of personal, social, private, and societal factors. For the purpose of our study, however, we concentrate on the one hand on the effect of the individuals' 'objective' skills, and on the other hand on their self-perception of skills, the individuals' 'subjective' evaluation of skills, so to say.

#### Aims and objectives:

By interviewing graduates at different points in time, we wanted to gain insights into gender-specific motivation for career aspirations, changes in individual career goals, and differences in career advancement.

#### Sample:

The study offers detailed information on about 500 students and graduates of the Faculty of Information Systems and Applied Computer Sciences at Bamberg University. For the purpose of this article which aims at investigating gendered career chances and aspirations in computer sciences, we make use only of the subsample of graduates who report to be employed at the time of their interview. Based on this definition, our analytical sample consists of 103 individuals.

#### Methodology:

Questionnaires for graduates were designed to collect longitudinal data on self-assessment, academic achievement, professional self-efficacy beliefs, career chances and career aspirations.

Hypotheses were tested using a series of binary logistic regression models. The results always report average marginal effects (AME) and predictive margins (PM). Predictive margins can be interpreted as average probabilities. Average marginal effects are the percentage point differences between the estimated predictive margins. They allow testing whether or not differences in the estimated predictive margins are statistically significant. For better clarity, we only report the results for the graduates' gender as well as gender-specific interaction effects though our models usually control additional factors.

#### Results:

The results of our empirical study show that the academic achievement of the female graduates of our sample is as good as that of male graduates. Despite this fact, however, female graduates display lower professional self-efficacy beliefs than male graduates, partly because low-achieving male graduates still display high professional self-efficacy beliefs irrespectively of their previous academic achievement at university. Additionally, we find differences in men's and women's career chances and career aspirations. Male graduates depend less on their academic achievement at university to fill (or aspire) a leadership position in computer sciences whereas female graduates have to be very ambitious to be able to hold a leadership position in computer sciences. All in all, we interpret these findings as signs for the relevance of gender-stereotypes in computer sciences (Steele, 1997; Spencer et al. 1999).

#### Scientific and applied significance:

The results show that women perceive themselves differently than men. Women have the same career ambitions as men, but are still disadvantaged with regard to their career prospects. For this reason measures such as career coaching for example are necessary to support women in their career ambitions and goals. In doing so more female role models in the field of computer science can be created and encourage female employees to opt for a leadership career.

### **Geschlechtsspezifische Sozialisationsprozesse und das Bewerbungsverhalten auf Ausbildungsplätze in Zeiten der Digitalisierung**

**Sebastian Thürer<sup>1</sup>, Gwendolin Blossfeld<sup>2</sup>**

<sup>1</sup>LifBi – Leibniz-Institut für Bildungsverläufe e.V | Leibniz Institute for Educational Trajectories, Deutschland; <sup>2</sup>Otto-Friedrich-Universität Bamberg; [sebastian.thuerer@lifbi.de](mailto:sebastian.thuerer@lifbi.de)

Technologischer Wandel verändert seit jeher die Arbeitswelt und die Arbeitsbedingungen für Menschen (Autor, 2015; Brynjolfsson & McAfee, 2014; Grass & Weber 2016) und beeinflusst damit soziale Ungleichheit in unserer Gesellschaft. Durch die zunehmende Verbreitung digitaler Technologien wurde ein Veränderungsprozess in Gang gesetzt, der seinesgleichen sucht und in der aktuellen Diskussion unter dem Schlagwort ‚Arbeit 4.0‘ allgegenwärtig ist (Hirsch-Kreinsen, 2014; Schietinger, 2016). In diesem Zusammenhang wird häufig von der Substituierbarkeit des Menschen in Arbeitsprozessen durch (rechnergestützte) Maschinen gesprochen (Frey & Osborne, 2017; Dengler & Matthes, 2015). Der Maximalfall der Substituierbarkeit wird erreicht, wenn Berufe oder ganze Berufssegmente durch Maschinen vollständig ersetzt werden. Als negative Begleiterscheinung kann kurz- bis mittelfristig die Arbeitslosigkeit massiv zunehmen (Möller, 2015), weswegen die Frage nach der Auswahl eines Ausbildungsberufes, der eine gewisse Sicherheit vor einer solchen Substituierbarkeit bietet, gegenwärtig bedeutender ist denn je. Denn das getätigte Investment in eine Ausbildung sollte im optimalen Fall dazu dienen, den weiteren Lebensunterhalt bestreiten zu können. Wird in eine Ausbildung investiert, die nach kurzer Zeit kein Einkommen mehr generiert, da der Beruf durch Maschinen ersetzt wird, führt dies nicht nur zu (monetären) Verlusten beim betroffenen Individuum, sondern belastet auch die jeweilige Volkswirtschaft.

Vernachlässigt werden in der aktuellen Diskussion häufig jene Schüler/innen, die das Schulsystem verlassen und vor dem Eintritt in den Arbeitsmarkt stehen. Sie sehen sich mit veränderten Arbeits(markt)bedingungen konfrontiert, auf die das



Bildungssystem aktuell noch nicht adäquat vorzubereiten scheint (Aktionsrat Bildung, 2017). Erschwerend kommen Geschlechterdifferenzen im Zugang zu und Umgang mit digitaler Technologie hinzu (Aktionsrat Bildung, 2009). Es gewinnt die Frage vermehrt an Bedeutung, ob tradierte Geschlechterrollen und Stereotype einen Effekt auf die Auswahl von Ausbildungsberufen im Bereich Informations- und Kommunikationstechnologie (IKT) haben (Aktionsrat Bildung, 2009).

Die Arbeit behandelt die Verteilung von Bewerber/innen auf unterschiedliche Ausbildungsplätze unter Berücksichtigung ihrer IKT-Kompetenzen und ihres Geschlechts. Betrachtet wird die Verteilung der unterschiedlichen Kompetenzwerte der Schüler/innen, die sich im Übergang von der Schule zum Arbeitsmarkt befinden. Haben verschiedene Befähigungen im Bereich der IKT-Kompetenzen einen signifikanten Unterschied aufzuweisen, wenn differenzierte Schultypen und Geschlechter betrachtet werden? Inwieweit haben die IKT-Kompetenzwerte einen Einfluss auf die Auswahl der zur Verfügung stehenden Ausbildungsberufe; mit anderen Worten: Wer bewirbt sich auf welche Arten von Ausbildungsplätzen? Zusammenfassend wird das Bewerbungsverhalten von Schüler/innen unter der Bedingung der vorhandenen IKT-Kompetenzen und dem sich durch den technologischen Fortschritt rasch wandelnden Arbeitsmarkt betrachtet.

Die empirischen Analysen basieren auf den Daten der Startkohorte 4 (SC 4) des Nationalen Bildungspanels (NEPS) (Blossfeld, Roßbach, & Maurice, 2011). Seit Herbst 2010 werden dort repräsentativ ausgewählte Schüler/innen der 9. Klasse aller allgemeinbildenden Schultypen regelmäßig befragt und in unterschiedlichen Domänen (z.B. IKT oder Mathematik) getestet. Für die empirischen Analysen sind von besonderem Interesse die Schüler/innen, die sich in der 9. Klasse aktiv auf einen Ausbildungsplatz bewerben. Die Daten enthalten Testungen der IKT-Kompetenzen der Schüler/innen in Klasse 9/1 sowie Auskünfte zu den Berufen, auf die sich die Schüler/innen tatsächlich beworben haben. Angereichert werden die NEPS SC 4 Daten mit Informationen zu den Substituierbarkeitspotenzialen der Berufe (Dengler & Matthes, 2015), auf die sich die Schüler/innen in der 9. Klasse beworben haben. Diese Informationen stammen aus der Expertendatenbank BERUFENET.

Erste Ergebnisse zeigen deutliche Geschlechterunterschiede im Bewerbungsverhalten auf einen Ausbildungsplatz. Während die IKT-Kompetenzen bei Schülern keinen Einfluss auf das Ausbildungsplatzbewerbungsverhalten haben, zeigt sich bei Schülerinnen ein weitaus differenzierteres Bild. Schülerinnen mit hohen IKT-Kompetenzen bewerben sich häufiger auf Berufe mit einem hohen Substituierbarkeitsniveau ( $\geq 70\%$  der Tätigkeit) und wesentlich seltener auf Ausbildungsberufe, die ein niedriges Substituierbarkeitsniveau ( $\leq 30\%$  der Tätigkeit) aufweisen.

## **Sind muslimische Jugendliche traditioneller eingestellt? Eine Studie zum Zusammenhang zwischen Konfessionslosigkeit, Konfessionszugehörigkeit und Religiosität mit der Geschlechterrollenorientierung**

**Tatjana Taraszow<sup>1</sup>, Cornelia Gresch<sup>2</sup>, Birgit Heppt<sup>3</sup>**

<sup>1</sup>Fachbereich Erziehungswissenschaft und Psychologie, Freie Universität Berlin, Deutschland; <sup>2</sup>Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Humboldt-Universität zu Berlin; <sup>3</sup>Institut für Erziehungswissenschaften, Humboldt-Universität zu Berlin; [t.taraszow@fu-berlin.de](mailto:t.taraszow@fu-berlin.de)

### Theoretischer Hintergrund und Fragestellung

Bekannterweise unterscheiden sich Mädchen und Jungen im Bildungserfolg (Hannover & Kessels, 2011). Eine Erklärung besteht in unterschiedlichen Geschlechterrollenorientierungen (GeRo) (Hannover & Kessels, 2008), die als individuelle Einstellung gegenüber Geschlechterrollen - geschlechterangemessene Verhaltensweisen - verstanden wird (Volter, Braun & Hannover, 2015). Dabei wird zwischen traditioneller und egalitärer Orientierung unterschieden (Brogan & Kutner, 1976). Es gibt verschiedene Thesen über die Genese der GeRo bei Heranwachsenden. Dieser Beitrag legt den Fokus auf einen bislang wenig untersuchten Bereich: Der Zusammenhang von religiösen Aspekten mit der GeRo.

Entsprechend der „strukturellen These“ sollten konfessionsfreie Personen egalitäre Einstellungen zu Geschlechterrollen aufweisen als Personen mit einer Konfession - unabhängig von der Konfessionsart (vgl. Vanfossen, 1977). Unterschiede zwischen Konfessionsarten (z.B. christlich, muslimisch) werden hingegen in der „These des Islam“ formuliert: Demnach gehe der Islam nicht mit der Gleichstellung der Geschlechter einher (vgl. Diehl, Koenig & Ruckdeschel, 2009), denn Muslime präferierten traditionelle Geschlechterrollen und seien somit traditioneller eingestellt als Personen, die einer anderen Konfession angehören (Inglehart & Norris, 2003). Die „These der Religiosität“ nimmt an, dass die Religiosität einer Person ihre GeRo beeinflusst, und zwar unabhängig davon, ob und welcher Konfession die Person angehört, (Read, 2003), wobei unter Religiosität die religiöse Überzeugung sowie der praktizierte Glaube verstanden wird (Voas, 2007).

Bisherige empirische Untersuchungen lieferten zu allen drei Thesen uneinheitliche Befunde. Insbesondere ist unklar, welcher religiöse Aspekt den besten Prädiktor der GeRo darstellt. Vor diesem Hintergrund beschäftigt sich der vorliegende Beitrag mit folgenden Fragen: (1) Geht Konfessionslosigkeit mit einer egalitären GeRo einher? (2) Sind muslimische Heranwachsende traditioneller eingestellt als nicht-muslimische Jugendliche? (3) Weisen stärker religiöse Heranwachsende eine traditionellere GeRo auf? (4) Welcher der drei religiösen Aspekte Konfessionslosigkeit, Konfessionsart und Religiosität tragen am stärksten zu einer traditionellen GeRo bei?

### Methode

Zur Bearbeitung der Forschungsfrage wurde der Datensatz der Jahrgangsstufe 9 des Nationalen Bildungspanels (Blossfeld, Roßbach & von Maurice, 2011) herangezogen. Die Stichprobe besteht aus 13 474 Jugendlichen (51% weiblich), die im Jahr 2011 durchschnittlich 16.48 Jahre (SD=0.62) Jahre alt waren. Zur Erklärung der GeRo durch die Konfessionslosigkeit (ja/nein), Konfessionsart (christliche, muslimische und andere Religion) und Religiosität (religiöse Überzeugung (Skala) und praktizierter Glaube (ja/nein)) wurden multiple Regressionsanalysen durchgeführt. Als Kontrollvariablen wurden Geschlecht, Geschlechterrollenorientierung der Peers, SES-Indikatoren und Geschlechterrollenorientierung der Eltern herangezogen (Taraszow, Gresch & Kempert, 2017). Fehlende Werte wurden mit dem Programm ice von Stata multipel imputiert (Royston, 2004).

### Ergebnisse

Von den Jugendlichen gaben 80% an, einer Konfession anzugehören, davon 90% einer christlichen Religion, 8% einer muslimischen Religion und 2% einer anderen Religion. Die religiöse Überzeugung lag auf einer Skala von gar nicht religiös (1) bis sehr religiös (4) bei durchschnittlich  $M=2.24$  ( $SD=0.91$ ), unabhängig davon, ob die Jugendlichen einer Konfession angehörten oder nicht. Die Mehrheit aller Jugendlichen praktizierte ihren Glauben nicht (79%).

In den Analysen können alle drei Thesen bestätigt werden. Konfessionsfreie Heranwachsende sind egalitärer eingestellt als Heranwachsende mit einer Konfession. Jugendliche, die einer muslimischen Religion angehören, sind im Vergleich zu Jugendlichen einer christlichen Religion traditioneller eingestellt. Je stärker die religiöse Überzeugung, desto traditioneller die GeRo. Jugendliche, die ihren Glauben praktizieren, sind traditioneller eingestellt. Bei gemeinsamer Betrachtung der Effekte von

Konfessionslosigkeit, Konfessionsart und Religiosität konnte die religiöse Überzeugung als Teil der Religiosität als stärkster Prädiktor der traditionellen GeRo identifiziert werden.

#### Diskussion

Die vorliegende Studie stützt die drei o.g. Thesen zeigt, dass Religiosität eine traditionelle GeRo verstärkt. Damit stehen die Ergebnisse in Einklang mit Studien, die zeigen, dass unter Berücksichtigung der Religiosität nicht die Konfessionsart sondern die Religiositätsstärke eine Rolle für Geschlechter-gerechtigkeit spielt und dass traditionelle Konfessionsarten mit geringerem Bildungserfolg einhergehen (Keysar & Kosmin, 1995; Schnabel, 2016). Diese Befunde können dazu beitragen, Lehrkräfte zu sensibilisieren, ihren stärker religiösen Schüler\*innen eine egalitäre GeRo zu vermitteln, z.B. durch die Auswahl geeigneter Unterrichtsmaterialien.

## D04

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Rosshofgasse, S02

### **Stressbegünstigende bzw. -mindernde Bedingungskonstellationen im kaufmännischen Unterricht – Eine Prozessstudie auf Basis des transaktionalen Stressmodells**

**Tobias Kärner, Julia Warwas**

Universität Bamberg, Deutschland; [tobias.kaerner@uni-bamberg.de](mailto:tobias.kaerner@uni-bamberg.de), [julia.warwas@uni-bamberg.de](mailto:julia.warwas@uni-bamberg.de)

Im Kontext von Schule und Hochschule widmeten sich bereits diverse Studien den lern- und leistungsbeeinträchtigenden Wirkungen von Stressoren, die für institutionalisierte Lernumwelten charakteristisch sind. Sie zeigten jedoch auch auf, dass nicht alle Lernenden gleichermaßen von den erwarteten Beeinträchtigungen betroffen sind (z. B. Akgun & Ciarrochi, 2003; Liu & Lu, 2012; Vaez & Laflamme, 2008). Aus einer präventiven Perspektive stellt sich deshalb die Frage, unter welchen Konstellationen situations- und personenspezifischer Bedingungen Schüler/innen im Unterricht unter akuten Stress geraten, um entsprechende Erkenntnisse letztlich für didaktisch-methodische Planungen sowie für die Stärkung schülerindividueller Stressresistenzen nutzen zu können.

Erklärungsangebote bieten die Belastungs-Beanspruchungs-Forschung (zsf. Van Dick & Stegmann, 2013) und insb. das transaktionale Stressmodell (Lazarus & Launier, 1981). Auf dieser Grundlage lassen sich situationsabhängige Stressoren (i. S. v. Belastungseinwirkungen des Unterrichtsgeschehens) von personenspezifisch verfügbaren Ressourcen (i. S. v. Bewältigungskapazitäten) abgrenzen. Weiterhin ist davon auszugehen, dass Stressreaktionen auftreten, wenn verfügbare Ressourcen angesichts aktueller Anforderungen als defizitär bewertet werden. Dies hat nicht nur zur Folge, dass verschiedene Individuen objektivierbar gleiche Unterrichtsmerkmale abhängig von ihrer Ressourcenausstattung als unterschiedlich stressreich empfinden, sondern auch, dass Stressreaktionen intraindividuell über verschiedene Anforderungssituationen hinweg an- und abschwellen können (siehe auch Nezlek, 2007).

Zwecks empirischer Analyse der Haupt- und Interaktionseffekte von unterrichtlichen Anforderungen und individuellen Ressourceneinschätzungen auf akute Stressreaktionen von Schüler/innen wurde daher eine prozessbegleitende, triangulative Unterrichtsstudie konzipiert. Neben einem psychischen Stressindikator (Stressempfinden) fand dabei auch ein vergleichsweise unspezifischer Indikator physiologischen Stresses (Herzrate) Berücksichtigung.

Den statistischen Auswertungen liegen Daten einer Gelegenheitsstichprobe von insgesamt 53 angehenden Industriekaufleuten (35 weiblich) mit einem durchschnittlichen Alter von 19.53 Jahren aus zwei Klassen der 10. Jahrgangsstufe einer deutschen Berufsschule (SD = 4.76) zugrunde. Sie wurden über einen Zeitraum von drei Wochen (9 × 45 Min. Unterricht) im Fach „Betriebswirtschaftliche Geschäftsprozesse“ videografiert. Während des Unterrichts machten die Lernenden im zehnmütigen Rhythmus Aussagen über ihre situationspezifischen Stressempfindungen und Bewältigungsfähigkeiten. Dies wurde mittels der Continuous-State-Sampling Methode über portable Dateneingabegeräte (Palm Tungsten E2®) realisiert. Die Herzrate der Schüler/innen wurde kontinuierlich während des Unterrichts mittels Brustgurten und portablen Speichergeräten gemessen (M = 96.68, SD = 16.05). Zur Kategorisierung situativ variierender Anforderungsmerkmale wurden sowohl beobachtbare Aktivitätsformen (schüler- vs. lehrerzentriert; Seidel, Dalehefte & Meyer, 2001) als auch Anforderungsniveaus der inhaltlichen Auseinandersetzung (angelehnt an die Lernzieltaxonomie von Bloom et al., 1956) von zwei trainierten Beobachtern kodiert. Bei guter Interrater-Reliabilität legt eine Hauptkomponentenanalyse eine Einfaktorlösung nahe, weshalb Aktivitätsform und Anspruchsniveau als ein standardisierter Faktorwert (i. S. d. situativen Anforderungsgehalts des Unterrichts) in die weiteren Berechnungen einfließen. Herzratenmessungen sowie Beobachtungsdaten wurden im Zuge der statistischen Auswertungen jeweils auf 10-minütige Zeitintervalle aggregiert, um sie mit den Erlebensdaten zu synchronisieren. Acht Wochen vor dem Unterricht wurden situationsübergreifend relevante Ressourceneinschätzungen (schulische Selbstwirksamkeit nach Jerusalem & Satow, 1999) sowie – als Kontrollvariablen – Alter, Geschlecht und das fachliche Vorwissen der Schüler/innen (wirtschaftskundlicher Bildungstest nach Beck & Krumm, 1998) fragebogen- bzw. testbasiert erfasst.

Mehrebenenanalysen (Fixed-Effects-Modelle mit autoregressiver Kovarianzstruktur) liefern folgende Befunde: Ein signifikant positiver Haupteffekt des situativen Anforderungsgehaltes des Unterrichts lässt sich sowohl für die Herzratenwerte als auch für die Stressempfindungen der Lernenden nachweisen. Negative Haupteffekte sowie moderierende Effekte individueller Ressourcenbewertungen lassen sich nur bei Verwendung des psychischen Stressindikators nachweisen. Auch unter Konstanzhaltung von Alter, Fachwissen und Klassenzugehörigkeit der Lernenden steigen deren Stressempfindungen theoriekonform mit zunehmenden Unterrichtsanforderungen umso stärker an, je weniger sie von eigenen Bewältigungsfähigkeiten überzeugt sind.

Diskutiert werden zum einen didaktische Maßnahmen, die sog. mastery experiences (Bandura, 2006) ermöglichen und damit zur Stärkung schülerseitiger Wirksamkeitsüberzeugungen beitragen. Zum anderen werden forschungsmethodische Herausforderungen einer validen und zugleich praktikablen Messung physiologischer Stresssymptome erörtert.

### **Leistungsangstverlauf im Vorfeld einer Klausur – kurz gefasst**

**Christin Lotz, Rebecca Schneider, Jörn Sparfeldt**

Universität des Saarlandes, Deutschland; [c.lotz@mx.uni-saarland.de](mailto:c.lotz@mx.uni-saarland.de)

Theorie: Im Umfeld schulischer und universitärer Prüfungen erleben Schülerinnen und Schüler sowie Studierende häufig Leistungsangst. Mit zunehmender zeitlicher Nähe zur Prüfung findet sich häufig ein moderater Leistungsangstanstieg. Zur Durchführung groß angelegter Studien besteht u.a. aus Ökonomiegründen großes Interesse an Kurzskalen mit sehr wenigen Items oder sogar nur einem Item. Allerdings wird entsprechenden Kurzskalen aus diagnostischer Perspektive teilweise mit Misstrauen begegnet. Es soll daher die Forschungsfrage betrachtet werden, ob sich eine 1-Item-Kurzskala diagnostisch vergleichbar gut wie eine Langskala zur Leistungsangsterfassung eignet.

Methode: Im Rahmen einer Vorlesung bearbeiteten N = 190 Studierende viermal im Semesterverlauf (Semesterbeginn, Semestermitte, Semesterende und direkt vor der Klausur) einen state-Leistungsangstfragebogen mit jeweils einer 8-Item-Langskala und einer 1-Item-Kurzskala. Die Leistungsangstskalen bezogen sich zum ersten, dritten und vierten Messzeitpunkt auf die zur Vorlesung gehörende Pflichtklausur. Zum zweiten Messzeitpunkt bezogen sich die Skalen einmalig auf eine unbenotete Übungsaufgabensammlung. Außerdem beantworteten die Studierenden zu Semesterbeginn einen trait-Leistungsangstfragebogen. Im Hinblick auf die Abschlussklausur konnten die Studierenden zwischen zwei Terminen wählen: entweder in der letzten Semesterwoche (n = 96) oder acht Wochen später in den Semesterferien (n = 94). Wir berücksichtigten daher die Klausurgruppe („frühe“ oder „späte“) in den Analysen, da die ersten drei Erhebungszeitpunkte je nach Gruppe unterschiedliche Abstände zur Klausur hatten.

Zur Analyse des Leistungsangstverlaufs berechneten wir getrennt für die Kurzsкала und die Langskala jeweils eine 2 (frühe vs. späte Klausurgruppe) × 4 (Messzeitpunkte T1 bis T4) mixed ANOVA. Im Falle eines statistisch signifikanten Haupteffekts der Klausurgruppe analysierten wir den state-Leistungsangstverlauf pro Gruppe mithilfe von Messwiederholungs-ANOVAs mit gerichteten a priori Kontrasten, die einseitig prüften, ob das state-Leistungsangst-Niveau (a) zu T1 höher als zu T2 (unbenotete Übungsaufgaben), (b) zu T2 niedriger als zu T3, (c) zu T3 höher als zu T1 und (d) zu T4 höher als zu T3 war. Diese mit Kurz- und Langskala erfassten Leistungsangstverläufe wurden verglichen. Weiterhin berechneten wir pro Messzeitpunkt die Korrelation der Kurz- und der Langskala. Außerdem verglichen wir die Korrelationen beider state-Leistungsangstskalen (Kurz- und Langskala) mit den trait-Leistungsangstfacetten Aufregtheit und Besorgtheit.

Ergebnisse: Vergleichbar zum Verlauf der Langskala ergab die Kurzsкала (a) einen Leistungsangstabfall von T1 zu T2 (unbenotete Übungsaufgaben), (b) einen Anstieg von T2 zu T3, (c) jedoch nicht den erwarteten Zuwachs von T1 zu T3 und (d) einen Leistungsangstanstieg von T3 zu T4. Wie bei der Langskala wies auch die Kurzsкала für die Studierenden der frühen im Vergleich zur späten Klausurgruppe höhere Leistungsangstwerte von meist kleiner bis mittlerer Effektgröße auf.

Lang- und Kurzsкала korrelierten in beiden Klausurgruppen pro Messzeitpunkt sehr hoch ( $.76 \leq r \leq .91$ ). Die Vergleiche der Korrelationskoeffizienten der trait-Leistungsangstfacetten Besorgtheit und Aufregtheit mit einerseits der Kurzsкала und andererseits der Langskala ergaben zumeist sehr kleine Unterschiede zwischen beiden Skalen ( $|q| < .10$ ).

Diskussion: Die Kurzsкала ergab einen zur Langskala vergleichbaren moderaten Leistungsangstverlauf über das Semester hinweg mit einem Angstanstieg kurz vor der Klausur. Insbesondere der Angsteinbruch zum zweiten Messzeitpunkt spricht für die Skalen-Sensitivität, Veränderungen detektieren zu können. Die diagnostische Eignung der Kurzsкала wird weiterhin belegt durch (a) die hohen Korrelationen zwischen Kurz- und Langskala pro Messzeitpunkt sowie (b) die vergleichbar hohen Korrelationen beider state-Leistungsangstskalen mit den trait-Leistungsangstfacetten. Das Potential von Kurzsкаlen als attraktive Alternative zu Langskalen in der Leistungsangstmessung wird diskutiert.

## **Hautleitfähigkeit und Mathematikbezogene Ängstlichkeit – Eine Anwendung des Kontroll-Wert-Ansatzes**

**Anselm Strohmaier, Anja Schiepe-Tiska, Kristina Reiss**

Technische Universität München, Deutschland; [anselm.strohmaier@tum.de](mailto:anselm.strohmaier@tum.de)

### Theoretischer Hintergrund

Leistungsemotionen sind wichtige Aspekte erfolgreicher Bildungsprozesse (Pekrun & Linnenbrink-Garcia, 2014). Mathematikbezogene Ängstlichkeit (engl. mathematical anxiety = MA) ist in deutschen Klassenzimmern sehr präsent (Schiepe-Tiska & Schmidtner, 2013). Gleichzeitig ist bekannt, dass MA langfristig und kurzfristig einen negativen Einfluss auf Leistungen in Mathematik hat (Dowker, Sarkar, & Looi, 2016). Um die Prozesse hinter diesem Einfluss zu verstehen, ist es sinnvoll, momentane (state) MA und habituelle (trait) Ängstlichkeit zu unterscheiden (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). Für die Messung momentaner Ängstlichkeit bieten sich physiologische Indikatoren an, sind aber bisher in der Forschung zu MA kaum verwendet worden (Hannula, 2016).

Elektrodermale Aktivität (EDA) ist ein Indikator für Stress, Aufregung und Aktivierung (Boucsein, 2012). Sie kann deshalb zur Beobachtung momentaner Ängstlichkeit in mathematischen Anforderungssituationen eingesetzt werden (Dew, Galassi, & Galassi, 1984). Im Gegensatz zu ersten Untersuchungen vor über 30 Jahren stehen heute die technischen Möglichkeiten zur Verfügung, um EDA in authentischen mathematischen Anforderungssituationen zu messen, ohne dabei auf unhandliche Laborapparate zurückgreifen zu müssen. Darüber hinaus wurden Theorien zum Einfluss von Leistungsemotionen entscheidend weiterentwickelt (Pekrun & Perry, 2014).

Die vorliegende Studie versteht sich als Erprobung einer Technik zur Beobachtung emotionaler Prozesse in Lern- und Leistungssituationen, die im Zusammenhang mit Ängstlichkeit und Stress stehen. Dadurch öffnen wir neue Perspektiven für die Berücksichtigung emotionaler Aspekte von Kompetenz, auch über die Mathematik hinaus. Die Studie folgt dabei dem Kontroll-Wert-Ansatz (Pekrun, 2006) für Leistungsemotionen.

### Fragestellung

Ausgehend von der anhaltenden Relevanz und steigenden Berücksichtigung emotionaler Aspekte erfolgreicher Bildungsprozesse zielt diese Studie auf die Beobachtung momentaner Aufregung in mathematischen Testsituationen. Es wird untersucht, ob ein Zusammenhang von EDA und habitueller MA beobachtet werden kann, ob also EDA ein sinnvoller Indikator momentaner Ängstlichkeit ist, die durch MA verursacht wird. Weiterhin wird untersucht, ob Ängstlichkeit hier gemäß dem Kontroll-Wert-Ansatz mit subjektiver Kontrolle und dem Wert der Aktivität in Verbindung steht.

### Methode

An der Studie nahmen 86 Studierende nicht-mathematischer Studiengänge teil (Alter  $M = 23,2$ ;  $SD = 4,07$ ; 62% weiblich). Sie wurden in vollem Umfang über die Messung der Hautleitfähigkeit informiert. Die Teilnehmerinnen und Teilnehmer absolvierten in Einzelsitzungen zunächst eine ca. 6-minütige Entspannungsübung, bevor zwei Mathematiktests bearbeitet wurden. Die Tests bestanden aus jeweils 6 Items der PISA- und TIMS-Studie und wurden unter Zeitdruck in jeweils 10 Minuten bearbeitet. Direkt nach jedem der beiden Tests wurden die subjektive Kontrolle und der Wert der Aktivität per Fragebogen abgefragt. Zuletzt wurden in einem weiteren Fragebogen Kontrollvariablen sowie habituelle MA mit der Skala ANXMAT erfasst (Ramm et al., 2006).

### Ergebnisse

Als Indikator momentaner Ängstlichkeit wurde als abhängige Variable die Frequenz unspezifischer Reaktionen (NSCR.freq) gewählt (Boucsein, 2012). Zunächst zeigte sich ein signifikanter Anstieg der EDA während der Tests verglichen mit der Entspannungsübung ( $F(1,85) = 17,23$   $p < 0,01$ ).

Zur weiteren Analyse wurde eine schrittweise multiple lineare Regression durchgeführt. Die durch die Kontrollvariablen aufgeklärte Varianz ( $R^2 = 0,576$ ) erhöhte sich nicht signifikant durch Hinzunahme von habitueller MA (Anstieg  $R^2 = 0,003$ ), jedoch durch Hinzunahme der Prädiktoren subjektive Kontrolle und Wert (Anstieg  $R^2 = 0,035^*$ ). Durch Hinzunahme der Interaktionsterme klärte das Modell erneut signifikant mehr Varianz auf (Anstieg  $R^2 = 0,038^*$ ). Dabei zeigte sich insbesondere eine signifikante Dreifachinteraktion aus MA, subjektiver Kontrolle und Wert ( $\beta = -0,212$ ;  $p < 0,01$ ).

### Diskussion

Erwartungsgemäß zeigte die erhöhte EDA während der Mathematiktests, dass diese Methode generell geeignet ist, um akademische Leistungssituationen zu beobachten. In Einklang mit dem Kontroll-Wert-Ansatz konnte ein indirekter Einfluss von habitueller mathematikbezogener Ängstlichkeit auf momentane Aufregung, vermittelt über subjektive Kontrolle und Wert der

Aktivität, bestätigt werden. Die Verwendung handlicher Messgeräte zur Erfassung von EDA scheint damit ein vielversprechendes Werkzeug zur Beobachtung momentaner Emotionen in Bildungssituationen zu sein.

## **Kognitive Belastung, Stress und Leistung bei der Auseinandersetzung mit molekularen Repräsentationen**

**Nina Minkley<sup>1</sup>, Tobias Kärner<sup>2</sup>, Atila Jojart<sup>1</sup>, Lasse Nobbe<sup>1</sup>, Moritz Krell<sup>3</sup>**

<sup>1</sup>Ruhr-Universität Bochum, Deutschland; <sup>2</sup>Universität Konstanz, Deutschland; <sup>3</sup>Freie Universität Berlin, Deutschland;  
[Nina.Minkley@rub.de](mailto:Nina.Minkley@rub.de), [moritz.krell@fu-berlin.de](mailto:moritz.krell@fu-berlin.de)

### Theoretischer Hintergrund

Die Repräsentationskompetenz ist ein wichtiger Teil naturwissenschaftlicher Grundbildung (scientific literacy; Gilbert & Treagust, 2009). Zum Beispiel helfen molekulare Repräsentationen, die Beziehungen zwischen Strukturen und Systemen darzustellen und zu verstehen (Cooper et al., 2010). Gleichzeitig zeigen Schülerinnen und Schüler (SuS), sowie auch Studierende Schwierigkeiten beim Verständnis molekularer Repräsentationen (z.B. Strukturformeln; Taskin & Bernholt, 2012).

Die Informationsverarbeitung beim Umgang mit Repräsentationen findet primär im Arbeitsgedächtnis statt, welches nur eine begrenzte Menge an Informationen parallel verarbeiten kann (Mayer, 2008; Schnotz, 2005). Unnötige oder redundante Elemente einer Repräsentation können die kognitive Beanspruchung erhöhen und folglich die Informationsverarbeitung behindern (vgl. cognitive load theory; Paas, Renkl, & Sweller, 2004; Sweller, 2005).

Ein weiterer Aspekt, der die Informationsverarbeitung beim Umgang mit molekularen Repräsentationen beeinträchtigen könnte, ist die Stressbelastung. Da SuS gerade mit komplexeren Repräsentationsarten (z.B. symbolisch-textlich) große Schwierigkeiten haben (Wu & Shah, 2004), kann die Konfrontation mit ihnen Versagensängste und damit einhergehend Stressreaktionen auslösen (Dickerson & Kemeny, 2004).

### Fragestellungen

Diese Studie untersucht den Zusammenhang zwischen der Art molekularer Repräsentationen in Testaufgaben (rein symbolisch vs. symbolisch-textlich), der kognitiven Belastung, der Stressreaktion und der Testleistung von SuS.

- Inwiefern präferieren SuS Tests mit rein symbolischen molekularen Repräsentationen gegenüber solchen mit symbolisch-textlichen Repräsentationen?
- Inwiefern führt die Art der Repräsentation in Testaufgaben zu einer unterschiedlich starken kognitiven Belastung und Stressreaktion bei SuS?
- Inwieweit hängt die Art der Repräsentation in Testaufgaben mit der kognitiven Belastung, der Stressreaktion und der Testleistung von SuS zusammen?

### Methode

SuS (N=93; 17.15±1.7 Jahre) bearbeiteten einen zehnminütigen Test mit Fragen zum molekularen Aufbau der DNA (22 Items im offenen Format). Zu Beginn des Tests konnten sie vermeintlich zwischen zwei Testversionen wählen (Testaufgaben mit rein symbolischen Repräsentationen vs. parallele Testaufgaben mit symbolisch-textlichen Repräsentationen). Anschließend wurden die Probanden zufällig, also unabhängig von ihrer Präferenz, einer der beiden Testvarianten zugeordnet.

Nach der Testbearbeitung wurde die aufgabenbasierte kognitive Belastung (mental load) erfragt (sechs Items mit siebenstufiger Likert-Skala; Krell, 2017). Die subjektive Stressbelastung wurde mit Hilfe einer visuellen Analogskala (VAS; Luria, 1975) vor und nach dem Test, sowie 15 Minuten später erhoben. Zur Ermittlung der physiologischen Stressreaktion wurden Speichelproben gesammelt und deren Cortisolkonzentration analysiert (vgl. Minkley et al., 2014). Zudem wurde während der gesamten Testbearbeitung kontinuierlich die kardiovaskuläre Aktivität der Probanden mit Hilfe eines Brustgurtes gemessen.

### Ergebnisse

Es zeigt sich eine deutliche Präferenz für die Testversion mit rein symbolischen Repräsentationen (82%;  $p < .01$ ,  $d = 1.70$ ;  $\chi^2$ -test). Zudem zeigen die SuS bei dieser Testversion bessere Testleistungen ( $p = .005$ ,  $d = .60$ ; t-test) und einen niedrigeren mental load ( $p = .001$ ,  $d = .69$ ; t-test).

Die Probanden beider Versuchsgruppen unterscheiden sich nicht signifikant hinsichtlich der subjektiven Stressbelastung (VAS) sowie der physiologischen Stressreaktionen (Cortisolkonzentration, kardiovaskuläre Aktivität).

Während der Bearbeitung der zweiten Hälfte des Tests zeigt sich jedoch bei den Probanden, die mit symbolisch-textlichen Repräsentationen arbeiteten, eine deutlich höhere kardiovaskuläre Stressbelastung. Kurz vor Testende (in den letzten Minuten) ist dieser Unterschied signifikant ( $p < .01$ ,  $d = -.70$ ; t-test). In diesem Teil des Tests waren im Vergleich zur ersten Hälfte komplexere Aufgaben zu bearbeiten (Erklärung von Prozessen vs. Benennung von Strukturen). Zudem weist eine durchgeführte Mediationsanalyse auf einen indirekten Effekt der Repräsentationsart auf die Testleistung hin, welcher über den mental load vermittelt wird.

Die höhere kognitive Belastung und die stärkere Stressreaktion bei der Bearbeitung von komplexeren Aufgaben mit symbolisch-textlichen Darstellungen unterstützt die Ergebnisse früherer Untersuchungen, welche Lernenden insgesamt große Schwierigkeiten mit molekularen Repräsentationen attestieren (vgl. Taskin & Bernholt, 2012). Aus unterrichtsmethodischer Perspektive stellt sich somit die Frage, wie der Umgang mit molekularen Repräsentationen im naturwissenschaftlicher Unterricht gestaltet werden kann, damit er zwar als heraus-, jedoch nicht als über- oder unterfordernd erlebt wird, und somit die von Bandura (1977) postulierten mastery experiences ermöglicht.

## D05

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Rosshofgasse, S01

### **Einfluss epistemologischer Überzeugungen auf die Reflexionskompetenz in verlängerten Praxisphasen während des Studiums**

**Judith Schellenbach-Zell, Michael Rochnia**

Universität Wuppertal, Deutschland; [zell@uni-wuppertal.de](mailto:zell@uni-wuppertal.de), [m.rochnia@uni-wuppertal.de](mailto:m.rochnia@uni-wuppertal.de)

#### 1. Theoretischer Hintergrund

Die Lehramtsausbildung in Deutschland begreift verlängerte Praxisphasen während des Studiums („Praxissemester“) als Gelegenheit, Theorie und Praxis miteinander zu verbinden. Das bedeutet für die universitäre Begleitung des Praxissemesters, dass entsprechende Reflexionsprozesse veranlasst, angeleitet und unterstützt werden müssen. Lerntagebücher stellen in diesem Zusammenhang ein geeignetes Instrument dar, zum einen zur schriftlichen Fixierung der Reflexionen und zum anderen zur Überwachung des eigenen Lernprozesses (Rambow & Nückles, 2002). An der Bergischen Universität Wuppertal werden die Praxissemester-Studierenden in der bildungswissenschaftlichen Begleitung dazu angehalten, für sie relevante schulische Situationen auszuwählen und zu reflektieren. Zur Erklärung von schulischen Situationen soll, im Sinne einer evidenzorientierten Pädagogik (vgl. Stark, 2017), nicht nur auf Erfahrungen und Einstellungen rekuriert werden, sondern explizit wissenschaftliches Wissen verwendet werden. Die Verbindung von Theorie und Praxis basiert im vorliegenden Beitrag daher auf der Nutzung von wissenschaftlichem Wissen für die Erklärung von schulischen Situationen und für die Entwicklung von Handlungskonsequenzen. Studierende tendieren jedoch dazu schulische Situationen eher auf der Grundlage subjektiver Theorien, statt wissenschaftlichen Wissens, zu reflektieren (Artmann, Herzmann, Hoffmann, Prose, 2013; Hatton & Smith, 1995; Schlag & Hartung-Beck, 2016). Mit Blick auf den Anspruch der Evidenzorientierung ist dies problematisch und birgt die Gefahr, dass nicht abgesichertes Wissen für schulisches Handeln zum Einsatz kommt. Eine mögliche Erklärung für das Verhaften in subjektiven Theorien könnten epistemologische Überzeugungen der Studierenden sein. Epistemologische Überzeugungen umfassen Kognitionen von Studierenden zu Wissen im pädagogischen Bereich (Hofer, 2000). Sie geben Einblick in die wahrgenommene Komplexität des Wissens sowie in die Nutzung und Bewertung möglicher Quellen des Wissens. Eine naive Epistemologie beinhaltet die Überzeugungen, dass durch eigenes Erleben sicheres Wissen über einen Sachverhalt produziert werden kann (Hofer, 2000). Für Studierende mit dieser Überzeugung erscheint die Konsultation von wissenschaftlichem Wissen wenig sinnvoll. Eine sophistische epistemologische Überzeugung, welche Wissen und seine Genese komplexer betrachtet, könnte vice versa die Konsultation von wissenschaftlichem Wissen nahelegen (vgl. Kienhues, Bromme & Stahl, 2008).

#### 2. Fragestellung

Vor dem dargestellten Hintergrund gehen wir der Frage nach möglichen Zusammenhängen zwischen epistemologischen Überzeugungen und der Nutzung von Wissen bei der Reflexion pädagogischer Situationen nach. Dabei gehen wir davon aus, dass Studierende mit sophistizierten epistemologischen Überzeugungen eher auf der Grundlage von wissenschaftlichem Wissen reflektieren, während Studierende mit eher naiven epistemologischen Überzeugungen stärker auf ihre subjektiven Theorien rekurrieren.

#### 3. Methode

Rund 260 Studierende wurden im Rahmen von 17 bildungswissenschaftlichen Begleitkursen angehalten, zu drei Zeitpunkten Lerntagebucheinträge mit schriftlichen Reflexionen von für sie relevanten pädagogischen Situationen einzureichen. Zum dritten Messzeitpunkt reflektierten die Studierenden zu einer standardisierten Situation im Sinne einer Fallvignette, die reichhaltige wissenschaftliche Bezüge ermöglicht. Die Reflexionen wurden von vier Personen hinsichtlich des genutzten Wissens eingeschätzt. Diese Ratings weisen eine ausreichende Interraterreliabilität auf. Die epistemologischen Überzeugungen wurden in einem Fragebogen mit den entsprechenden Skalen von Hofer (2000) erhoben. Reliabilitätsanalysen verdeutlichen, dass sich die Items zu den epistemologischen Überzeugungen theoriekonform skalieren lassen. In Vorbereitungsseminaren erhielten alle Studierenden eine Instruktion dazu, wie pädagogische Situationen durch wissenschaftliche Theorien reflektiert werden sollen, um eine vergleichbare Theorienutzungscompetenz (Klein, Wagner, Klopp & Stark, 2015; Ohlsson, 1992) aller Studierenden sicherzustellen. Für die Datenanalyse wurden die Ratings der Reflexion in quantitative Daten transformiert und mit den epistemologischen Überzeugungen sowie Kontrollvariablen (Alter, Geschlecht, Schulform und Dozierende des Vorbereitungskurses) in Regressionsmodellen betrachtet.

#### 4. Ergebnisse

Erste Auswertungen zeigen, dass epistemologische Überzeugungen mit dem erworbenen wissenschaftlichen Wissen (PUW, vgl. König & Blömeke, 2009) in Zusammenhang stehen. Gerade Studierende mit sophistizierten epistemologischen Überzeugungen weisen eine breitere Wissensbasis auf. Dieser Befund legt nahe, dass im Rahmen der schriftlich fixierten Reflexionen in den Lerntagebüchern stärker auf das erworbene Wissen zurückgegriffen wird. Entsprechende Analysen werden im Beitrag vorgestellt und vor dem Hintergrund des Forschungsstandes diskutiert.

### **Effekte einer Nützlichkeits-Wert-Induktion bei der evidenzorientierten Bearbeitung schulischer Problemsituationen bei Lehramtsstudierenden**

**Lisa Stark<sup>1,2</sup>, Martin Klein<sup>1</sup>, Jennifer Uhlenbrock<sup>1</sup>, Robin Stark<sup>1</sup>**

<sup>1</sup>Universität des Saarlandes, Deutschland; <sup>2</sup>Hochschule für Musik Saar, Deutschland; [lisa.stark@uni-saarland.de](mailto:lisa.stark@uni-saarland.de)

Ausgehend von Forderungen nach einer evidenzorientierten pädagogischen Praxis (Stark, 2017; KMK, 2014) stellt evidenzorientiertes Erklären und Lösen von Problemsituationen eine wichtige Professionalisierungskomponente angehender Lehrkräfte dar (KMK, 2014). Allerdings besitzen (angehende) Lehrkräfte eher negative Einstellungen gegenüber bildungswissenschaftlichem Wissen (Stark et al., 2010) und ziehen dieses nur selten in Problemsituationen heran (Hetmanek et al., 2015). Durch den Einsatz fehlerbasierten Lernens (Oser & Spychiger, 2005) im Kontext integrierter Lernumgebungen (Reinmann & Mandl, 2006) konnten bereits Kompetenzen zum evidenzorientierten Erklären schulischer Situationen gefördert werden (z.B. Klein et al., 2017; Wagner et al., 2014).

Die vorliegende Studie untersucht, inwieweit die Effektivität dieses fehlerbasierten Ansatzes durch die Induktion eines hohen Nützlichkeits-Wertes (Eccles & Wigfield, 2002) gesteigert werden kann. Im Fokus standen Effekte auf den Erwerb der Kompetenz zum evidenzorientierten Erklären und Lösen von Problemsituationen sowie auf Einstellungen zu Theorien.

In einem einfaktoriellem experimentellen Zwei-Gruppen-Design bearbeiteten N=70 Lehramtsstudierende (68.6% weiblich, Alter: M=23.34, SD=3.29) eine integrierte Lernumgebung. Die Studierenden wurden bei der Bearbeitung von Fallvignetten mit typischen Fehlern konfrontiert, welche beim evidenzorientierten Erklären und Lösen schulischer Problemsituationen auftreten (vgl. Oser & Spychiger, 2005). Durch Identifikation und Nachvollzug dieser Fehler sollte die Kompetenz zum evidenzorientierten Erklären und Lösen von Problemsituationen und Fehlervermeidungswissen erworben werden.

Der experimentelle Faktor Wertinduktion wurde mithilfe eines Textes vor dem Lernen umgesetzt. Die Gruppe mit Induktion eines hohen Nützlichkeits-Wertes (EG) erhielt einen Text, der die Nützlichkeit der Lernumgebungsinhalte für den späteren Lehrberuf herausstellte. Die Kontrollgruppe (KG) las einen Text ähnlicher Länge, welcher andere Lernumgebungsthemen vorstellte. Abgesehen von der Variation der Wertinduktion vor der Lernphase war die Lernumgebung für beide Gruppen identisch.

Erhoben wurden konzeptuelles Wissen über Fehler, strategisches Fehlervermeidungswissen, prozedurales Problemlösewissen und situationales Fehlerwissen (De Jong & Ferguson-Hessler, 1996) sowie die Kompetenz zum evidenzorientierten Erklären und Lösen einer Problemsituation. Die wahrgenommene Nützlichkeit der Lerninhalte wurde als Baseline (t1), als Manipulationscheck nach der Wertinduktion (t2) sowie nach Bearbeitung der Lernumgebung (t3) anhand einer Drei-Item-Skala eingeschätzt. Einstellungen zu Theorien wurden anhand einer 16-Item-Skala vor und nach dem Lernen erhoben. Als Kontrollvariablen wurde neben den Baseline-Messungen das Vorwissen bezüglich konzeptuellem und strategischem Wissen sowie der Kompetenz zum evidenzorientierten Erklären und Lösen einer Problemsituation erfasst. Alle Skalen waren hinreichend reliabel.

Es gab keine signifikanten Gruppenunterschiede in den Kontrollvariablen,  $p > .10$ . Eine MANOVA mit den fünf Wissensdimensionen als AVs und Gruppe als UV zeigte einen signifikanten Gruppenunterschied zugunsten der EG,  $F(5,64)=2.86$ ,  $p=.021$ ,  $\eta^2=.183$ . Univariate Analysen zeigten die Überlegenheit der EG in den Tests zum strategischen Fehlervermeidungswissen,  $F(1,68)=4.07$ ,  $p=.048$ ,  $\eta^2=.056$ , dem situationalen Fehlerwissen,  $F(1,68)=9.39$ ,  $p=.003$ ,  $\eta^2=.121$ , und im Kompetenztest,  $F(1,68)=3.99$ ,  $p=.050$ ,  $\eta^2=.055$ . Bezüglich konzeptuellem Fehlerwissen und prozeduralem Problemlösewissen gab es keine signifikanten Gruppenunterschiede,  $F_s < 1$ . In der wahrgenommenen Nützlichkeit der Lerninhalte unterschieden sich die Gruppen zu keinem der Messzeitpunkte signifikant,  $p > .15$ . Eine RM-ANOVA mit drei Messzeitpunkten und Gruppe als Zwischensubjektfaktor zeigte keinen Effekt der Gruppe und keinen Interaktionseffekt der Gruppe mit dem Messwiederholungsfaktor, beide  $F_s < 1$ . Der Effekt des Messwiederholungsfaktors,  $F(2,132)=3.30$ ,  $p=.040$ ,  $\eta^2=.048$ , deutete auf einen Anstieg der wahrgenommenen Nützlichkeit über die Dauer der Lernphase hin. Eine weitere RM-ANOVA zeigte, dass sich die Gruppen nicht signifikant in ihren Einstellungen zu Theorien unterscheiden,  $F < 1$ , auch der Interaktionseffekt mit dem Messwiederholungsfaktor war nicht signifikant,  $F(1,64)=2.43$ , n.s.. Der Haupteffekt des Messwiederholungsfaktors zeigte, dass sich Einstellungen gegenüber Theorien insgesamt verbessert haben,  $F(1,64)=10.12$ ,  $p=.002$ ,  $\eta^2=.137$ .

Die Ergebnisse deuten darauf hin, dass die Induktion eines hohen Nützlichkeits-Wertes evidenzorientiertes Erklären und Lösen von Problemsituationen unterstützen kann. Eine Verbesserung der Einstellungen zu Theorien und ein Anstieg der wahrgenommenen Nützlichkeit der Lerninhalte zeigten sich unabhängig von der Wertinduktion. Durch die für die Lehramtsstudierenden klar erkennbare Relevanz der Lerninhalte hatte die Wertinduktion auf diese subjektiven Maße möglicherweise keinen zusätzlichen Einfluss. Um dennoch die Unterschiede in der Lernleistung erklären zu können, stehen Prozessdatenanalysen an.

## **Wissensverknüpfung „relevant machen“: Förderung der Wissensintegration bei angehenden Lehrkräften durch Relevanz-Instruktionen**

**Helene Rieche, Eileen Spitzmesser, Felicitas Biwer, Alexander Renkl**

Universität Freiburg, Deutschland; [helene.rieche@psychologie.uni-freiburg.de](mailto:helene.rieche@psychologie.uni-freiburg.de)

Fachdidaktisches Wissen und pädagogisch-psychologisches Wissen gelten als wichtige Bereiche des Lehrerwissens. Beide Wissensbereiche beziehen sich auf Konzepte zum Lehren und Lernen – einerseits für ein spezielles Fach, andererseits für das Unterrichten allgemein – und haben somit vielfältige Anknüpfungspunkte. Beispielsweise ist beim Umgang mit Schülerinnen und Schülern, die überzeugt sind, ein bestimmtes Fach wie Mathematik oder Musik „einfach nicht lernen zu können“, Wissen aus beiden Bereichen hilfreich: Fachdidaktisches Wissen dient dem Verständnis von fachtypischen Überzeugungen, während aus der pädagogisch-psychologischen Forschung Ideen für Interventionsmöglichkeiten abgeleitet werden können. Lehrkräfte sollten daher in der Lage sein, ihr Wissen sinnvoll zu verknüpfen und kombiniert anzuwenden. Ein Problem der Lehramtsausbildung besteht jedoch darin, dass diese Wissensbereiche meist getrennt voneinander vermittelt werden, was eine integrierte Anwendung erschwert. In bisherigen Studien (Harr, Eichler, & Renkl, 2014, 2015) wurde die Wissensintegration beispielsweise durch eine vernetzte Präsentation der Inhalte sowie durch spezielle Hinweise (Prompts) nach dem Lernen angeregt. Beide Ansätze erwiesen sich als effektiv, waren jedoch mit Nachteilen wie geringer Praktikabilität und erhöhtem Zeitaufwand verbunden. Eine praktikablere und effizientere Methode wäre eine kurze Instruktion vor dem Lernen, in welcher die Wissensintegration „relevant gemacht“ wird. Relevanz-Instruktionen unterstützen Lernende dabei, sich sinnvolle Lernziele zu setzen, ihre kognitiven Ressourcen dementsprechend einzusetzen und die durch die Instruktion relevant gemachten Inhalte besser zu lernen (s. Relevanz-Effekt; McCrudden & Schraw, 2006). In zwei experimentellen Studien untersuchten wir, inwiefern die Verknüpfung von fachdidaktischem und pädagogisch-psychologischem Wissen bei angehenden Lehrkräften durch eine Relevanz-Instruktion gefördert werden kann. Der Lerninhalt bezog sich dabei auf Fähigkeitsüberzeugungen von Schülerinnen und Schülern aus fachdidaktischer und pädagogisch-psychologischer Sicht.

In Studie 1 bearbeiteten 69 Mathematik-Lehramtsstudierende eine computergestützte Lernumgebung. In der Integrationsbedingung ( $n = 36$ ) erhielten die Lehramtsstudierenden eine Instruktion zu Beginn der Lernphase, in welcher die Wichtigkeit der Wissensintegration betont wurde. In der Kontrollbedingung ( $n = 33$ ) enthielt die Instruktion keine Hinweise auf Wissensintegration. Die Lehramtsstudierenden bearbeiteten einen konzeptuellen Vorwissenstest und sahen anschließend zwei 15-minütige Lehrvideos zu subjektiven Intelligenztheorien (pädagogisch-psychologisches Wissen; s. Dweck, 2000) und zu Einstellungen im Mathematikunterricht (mathematikdidaktisches Wissen; s. Di Martino & Zan, 2010). Im darauffolgenden Posttest wurde sowohl konzeptuelles Wissen als auch Anwendungswissen anhand von Fallvignetten erhoben. Bezüglich der kombinierten Anwendung beider Wissensbereiche gab es jedoch keine Unterschiede zwischen den Gruppen,  $F(1,67) = 0.13$ ,  $p = .724$ .

An Studie 2 nahmen 72 Musik-Lehramtsstudierende teil. Die Lernumgebung glich im Aufbau der zuvor beschriebenen, wobei alle fachdidaktischen Inhalte auf das Schulfach Musik zugeschnitten waren. Im fachdidaktischen Lehrvideo wurde das musikalische Selbstkonzept thematisiert (s. Spychiger, 2017). Im Unterschied zu Studie 1 erhielten die Lehramtsstudierenden in der Integrationsbedingung ( $n = 36$ ) nicht nur eine Relevanz-Instruktion zu Beginn der Lernphase, sondern zusätzlich eine

Erinnerung an die Instruktion vor der zweiten Lerneinheit. In der Kontrollbedingung ( $n = 36$ ) enthielt die Instruktion – wie zuvor – keine Hinweise auf Wissensintegration. Die Auswertung der Anwendungsaufgaben zeigte, dass die Lehramtsstudierenden in der Integrationsbedingung häufiger beide Wissensbereiche kombinierten als die Lehramtsstudierenden in der Kontrollbedingung,  $F(1,69) = 7.97, p = .007$ .

Insgesamt zeigt sich somit ein gemischtes Bild: Während eine einmalige Instruktion bei angehenden Mathematiklehrkräften keinen Effekt hatte (Studie 1), führte eine zweifach dargebotene Instruktion bei angehenden Musiklehrkräften zur erwarteten integrierten Wissensanwendung (Studie 2). Eine mögliche Erklärung könnte sein, dass solche Instruktionen mehrfach präsentiert werden müssen, um wirksam zu sein. In einer Folgestudie soll daher untersucht werden, ob der erwartete Effekt bei angehenden Mathematiklehrkräften auftritt, wenn die Instruktion wiederholt dargeboten wird. Darüber hinaus könnten genauere Erkenntnisse darüber, unter welchen Bedingungen Relevanz-Instruktionen wirksam sind, für die Gestaltung der Lehramtsausbildung von Bedeutung sein. Besonders im Hinblick auf den geringen Zeitaufwand und die hohe Praktikabilität stellen solche Instruktionen einen vielversprechenden Ansatz dar, um die Verknüpfung verschiedener Wissensbereiche zu unterstützen.

## **Erhebung des latenten Konstrukts „Interesse an Reflexion“ von Lehramtsstudierenden in quantitativen Studien: Ergebnisse einer Instrumentenanalyse**

**Katrin Kreilinger, Birgit Ziegler**

TU Darmstadt, Deutschland; [kreilinger@bp.tu-darmstadt.de](mailto:kreilinger@bp.tu-darmstadt.de)

Ziel ist die Untersuchung des Zusammenhangs von Einstellungen zur Reflexion und dem Belief an die Relevanz der Persönlichkeit für den Berufserfolg von Lehrkräften (vgl. Hertrampf & Herrmann 1999) bei Studienanfänger\*innen im Lehramt. Es werden Ergebnisse einer laufenden Panelstudie zur Studieneingangsphase (1. Studienjahr) berichtet. Studierende werden über die Eingangsberatung, praktikumsbegleitende Seminare sowie die Arbeit mit Kompetenzportfolios für die Bedeutung der Reflexionsfähigkeit im Lehramt sensibilisiert und zur Entwicklung von Reflexionskompetenz angeleitet. Um Effekte dieser Maßnahmen zu untersuchen, wird ein Instrument genutzt, das verschiedene Aspekte der Einstellung zur Reflexion abbildet.

Theoretische Grundlage dieses Instruments bildet das Modell reflexiven Handelns von Griffiths & Tann (1992). Die Autorinnen unterscheiden mehrere, hierarchisch geordnete Ebenen von Reflexion, die sich anhand ihrer Denk- und Handlungsprozesse unterscheiden lassen. Level 1-2 findet als Reflection-in-action statt, Reflection-on-action (Level 3-5) dagegen nach einer kritischen Situation. Mit jedem Level steigt das Niveau an Systematik, sowohl im Prozess der Informationssammlung als auch -analyse. Bis Level 4 können subjektive Theorien/Beliefs die Grundlage von Analysen im Reflexionsprozess bilden. Nur auf Level 5, Rethorising and reformulating, werden Beliefs selbst Gegenstand des Reflexionsprozesses, indem sie wissenschaftlichen Theorien gegenübergestellt und infolgedessen verändert werden.

Basierend auf diesem Modell entwickelte Niggli (2004) ein Instrument, bestehend aus drei Skalen, die sich auf Reflection-on-action beziehen: Interesse an Reflexion (5 Items) bildet die allgemeine Bereitschaft zur Reflection-on-action ab. Die Einschätzung Studierender zur Relevanz wissenschaftlicher Theorien für die Lehrtätigkeit wird in der Skala Theoriedistanz (6 Items) erfasst. Eine geringe Ausprägung auf dieser Skala bildet die Voraussetzung für Reflexion auf Level 5. Die dritte Skala, Pragmatismus (5 Items), erfragt die Tendenz, vorwiegend aus praktischen Erfahrungen berufliche Kompetenz erwerben zu wollen. Auf Grundlage theoretischer Überlegungen, ebenfalls basierend auf Griffiths & Tann (1992), Ergebnissen eigener Datenanalysen ( $N = 174$ ) sowie von Niggli (2004) berichteten Ergebnissen, wurden einige Änderungen am Modell vorgenommen: Die Skala Pragmatismus wurde aufgrund inhaltlicher Skepsis und geringer Skalengüte aus dem Instrument entfernt. Die ausgeprägte Zustimmungstendenz der Skala Interesse an Reflexion ( $M: 4.89, SD: .57, \alpha .55$ , Skala 1 – 6, eigene Analysen; auch Niggli 2004, 352f) soll durch zusätzliche Items verringert werden. Während die Skala Theoriedistanz die Bewertung der Relevanz wissenschaftlicher Kenntnisse im Lehrerberuf zufriedenstellend abbildet ( $\alpha .79$ , eigene Analysen; auch Niggli 2004, 353), fehlt eine Bezugnahme auf die Integration wissenschaftlicher Theorien in Reflexionsprozesse. Daher wurde eine zusätzliche Skala Bereitschaft zur systematischen Informationssammlung und -analyse (3 Items) konstruiert, anhand derer die Bereitschaft anspruchsvolle Reflexionsprozesse zu vollziehen, erfasst werden soll. Zusätzlich wurde als weitere Komponente der Belief in die Lehrerpersönlichkeit (5 Items,  $\alpha .70$ ) in das Modell aufgenommen. Damit wird eine unter Lehrkräften verbreitete Haltung erfasst, nach der insbesondere angeborene Charakteristika über den Berufserfolg entscheiden (vgl. Hertrampf & Herrmann 1999). Beliefs können, wie beschrieben, Gegenstand von Reflexion sein (vgl. Griffiths & Tann 1992). Eine Verringerung des Beliefs in die Lehrerpersönlichkeit im Zeitverlauf dient somit als Indikator für den Ablauf anspruchsvoller Reflexionsprozesse. Ein abnehmender Belief in die Lehrerpersönlichkeit sollte durch steigendes Interesse an Reflexion und steigende Bereitschaft zur systematischen Informationssammlung und -analyse vorhergesagt werden können. Zugleich sollte ein stark ausgeprägter Belief in die Lehrerpersönlichkeit mit einer markanten Theoriedistanz einhergehen (vgl. Cramer 2013).

Unsere Forschungsfragen lauten: Lässt sich die postulierte Faktorstruktur empirisch abbilden? Konnte der Zustimmungsbias der Skala Interesse an Reflexion verringert werden? Lassen sich die postulierten Zusammenhänge, die sich in Vorstudien zeigten, mit dem erweiterten Modell replizieren?

Datengrundlage bilden zwei Studienanfängerkohorten (2016, 2017) der TU Darmstadt im Lehramt (Gymnasium, berufliche Schulen), die zu zwei Messzeitpunkten schriftlich befragt wurden. Datenerhebungen finden von Oktober-November 2017 statt.

Skalenanalysen umfassen explorative Faktorenanalysen und Reliabilitätsanalysen (Cronbachs  $\alpha$ ). Anhand einer Teilstichprobe werden erste Ergebnisse der Testung der Konstruktvalidität berichtet. Zur Analyse der Zusammenhänge zwischen den Skalen werden Ergebnisse von Korrelationsanalysen präsentiert.



## D06

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus S212

### Untersuchung zur professionellen Unterrichtswahrnehmung der Studierenden im Lehr-Lern-Labor Seminar

**Florian Treisch, Thomas Trefzger**

Universität Würzburg, Deutschland; [florian.treisch@physik.uni-wuerzburg.de](mailto:florian.treisch@physik.uni-wuerzburg.de)

#### Theoretischer Hintergrund

Die Professionelle Unterrichtswahrnehmung (PU) ist die Fähigkeit einer Lehrkraft, relevante Unterrichtssituationen zu erkennen und diese theoriebezogen zu begründen (van Es & Sherin, 2002). Sie gilt als zentrale Voraussetzung, im Unterricht professionell zu handeln (Seidel & Stürmer, 2014) und dient als Bindeglied zwischen dem professionellen Wissen sowie den Erfahrungen einer Lehrkraft und dem Handeln, da sie auf Grundlage kognitiver Schemata Handlungsscripts, also mentale Repräsentationen von Handlungsabfolgen aktivieren können (Kopp & Mandl, 2005; Schwindt, 2008). Die PU lässt sich in die Dimensionen des „Noticing“ und des „Knowledge based Reasoning“ unterteilen. Das Noticing beschreibt die Fähigkeit relevante Unterrichtssituationen zu erkennen. Das Reasoning bezieht sich auf die Fähigkeit, das zugrunde liegende theoretische Wissen der Lehrkraft auf die Begründung der beobachteten Situation anzuwenden (Seidel & Stürmer, 2014). Die in dieser Studie berücksichtigten Theorien sind die Unterrichtsmerkmale „Zielorientierung“, „Lernbegleitung“ und „Lernatmosphäre“, die einen Einfluss auf das Kompetenzerleben, die Autonomie und die soziale Eingebundenheit und somit auf die Motivation der Schülerinnen und Schüler haben (Seidel & Shavelson, 2007; Deci & Ryan, 2002). Das Reasoning lässt sich nochmals in die Facetten Beschreiben, Erklären und Vorhersagen einteilen (Seidel & Stürmer, 2014).

Basis der Untersuchung ist das Lehr-Lern-Labor Seminar (LLL) für Physik-Lehramtsstudierende an der Universität Würzburg. In diesem Seminar bereiten Studierende in der Vorbereitungsphase Experimentierstationen zu einem Themengebiet aus den bayrischen Lehrplänen vor. In der anschließenden Praxisphase besuchen Schulklassen das Seminar und durchlaufen die einzelnen Stationen in kleinen Gruppen. Dabei werden sie in einer Microteaching-Einheit von den Studierenden betreut. Die Betreuungen werden nach jeder Durchführung mit den Kommilitonen und zusammen mit den Dozenten reflektiert.

Studien zeigen, dass schon in der universitären Ausbildung die PU von angehenden Lehrkräften gefördert werden kann (Gold, Förster und Holodinski, 2016; Krammer et al., 2016; Sunder, Todorova & Möller, 2016, Santagata & Guarino, 2011). Dazu wurden meist Videoanalysen, teilweise in Kombination mit Praxisphasen im Rahmen eines Seminars durchgeführt (Stürmer, Seidel & Schäfer, 2013). Auf Grundlage dieser Erkenntnisse soll in dieser Studie untersucht werden, inwieweit das LLL mit ihrer iterativen Praxis kombiniert mit zusätzlichen theoriebezogenen Videoanalysen der eigenen Betreuungen, die PU der Lehramtsstudierenden fördern kann.

#### Fragestellung

1. Hat das Lehr-Lern-Labor Seminar einen Einfluss auf die PU der Lehramtsstudierenden?
2. Inwieweit beeinflusst das Lehr-Lern-Labor Seminar und eine zusätzliche Videoanalyse in der Praxisphase die PU der Studierenden?

#### Methode

Zur Beantwortung der Forschungsfragen wurden während den Betreuungen die Hälfte der Studierenden videografiert, sodass sie zusätzlich zwischen den Durchführungen Videoanalysen ihrer eigenen Betreuung und der ihrer Kommilitonen bezogen auf die Unterrichtsmerkmale in kleinen Gruppen fragenbasierend durchführen können. Insgesamt gibt es drei Gruppen, die miteinander verglichen werden können:

- LLL-Gruppe: Studierende, die das Seminar besuchen aber keine Videoanalysen durchführen
- LLLV-Gruppe: Studierende, die das Seminar besuchen und zusätzlich Videoclips analysieren
- Baseline: Physik-Lehramtsstudierende, die das Seminar nicht besuchen

Die PU wurde im Pre-Post-Design mit dem Observer-Tool gemessen (Seidel, Blomberg & Stürmer, 2010). Zu jedem Studierenden gibt es Messwerte bezüglich des Gesamttests, sowie der drei Dimensionen des Reasoning.

#### Ergebnisse

An der Studie nahmen 71 Physik-Lehramtsstudierende teil. Multiple Regressionsanalysen zeigen, dass die Zugehörigkeit zur LLLV-Gruppe im Gegensatz zur LLL-Gruppe einen signifikanten Einfluss auf den Post-Test des Gesamttestwertes der PU unter Kontrolle des Pre-Tests im Bezug zur Baseline ausübt:

abh. Variable PU (Post-Test),

unabh. Variablen:

PU (Pre-Test) ( $B=0,799$   $\beta=0,606$ ,  $p=0,000$ )

LLL-Gruppe ( $B=0,041$   $\beta=0,125$ ,  $p=0,289$ )

LLLV-Gruppe ( $B=0,082$   $\beta=0,239$   $p=0,045$ )

$R^2 = 0,390$

Die Dimensionen des Reasoning zeigen ähnliche Tendenzen, die im Vortrag vorgestellt werden. Außerdem werden in das Modell mögliche Einflussfaktoren, wie Unterrichtserfahrung oder besuchte Lehrveranstaltungen aufgenommen. Es deutet sich an, dass eine Kombination aus Praxisphase im LLL und eine theoriebezogene Videoanalyse der eigenen Betreuungen die PU der Studierenden verbessern kann.

### Professionalisierung in der einphasigen Waldorflehrerausbildung: Lerngelegenheiten im Schulfeld in Abhängigkeit von personalen Merkmalen der Mentoren

**Christian Reintjes, Kerstin Bäuerlein**

PH FHNW, Schweiz; [christian.reintjes@fhnw.ch](mailto:christian.reintjes@fhnw.ch)

#### Theoretischer Hintergrund

Angesichts einer zunehmenden Kompetenzorientierung wird in der Debatte zur Lehrerbildungsreform kritisch hinterfragt, ob die Lehramtsstudiengänge (staatlich/nicht-staatlich, ein-/zweiphasig) hinreichend auf das Berufsfeld ausgerichtet sind (u.a. Arnold et al. 2014). Studienergebnisse, die Aussagen darüber zulassen, wann und unter welchen organisatorischen Gegebenheiten Praktikumsphasen besonders erfolgreich sind und welche an Studierende gestellten Anforderungen besonders lernwirksam sind, sind aber kaum vorhanden.

Ungeachtet dessen sehen gerade Studierende die Praxisphasen als Ort des Aufbaus von Handlungskompetenzen, die ihnen den Berufseinstieg erleichtern. Von Universitätsseite besteht hingegen begründete Skepsis gegenüber rezepthaften Praxisroutinen, die sich einer theoretischen Fundierung und empirischen Wirkungsanalyse entziehen (u.a. Reintjes & Bellenberg 2016).

In der einphasigen Waldorflehrerbildung soll der Professionalisierungsprozess durch eine Stärkung der von Mentoren angeleiteten Reflexion, welche als wichtiger Mediator zwischen Wissen und Können angenommen wird, verbessert werden. Charakteristisch für die 4-jährige Lehrerbildung am Institut für Waldorf-Pädagogik (Witten) sind umfangreiche Praxisphasen. Die schulische Ausbildung und Betreuung der Praktikanten tragen das Ausbildungsinstitut und die Ausbildungsschulen gemeinsam.

Im vorliegenden Beitrag werden ausgewählte Befunde der längsschnittlichen Evaluationsstudie präsentiert, die eine multiperspektivische Bestandsaufnahme von Lerngelegenheiten leistet (12/2012-06/2016) und eine erste Bilanz der einphasigen, nicht-staatlichen Lehrerbildung ermöglicht. Im Fokus stehen unterrichtsbezogene Lerngelegenheiten, welche von den schulischen Mentoren unterbreitet werden. Anknüpfend an die LEK-Studie (König & Seifert, 2012) sowie die EMW-Studie (König & Rothland, 2013) wird als theoretischer Bezugsrahmen ein Mehrebenen-Modell genutzt.

Fragestellungen

1. Wie gestalten die schulischen Mentoren die (unterrichtsbezogenen) Lerngelegenheiten in den Praxisphasen?
2. Inwiefern beeinflussen personale Voraussetzungen der Mentoren die Gestaltung unterrichtsbezogener Lerngelegenheiten in Praxisphasen der Waldorflehrerbildung?

Methode

Zur Beantwortung dieser Fragestellungen wurde eine freiwillige Online-Befragung der Mentoren durchgeführt, die Folgendes enthielt:

1. Fragen zur eigenen Person: u.a. Geschlecht, Alter, Erfahrung als Lehrer, Erfahrung als Mentor/in
2. Skalen zur eigenen Persönlichkeit und Einstellung: u.a. FIT-Choice (Richardson & Watt, 2006), LPA (Brandstätter & Mayr, 1994), LIS (Mayr, 1998), Mentoren-Selbstwirksamkeit (Riggs 2000)
3. Fragen zu Rahmenbedingungen des Mentoring und Gestaltung der Lerngelegenheiten: u.a. Begleitung und Beratung im Praktikum, institutionalisierte Gesprächsanlässe im Unterricht unter Anleitung

Personenmerkmale der Mentoren und die Gestaltung der Lerngelegenheiten wurden deskriptiv analysiert. Spearman-Rang-Korrelationen dienten der Ermittlung von Zusammenhängen zwischen der Häufigkeit, mit der Inhalte in Vor-/Nachbesprechungen thematisiert wurden, und der Dauer der Mentoringgespräche. Auf Basis der Gesprächsdauer gebildete Gruppen (lange vs. kurze Gespräche) wurden mittels t-Test für unabhängige Stichproben hinsichtlich der Häufigkeit der Thematisierung verschiedener Inhalte verglichen. Zusammenhänge zwischen Personenmerkmalen und der Gestaltung der Lerngelegenheiten wurden mittels Korrelationen überprüft. Gemäss einer Clusteranalyse (Ward-Methode, quadrierte Euklidische Distanz) basierend auf verschiedenen Personenmerkmalen wurden zwei Mentorengruppen gebildet, welche mittels t-Test für unabhängige Stichproben hinsichtlich der Gestaltung der Lerngelegenheiten verglichen wurden.

Ergebnisse

- In die Auswertungen flossen Angaben von 69 Mentoren (Ausschöpfungsquote: 55%) ein.
- Die Mentoren schätzten sich selbst als eher selbstkontrolliert, stabil und kontaktbereit ein. Auch die Mentoren-Selbstwirksamkeit scheint eher hoch ausgeprägt.
- Die beiden clusteranalytisch (basierend auf den Personenmerkmalen Erfahrung als Lehrer, Erfahrung als Mentor/in, Lehrerpersönlichkeit, Selbstreflexion und Mentoren-Selbstwirksamkeit) gebildeten Gruppen unterscheiden sich hinsichtlich aller Clustervariablen bedeutsam bis auf Kontaktbereitschaft. Gruppe 1 (n=26) hat auf allen Merkmalen eine höhere Ausprägung als Gruppe 2 (n=36).
- Hinsichtlich der Zeitstunden, welche die Mentoren nach eigener Schätzung für die Betreuung der Studierenden eingesetzt haben, unterschieden sich die Mentoren-Gruppen nicht ( $t_{57}=0.69$ ,  $p=.491$ ); Gruppe 1 berichtete aber erwartungskonform von längeren Vor- und Nachbesprechungen als Gruppe 2 ( $t_{55}=2.83$ ,  $p=.006$ ,  $d=0.76$  bzw.  $t_{58}=2.68$ ,  $p=.010$ ,  $d=0.71$ ).
- Insbesondere in Vorbesprechungen wurden in Gruppe 1 zudem für die Unterrichtsplanung relevante Themen außer Stundenablauf und Medienwahl signifikant häufiger thematisiert als in Gruppe 2. In Nachbesprechungen wurden lediglich Zielsetzungen der Stunde in Gruppe 1 häufiger thematisiert als in Gruppe 2 ( $t_{54}=2.03$ ,  $p=.047$ ,  $d=0.57$ ).

Die Ergebnisse werden in Hinblick auf das Professionalisierungskonzept des Institutes für Waldorfpädagogik diskutiert.

## **Theorie-Praxis-Verknüpfung im Lehr-Lern-Labor (LLL) als Ausgangspunkt erster Professionalisierungsschritte**

**René Dohrmann**

Frei Universität Berlin, Deutschland; [rene.dohrmann@fu-berlin.de](mailto:rene.dohrmann@fu-berlin.de)

Ausgangslage:

Damit Unterricht professionell durchgeführt werden kann, gilt der Erwerb fachdidaktischen Wissens als eine wichtige Gelingensbedingung (vgl. Baumert & Kunter, 2006; Borowski et al., 2010). Dies geschieht vornehmlich, wenn die angehenden Lehrkräfte genötigt sind, theoretische Wissensanteile in der Praxis umzusetzen (vgl. Fischler, 2008; Nölle, 2002; Gudmundsdottir, 1995). Die Wirklichkeit an den meisten Hochschulen bietet bisher jedoch nur sehr wenige solcher Lerngelegenheiten (vgl. Fischler, 2008), sodass es „offensichtlich nicht oder nur unzureichend [gelingt], die erwünschten Einstellungen zu sichern und eine tragfähige Handlungskompetenz zu entwickeln“ (Messner, 1999). Darüber hinaus erfährt ein Teil der angehenden Lehrkräfte beim Belegen von Praktika einen „Praxis- bzw. Realitätsschock“ (vgl. Dicke et al., 2016; Rabe et al., 2013; Merzyn, 2006; Messner, 1999; Tschannen-Moran, 1998). Deshalb fordern nicht nur Studierende (vgl. Weyland,

2014; Makrinus, 2013; Hascher, 2011; Hoppe-Graff et al., 2008), sondern auch Ausbildungslehrkräfte (vgl. Völker & Trefzger, 2010) sowie die Kultusministerkonferenz (vgl. KMK, 2004, 2008) eine verstärkte Praxisorientierung im Studium. Will man dieser Forderung nachkommen, müssen Praktika theoriegeleitet durchgeführt werden, da sonst die Möglichkeit einer Deprofessionalisierung besteht (vgl. Weyland, 2014; Hascher, 2011). Um darüber hinaus einem „Realitätsschock“ vorzubeugen ist es förderlich praktische Ausbildungsanteile in ihrer Komplexität sukzessive zu steigern (vgl. Krofta et al., 2013; Tschannen-Moran, 1998).

Es wird angenommen, dass LLL die o.g. Bedingungsfaktoren erfüllen, d.h. dass sie sowohl der Forderung nach praktischen Ausbildungsanteilen entgegen, die Ausprägung fachdidaktischen Wissens fördern und dem „Realitätsschock“ vorbeugen (vgl. Dohr-mann & Nordmeier, 2017, 2016; Krofta et al., 2013). Zusätzlich kann ein solches Format zur Verbesserung der Lehrkompetenzen der Teilnehmer\*innen führen (vgl. Gröschner et al., 2013).

Hypothesen und Methodik:

Auf Basis der o.g. Ziele des LLL sowie einer explorativen Vorstudie wurden theorie- und evidenzbasiert verschiedene Hypothesen aufgestellt, u.a.:

- H1: In einem „geschützten“ LLL-Setting kommt es nicht zum „Praxisschock“, sondern zu einer positiven Entwicklung der Selbstwirksamkeitserwartungen (SWE).
- H2: Die Teilnahme am LLL bewirkt einen Anstieg des fachdidaktischen Wissens bei den Teilnehmer\*innen.
- H3: Die Teilnahme am LLL bewirkt eine Verbesserung der (selbsteingeschätzten) Unterrichtskompetenzen bei den Studierenden.
- H4: Die Teilnahme am LLL führt zu positiven Einstellungen gegenüber dem Reflektieren unterrichtlicher Handlungsprozesse.

In der Hauptstudie wurden sowohl eine Fragebogenerhebung im Pre-Post-Design, als auch leitfadengestützte Interviews durchgeführt und inhaltsanalytisch ausgewertet. Die Resultate aus beiden Erhebungen wurden anschließend trianguliert.

Ergebnisse:

Die Ergebnisse zeigen erste Tendenzen in Bezug auf die Wirksamkeit der Lehrveranstaltung. So scheint der Komplexitätsreduzierte und geschützte Rahmen nicht nur zur Stabilisierung der SWE, sondern sogar zu einer leichten Steigerung in Bezug auf die Planung und Durchführung von Unterricht zu führen ( $d=0,5$ ). Dieses Ergebnis konnte durch die Auswertung der Interviews bestätigt werden. 87 % der Ratings fallen in die Kategorie „Zunahme der SWE“, 13 % entfallen auf „Keine Änderung der SWE“ und kein einziges Rating gibt es in der Kategorie „Abnahme der SWE“ (Cohen's Kappa: 0.68).

Positive Ergebnisse konnten auch beim fachdidaktischen Wissen in Bezug auf die Veranstaltungsinhalte festgestellt werden. Insgesamt wurden 480 Antworten ausgewertet und mit einer nahezu vollständigen Übereinstimmung kodiert (0.91). Der T-Test zeigt einen signifikanten Mittelwertunterschied bei einer hohen Effektstärke ( $d=0.87$ ). Die Selbsteinschätzung der Unterrichtskompetenzen zeigt im Pre-Post-Vergleich einen nicht signifikanten Zuwachs. Darüber hinaus scheint sich bei den Teilnehmer\*innen eine positive Einstellung gegenüber dem Reflektieren im Laufe des LLL zu entwickeln.

Die Befunde deuten auf erste Professionalisierungsschritte hin und zeigen Tendenzen in Bezug auf die Wirksamkeit des Veranstaltungskonzepts auf. Einschränkend ist zu sagen, dass die Stichprobe nicht zufällig bestimmt wurde und sehr klein ist. Dazu kommt, dass die Intervention sehr kurz ist und die Ergebnisse nur eine geringe Reichweite haben. Zwei weitere Erhebungen sind in Planung (N=70).

## **Peer-to-Peer-Mentoring in lokalen Arbeits- und Lerngemeinschaften – Potential partnerschaftlichen Lehrens und Lernens angehender und erfahrener Lehrpersonen im Schulfeld**

**Urban Fraefel, Kerstin Bäuerlein, Nils Bernhardsson-Laros**

Pädagogische Hochschule FHNW, Schweiz; [urban.fraefel@fhnw.ch](mailto:urban.fraefel@fhnw.ch)

Kontext und theoretischer Hintergrund

In den Praktika der Lehrpersonenbildung herrscht überwiegend die Rollenverteilung vor, wonach berufserfahrene Lehrpersonen die Studierenden begleiten, beraten und beurteilen, um den Aufbau erwünschten unterrichtlichen Handelns zu unterstützen. Hingegen sind wichtige weitere Dimensionen von Professionalität – Nutzung und Integration vielfältiger Wissensressourcen sowie die Fähigkeit kooperativen Arbeitens – zwar als Notwendigkeit erkannt, aber wenig entwickelt (Fraefel, 2012). Institutionelle Bedingungen leisten in der Regel eher einer asymmetrischen Beziehung von Studierenden und erfahrenen Lehrpersonen Vorschub.

Dem steht ein zunehmend diskutierter Ansatz gegenüber, wonach mentorierende Lehrpersonen (=Praxislehrpersonen) und Studierende in lokalen Arbeits- und Lerngemeinschaften auch eine reziproke Beziehung eingehen können (Wenger, 1998). Sie verstehen sich als „near-peers“ (Whitman, 1988), die sich der Erfüllung derselben Aufgabe – Lernen und Entwicklung der Schüler/innen zu fördern – widmen und sich dazu in einem gegenseitigen Co-Mentoring engagieren. Mit dieser Mentoring-Beziehung „auf Augenhöhe“ geht insbesondere für erfahrene Lehrpersonen ein Rollenwandel einher: Die evaluative Begleitung von Studierenden tritt in den Hintergrund, während das gemeinsame Bewältigen von realen Problemen des Berufsalltags beiderseitige Lernprozesse auslösen kann (Le Cornu, 2005; Fraefel & Bernhardsson-Laros, 2016; Gallo-Fox & Scantlebury, 2016).

Fragestellung

Diese Studie ist Teil der Interventionsstudie „Partnerschulen“ (Fraefel et al., 2017) und fokussiert die Kooperation angehender und erfahrener Lehrpersonen. Mikroteams aus jeweils einer Praxislehrperson und zwei Studierenden arbeiten im Partnerschuljahr über mehrere Praktika hinweg an der Qualität des professionellen Handelns und des Unterrichts, unterstützt von Mitarbeitenden der Hochschule. Um zu untersuchen, wie sich dieser Mentoring-Ansatz mit Einstellungen zusammenhängt, wurden folgende Forschungsfragen formuliert:

1. Besteht ein Zusammenhang zwischen der Intensität der Kooperation in Mikroteams und der Bereitschaft, Wissensressourcen zu aktivieren und zu verknüpfen?
2. Wie gehen Praxislehrpersonen mit der Erwartung um, dass sie sich in einem symmetrischen Peer-Mentoring mit Lehramtsstudierenden engagieren sollen?

Methode

Es wurden zwei komplementäre methodische Zugangsweisen gewählt.

In der quantitativen Erhebung zu Fragestellung 1 beantworteten angehende Lehrpersonen (N=82) in 10 Partnerschulen dreimal einen Online-Fragebogen, jeweils nach den Praktika im Partnerschuljahr. Einerseits wurden sie nach der Intensität der Zusammenarbeit innerhalb des Mikroteams befragt, andererseits enthielten die Fragebögen Skalen zur Einschätzung ihrer Bereitschaft, Wissensressourcen zu aktivieren und zu verknüpfen. Die kumulierten Werte der Kooperations-Skalen (Co-Planning, Co-Teaching, Team-Kohäsion, programmatische Kooperation) wurden mit den Wissensintegrationskalen (Theorie-Praxis-Integration und Nutzung von Wissensressourcen) zum Jahresende korreliert.

Die Fragestellung 2 wurde mit qualitativen Methoden untersucht. Leitfadengestützte Interviews (Helfferich, 2011) mit 27 Praxislehrpersonen aus sechs Schulen wurden in Anlehnung an die dokumentarische Methode (Nohl, 2013) ausgewertet. Die Methode konzentriert sich auf das implizite und integrierte Wissen, das dem menschlichen Handeln Orientierung gibt und sich in habituellen Handlungsmustern manifestiert. So können typische Strategien der Menschen identifiziert werden, mit denen sie den Herausforderungen des kooperativen Peer-to-Peer-Ansatzes begegnen.

#### Ergebnisse

ad 1: Zu allen drei Messzeitpunkten bestehen signifikante Zusammenhänge zwischen der Kooperation im Mikroteam und der Wissensintegration. Eine intensivere Kooperation im Mikro-Team (Co-Planning, programmatische Kooperation und Kohäsion) geht mit einer positiveren Einstellung bzw. stärkeren Orientierung bezüglich der Nutzung von Wissensquellen, der Theorieorientierung bei Entscheidungen im Unterricht und der Verknüpfung von Theorie und Praxis einher. Zum dritten Messzeitpunkt liegen die Korrelationswerte überwiegend im Bereich mittlerer Effektgrößen (Cohen, 1992), was für eine praktische Bedeutung der Befunde spricht. Diese Ergebnisse sind ein Hinweis auf das Potenzial der intensiven Kooperation im Peer-to-Peer-Mentoring, wenn es um die Einstellung und Orientierung bezüglich der Integration von theoretischem Wissen und praktischem Handeln geht.

ad 2: Ermittelt wurden drei unterschiedliche Typen, die beschreiben, wie Praxislehrpersonen auf die Konfrontation mit der neuen Rolle reagieren (1. „Schulmeister“, 2. „Lehrperson mit eigenem Stil“, 3. „Engagierte“). Lehrpersonen, die sich bei der Bearbeitung von Problemen des Schul- und Unterrichtsalltags auf eine intensive Peer-to Peer-Kooperation einlassen (Typus 3), sehen eher einen Gewinn in der Zusammenarbeit mit den Studierenden als solche, die das gewohnte Meister-Lehrlings-Verhältnis fortsetzen.

## D07

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Bernoullianum 32, H120

### Fit für digitales Lernen? ICT-Kompetenzniveaus angehender und aktuell Studierender

**Christian Schöber, Martin Senkbeil, Jan Marten Ihme**

Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Deutschland; [schoeber@ipn.uni-kiel.de](mailto:schoeber@ipn.uni-kiel.de)

#### Theoretischer Hintergrund

Digitales Lehren und Lernen gewinnt als integraler Bestandteil der Hochschullehre zunehmend an Bedeutung, beispielsweise gehören digitale Lehrformate wie Blended Learning bereits zum Hochschulalltag (Wannemacher, Jungermann, Scholz, Tercanli & Vielliz, 2016). Lange Zeit wurde vermutet, dass sich angehende Studierende die notwendigen computer- und informationsbezogenen Kompetenzen selbstständig aneignen und eine hohe Bereitschaft zur Nutzung digitaler Medien mitbringen (Kennedy, Judd, Churchward, Gray & Krause, 2008). Diese Annahmen widersprechen einer Reihe empirischer Befunde, wonach Studierende zwar seit frühester Kindheit mit digitalen Medien vertraut sind, aber nicht durchgängig über die notwendige ICT-Literacy für eine erfolgreiche Umsetzung ihrer beruflichen und persönlichen Zielvorstellungen verfügen (Lorenzo, Oblinger & Dziuban, 2006). Häufig nutzen sie digitale Medien sehr eingeschränkt und überwiegend zu privaten Zwecken (Unterhaltung, sozialer Austausch), sind aber nicht in der Lage, ihre zumeist autodidaktisch erworbenen Kenntnisse und Fertigkeiten auf studienbezogene Lerninhalte zu übertragen (Margaryan, Littlejohn & Voijt, 2011). Entsprechend haben Studierende in der Mehrzahl nur ein begrenztes Verständnis über innovative und digitale Lehr-Lernszenarien und weisen zudem häufig motivationale Defizite in Bezug auf die Verbesserung ihrer ICT-Literacy auf (Bullen, Morgan & Qayyum, 2011).

Vor diesem Hintergrund soll im Rahmen des vom BMBF geförderten Projekts CiBaS anhand repräsentativer Stichproben u.a. der Frage nachgegangen werden, über welche ICT-Kompetenzniveaus angehende (Ende Sekundarstufe II) und aktuell Studierende (sechstes Fachsemester) in Deutschland verfügen.

#### Fragestellungen

Konkret beschäftigt sich die vorgestellte Studie mit den folgenden empirischen Fragestellungen:

1. Wie groß sind die Anteile angehender Studierender und aktuell Studierender, die über rudimentäre, ausreichende oder weiterführende ICT-Literacy verfügen?
2. Ist der Anteil ohne ausreichende ICT-Literacy bei aktuell Studierenden geringer als bei angehenden Studierenden?
3. Lassen sich für a priori definierte Subgruppen (Studienfach, Geschlecht, Migrationshintergrund, Sozioökonomischer Status) ICT-Kompetenzniveauunterschiede finden?

#### Methode

Zur Beantwortung der Fragestellungen wurden Daten des Nationalen Bildungspanels (NEPS) herangezogen. Berücksichtigt wurden Personen, die den im NEPS eingesetzten ICT-Literacy-Test bearbeiteten und entweder der Startkohorte 4 (angehende Studierende, n = 3862) oder der Startkohorte 5 (aktuell Studierende, n = 1998) angehören. Zur Bestimmung von Kompetenzniveaus wurde ein Standard-Setting-Verfahren unter Anwendung der Item-Descriptor-Matching-Methode (Ferrara & Lewis, 2012) mit Expertinnen und Experten aus den Bereichen ICT-Literacy und digitales Lehren und Lernen an Schule und Hochschule durchgeführt. Für die Evaluation des Standard-Setting-Verfahrens wurden Gütekriterien nach Pitoniak (2003) herangezogen. Analysen weisen auf eine befriedigende prozedurale und internale Validität des Verfahrens hin.

#### Ergebnisse

In Bezug auf die angehenden Studierenden zeigt sich, dass ein substanzieller Anteil (20%) nur über rudimentäre ICT-Literacy verfügt. Für sie sind eine erfolgreiche Nutzung digitaler Lehr- und Lernformate an der Hochschule sowie eine erfolgreiche Umsetzung ihrer beruflichen Vorstellungen zumindest anzuzweifeln. Für die aktuell Studierenden beträgt der Anteil mit rudimentärer ICT-Literacy immerhin noch 4%. In beiden Gruppen war dabei der Anteil der Frauen etwas größer (23% bzw. 5%; Männer: 16% bzw. 2%). Ein ähnliches Muster zeigte sich hinsichtlich der Personen mit Migrationshintergrund im Vergleich zum Anteil der Personen ohne Migrationshintergrund sowie einem geringen oder hohen sozioökonomischen Status der Eltern (Mediansplit).

Die Anteile mit ausreichender ICT-Literacy (Basiskenntnisse, die ein erfolgreiches Weiterlernen ermöglichen) betragen 62% (angehende Studierende) bzw. 51% (aktuell Studierende). Weiterführende ICT-Literacy, die in etwa den von der Hochschulkonferenz (HRK, 2012) definierten Teilfertigkeiten der Informationskompetenz entspricht, weisen 18% der angehenden und 45% der aktuell Studierenden auf. Besonders die Studierenden der Sozialwissenschaften, Mathematik, Informatik und dem Ingenieurwesen mit Anteilen von jeweils über 50% auf weiterführendem Niveau zeigten sich gut auf digitale Herausforderungen vorbereitet.

Insgesamt weisen die Ergebnisse darauf hin, dass nicht alle angehenden Studierenden über die notwendige ICT-Literacy für eine erfolgreiche Nutzung digitaler Lehr- und Lernformate an der Hochschule verfügen und einer gezielten Schulung bedürfen. Weiterhin erreicht nur knapp die Hälfte der Studierenden im sechsten Fachsemester das von der HRK empfohlene Kompetenzniveau, was eine stärkere Vermittlung von ICT-Literacy in den universitären Fachbereichen als bisher empfiehlt.

### Effekte der Nutzung digitaler Medien in der Hochschullehre: Was sagen die Studierenden?

**Christina Wekerle, Martin Daumiller, Ingo Kollar**

Universität Augsburg, Deutschland; [christina.wekerle@phil.uni-augsburg.de](mailto:christina.wekerle@phil.uni-augsburg.de), [martin.daumiller@phil.uni-augsburg.de](mailto:martin.daumiller@phil.uni-augsburg.de)

#### Theoretischer Hintergrund

Digitalen Medien wird ein erhebliches Potenzial zugeschrieben, wenn es um die Förderung von individuellem und kollaborativem Lernen von Studierenden geht (vgl. Schmid et al., 2014). Allerdings ist anzunehmen, dass Hochschullehrende im deutschsprachigen Raum ähnlich wie Schullehrkräfte eher zu dozentenorientierter Lehre neigen (Balanskat, Blamire & Kefele, 2006) und digitale Medien wenn überhaupt auf eine Weise verwenden, die mit derartigen dozentenorientierten Lehrformen kompatibel ist (Schmotz, 2009). Entsprechend ist zu befürchten, dass Gelegenheiten zu einem Engagement in tiefergehenden Lernprozessen auch beim Einsatz von digitalen Medien nur selten gegeben sind (Seidel et al., 2002). Chi und Wylie (2014) zufolge werden tiefergehende Wissenskonstruktionsprozesse insbesondere dann gefördert, wenn Lernende zu konstruktiven (z.B. Online-Concept-Map erstellen) und interaktiven Lernaktivitäten (z.B. Online-Concept-Map in Kleingruppen diskutieren) anstatt aktiven (z.B. zu einem Onlinevideo Notizen machen) und passiven (z.B. Onlinevideo ansehen)

Lernaktivitäten bewegt werden. Inwiefern diese unterschiedlichen Lernaktivitätsklassen in Zusammenhang mit unterschiedlichen Lernzielvariablen stehen (z.B. kognitive, motivational-affektive, soziokommunikative und selbstregulative Fähigkeiten), ist jedoch noch weitgehend unklar. Der Beitrag geht deshalb der Frage nach, inwiefern Hochschullehrende die Potenziale digitaler Medien zur Anregung hochwertiger Lernaktivitäten nutzen und diese sich positiv auf unterschiedliche Dimensionen des subjektiven Lernerfolgs von Studierenden auswirken.

Fragestellungen

Entsprechend wurden die folgenden Fragestellungen untersucht:

(1) Inwiefern werden im Rahmen von Hochschullehrveranstaltungen in Veranstaltungsphasen, in denen digitale Medien (DM) eingesetzt werden, mehr hochwertige Lernaktivitäten (d.h. konstruktive und interaktive Lernaktivitäten) auf Studierendenseite angeregt als in Phasen, in denen keine digitalen Medien eingesetzt werden (ODM)?

(2) Inwiefern sind Lernaktivitäten, zu welchen Studierende in Veranstaltungsphasen mit digitalem Medieneinsatz angeregt werden, enger mit dem Erwerb kognitiver, motivational-affektiver, soziokommunikativer und selbstregulativer Fähigkeiten verbunden als Lernaktivitäten, die in Veranstaltungsphasen ohne digitalen Medieneinsatz gezeigt werden?

Methode

396 Lehramtsstudierende (durchschnittliches Alter: 23.08, SD = 3.82, Jahre; weiblich= 79,9%) nahmen an einer Paper-Pencil-Studie in 30 Seminaren teil. Die Studierenden wurden mithilfe von jeweils 21 Likert-basierten selbst entwickelten Items gefragt, zu welchem Ausmaß sie sich in Phasen mit und ohne Unterstützung digitaler Medien zu verschiedenen passiven, aktiven, konstruktiven und interaktiven Lernaktivitäten angeregt fühlen. Die addierten, gewichteten Mittelwerte (Multiplikationsfaktor: passiv=1, aktiv=2, konstruktiv=3, interaktiv=4) der jeweiligen Lernaktivitätsklassen wurden zu einem Lernaktivitätsscore für DM und ODM verrechnet. Ebenfalls sollten die Studierenden ihren subjektiven Lernerfolg hinsichtlich kognitiver, motivational-affektiver, sozio-kommunikativer und selbstregulativer Dimensionen mithilfe von insgesamt 20 selbst entwickelten Likert-skalierten Items einschätzen. Die internen Konsistenzen der Skalen zum Fähigkeitserwerb sind als gut bis sehr gut zu bewerten (Cronbachs Alpha: .81–.90). Die Auswertung erfolgte mit Mplus, wobei die genestete Datenstruktur durch type = complex berücksichtigt wurde (Muthén & Muthén, 2014).

Ergebnisse

(1) Es zeigte sich, dass sich Studierende in Phasen, in denen digitale Medien zur Anwendung kamen, signifikant häufiger zu höherwertigen Lernaktivitäten (d.h. mehr konstruktive und interaktive statt passive und aktive Aktivitäten) angeregt fühlten als in Phasen, in denen keine digitalen Medien eingesetzt wurden (Wald  $\chi^2=31.28$ ,  $p<.001$ ). (2) Eine Modellierung der Effekte der in Phasen mit vs. ohne Einsatz digitaler Medien berichteten Lernaktivitäten auf die vier abhängigen Lernzuwachsvariablen in einem Pfadmodell indizierte ferner, dass die in Phasen mit digitalem Medieneinsatz gezeigten Lernaktivitäten einen fast doppelt so großen, positiven Einfluss auf den selbsteingeschätzten Erwerb kognitiver ( $\beta_{DM}=.46$ ,  $\beta_{ODM}=.24$ ,  $SE=.05$ ), motivational-affektiver ( $\beta_{DM}=.46$ ,  $\beta_{ODM}=.24$ ,  $SE=.05$ ), sozio-kommunikativer ( $\beta_{DM}=.41$ ,  $\beta_{ODM}=.25$ ,  $SE=.05$ ) und selbstregulativer Fähigkeiten ( $\beta_{DM}=.46$ ,  $\beta_{ODM}=.28$ ,  $SE=.05$ ) hatten als diejenigen Lernaktivitäten, die während Veranstaltungsphasen ohne Einsatz digitaler Medien berichtet wurden.

Diskussion

Die Ergebnisse weisen darauf hin, dass aus Studierendensicht die Potenziale digitaler Medien von Hochschullehrenden zu einem deutlich größeren Ausmaß als angenommen genutzt werden und unterstreichen die Bedeutung digitaler Medien für Lehramtsstudierende und ihren subjektiven Wissenserwerb. Da die vorliegende Studie allerdings nur auf selbstberichteten Daten beruht, sollten weitere Beobachtungsstudien durchgeführt werden, um einen objektiveren Einblick in die Nutzung und Wirkung digitaler Medien in der Hochschullehre zu erhalten.

## **Der Zusammenhang von Lernmotivation und Nutzung von E-Quizen sowie deren Feedbackfunktion in einer Großveranstaltung – eine Untersuchung im Längsschnitt**

**Manuel Förster<sup>1</sup>, Andreas Maur<sup>2</sup>**

<sup>1</sup>Otto Friedrich Universität Bamberg, Deutschland; <sup>2</sup>Johannes Gutenberg-Universität, Mainz; [manuel.foerster@uni-bamberg.de](mailto:manuel.foerster@uni-bamberg.de)

Theorie und Fragestellung

Verschiedene Forschungsbefunde zeigten bereits, dass Studierende das Fach Statistik in den Sozialwissenschaften oftmals nur mit geringer Motivation beginnen und den Kurs nur mit geringem Selbstkonzept absolvieren (Förster & Maur, 2015; Hanna, Shevlin & Dempster, 2008). In dem Vortrag wird dabei unter der Motivation die Verknüpfung von Erwartungs- und Wertkomponenten nach Wigfield und Eccles (2002) verstanden, welche sich auch in empirischen Studien auf das fachstatistische Leistungsverhalten der Studierenden auswirkt (Macher et al., 2012, Tempelaar, 2004). Hierbei konstituieren sich einerseits die Erfolgserwartungen aus den fachspezifischen Fähigkeitsselbstkonzepten als auch aus der subjektiven Einschätzung der Aufgabenschwierigkeit, welche gemeinsam den verhaltenswirksamen Einfluss der Erwartungskomponenten auf das Leistungsverhalten darstellen (Jerusalem, 2002). Andererseits wird mittels der Wertkomponenten der persönliche Anreiz einer Aufgabenbearbeitung und die damit verbundenen Kosten, den Nutzen und das Interesse ausgedrückt (Wigfield et al., 2009).

Die zwar vereinzelt nachgewiesene zeitliche Variabilität der Motivation zur Statistik im Laufe eines Semesters (Chiesi & Primi, 2010; Ramirez, Schau & Emmioglu, 2012) deutet dabei jedoch auf die bislang kaum empirisch untersuchte Frage nach möglichen veranstaltungsimmanenten Stellgrößen solcher Veränderungen im Längsschnitt. In der vorliegenden Studie werden hierfür regelmäßige E-Quize herangezogen, welche sich gerade in statistischen Großveranstaltungen eignen, um Studierenden automatisierte Feedbacks zu geben und diese zur frühzeitigen Auseinandersetzung mit den Inhalten und Selbsteinschätzung zu motivieren sowie diesen frühzeitig den Wert statistischer Fachinhalte und die Lösbarkeit statistischer Aufgaben zu verdeutlichen. An diesen genannten Forschungslücken anknüpfend wird in dem Beitrag der Zusammenhang zwischen der Motivation der Studierenden zum Fach Statistik sowie der quantitativen und qualitativen Teilnahme der E-Quize im Längsschnitt untersucht.

Stichprobe/ Instrument

In einer längsschnittlichen Studie im Sommersemester 2017 wurden die Studierenden der Wirtschaftswissenschaften in einer Statistikgroßveranstaltung zu Beginn ( $n = 562$ ), in der Mitte ( $n = 578$ ) und zum Ende des Semesters ( $n = 412$ ) zu Ihrer Motivation im Fach Statistik befragt. Die Motivation wurde dabei mittels des Survey of Attitudes Towards Statistics-36 (SATS-36; Schau, 2003) erfasst. Die Erwartungskomponente der Motivation (Wigfield & Eccles, 2002) wurde dabei mittels der

wahrgenommenen Schwierigkeit statistischer Aufgaben (Difficulty) sowie dem statistikbezogenen Selbstkonzept (Cognitive Competence) operationalisiert. Die Wertkomponenten werden anhand des Interesses (Interest), dem wahrgenommenen Nutzen (Value) sowie der Einstellung (Affect) zum Fach Statistik als auch anhand der mit der Aufgabenbewältigung verbundenen Kosten (Cost; Trautwein et al., 2012) gemessen.

Das Leistungsverhalten wird mittels der Teilnahme an den E-Quizen und den dabei erzielten Ergebnissen erfasst. Dazu bearbeiten die Studierenden vier E-Quizze während des Semesters ( $n_1 = 675$ ,  $n_2 = 623$ ,  $n_3 = 606$ ,  $n_4 = 595$ ). Jedes Quiz beinhaltete vier bis fünf Aufgaben, die in den jeweils davorliegenden Statistikvorlesungen behandelt wurden. Die Teilnehmenden erhielten nach jedem Quiz ein automatisiertes Feedback zu den gegebenen und tatsächlich richtigen Antworten der Multiple-Choice- und numerischen Lücken sowie einen Verweis zur Lösungsmethode. Zudem wurden Geschlecht, Note der Hochschulzugangsberechtigung, Mathematiknote, etc. kontrolliert.

Die Daten der Studierenden aus den verschiedenen Befragungszeitpunkten wurden administrativ über einen anonymisierten, eindeutigen Benutzercode verknüpft. Insgesamt konnte die Motivationsentwicklung von 331 Studierenden vollständig über alle Zeitpunkte verknüpft werden und weitere Daten von zahlreichen Studierenden liegen über zwei Zeitpunkte vor. 504 Studierende nahmen an allen vier E-Quizen und 108 weitere Studierende an drei E-Quizen teil.

#### Analyseverfahren/ Ergebnisse

Über autoregressive Strukturgleichungsmodelle wird untersucht, in welchem Maße sich die Erwartungs- und Wertkomponenten und die quantitative und qualitative Teilnahme an den E-Quiz gegenseitig bedingen. Dabei werden weitere relevante Kovariate wie Geschlecht, Vornoten, Migrationshintergrund und Bildungshintergrund kontrolliert. Es wird insbesondere der Frage nachgegangen, ob sich die E-Quiz in hochschulischen Großveranstaltungen, in denen Feedback zum individuellen Lernstand in üblichen traditionellen Arrangements zumeist ausbleibt, zur motivationalen Förderung der Lernprozesse im Laufe des Semesters eignen.

## D08

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Bernoullianum 32, H103

### **Aufgabenbearbeitungsprozesse als Indikator für die Trennbarkeit der PIAAC Kompetenzen von kognitiver Grundfähigkeit**

**Lena Engelhardt<sup>1</sup>, Frank Goldhammer<sup>1,2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland; <sup>2</sup>Zentrum für internationale Vergleichsstudien (ZIB); [engelhardt@diipf.de](mailto:engelhardt@diipf.de)

Inwiefern Kompetenzaufgaben mehr als kognitive Grundfähigkeit erfordern ist eine andauernde Debatte in der Bildungsforschung (vgl. Baumert, Lüdtke, Trautwein, & Brunner, 2009). Das Ziel dieser Studie ist es anhand eines Indikators des Aufgabenbearbeitungsprozesses - der Bearbeitungszeit - zu untersuchen, ob die PIAAC Kompetenztests für Lesekompetenz und alltagsmathematische Kompetenz etwas anderes als kognitive Grundfähigkeit abbilden.

#### Theoretischer Hintergrund

Der Zusammenhang zwischen Bearbeitungszeit und Aufgabenerfolg beschreibt einen grundsätzlichen Aspekt des Bearbeitungsprozesses. Bisherige Forschung zeigt, dass dieser Zusammenhang von Personenfähigkeit und Itemleichtigkeit abhängt und er für fähigere Personen und leichtere Items negativer wird (vgl. Goldhammer, Naumann, Stelter, Toth, Rölke, & Klieme, 2014). Personen mit einer höheren Fähigkeit lösen Aufgaben demnach eher wenn sie diese in kürzerer Zeit beantworten, das heißt eine längere Bearbeitungszeit ist nicht mit einer höheren Wahrscheinlichkeit, die Aufgabe zu lösen, verbunden. Personen mit niedrigerer Fähigkeit lösen dahingegen Aufgaben eher wenn sie mehr Zeit verwenden. Personenvariablen, für die der Zusammenhang zwischen Bearbeitungszeit und Aufgabenerfolg für eine höhere Ausprägung negativer wird, können demnach als bedeutsam für den Bearbeitungsprozess und infolgedessen die Aufgabenlösung angesehen werden.

#### Fragestellung

Verarbeitungsgeschwindigkeit wird als kognitive Basis von fluider Intelligenz verstanden (Schweizer & Koch, 2002) und sollte daher für kognitive Grundfähigkeit eine relevante Personenvariable bei der Aufgabenlösung darstellen. Eine höhere Verarbeitungsgeschwindigkeit sollte also in Aufgaben, die in erster Linie kognitive Grundfähigkeit erfordern, zu einem negativeren Zusammenhang zwischen Bearbeitungszeit und Erfolg führen (Hypothese 1). Kompetenzaufgaben sollten über kognitive Grundfähigkeit hinaus Ergebnisse aus Lerngelegenheiten abbilden (OECD, 2016). Deshalb sollte der Zusammenhang zwischen Bearbeitungszeit und Erfolg für Kompetenzaufgaben weniger von der Verarbeitungsgeschwindigkeit (Hypothese 2), als viel eher vom Bildungsniveau abhängen (Hypothese 3), da dieses entscheidend für die Lösung von Kompetenzaufgaben sein sollte.

#### Methode

Im Rahmen von PIAAC-L (Rammstedt, Martin, Zabal, Carstensen, & Schupp, 2017), einer längsschnittlichen Weiterverfolgung der PIAAC 2012 Stichprobe in Deutschland, bearbeiteten Personen 2015 PIAAC-Aufgaben zur Lesekompetenz und alltagsmathematischen Kompetenz. Ein Jahr später erhielten von diesen Personen N = 910 einen Zahlenreihentest (McArdle & Woodcock, 2009), der kognitive Grundfähigkeit abbilden soll, und einen Zahlen-Zeichen-Test (Schupp, Herrmann, Jaensch, & Lang, 2008), der Verarbeitungsgeschwindigkeit abbilden soll (vgl. McGrew, 2009). Mittels generalisierter linearer Mischmodelle (GLMMs; Wilson, De Boeck, & Carstensen, 2008) wird der Zusammenhang von Bearbeitungszeit und Erfolg untersucht, wobei der Zusammenhang von Bearbeitungszeit und Erfolg über Items und auch Personen variieren darf.

#### Ergebnisse

Wie erwartet lösten Personen mit einer höheren Verarbeitungsgeschwindigkeit Aufgaben, die kognitive Grundfähigkeit erfassen sollten, eher, wenn sie zur Bearbeitung weniger Zeit brauchten (Hypothese 1;  $\beta = -0.23$ ,  $p < .001$ ). Dieser Zusammenhang zeigte sich entgegen der Erwartung auch für Lesekompetenz ( $\beta = -0.09$ ,  $p = .007$ ). Wie erwartet (Hypothese 2) beeinflusste die Verarbeitungsgeschwindigkeit den Zusammenhang zwischen Bearbeitungszeit und Aufgabenerfolg nicht für alltagsmathematische Kompetenz ( $\beta = 0.02$ ,  $p = .517$ ). Dahingegen war der Zusammenhang wie erwartet in beiden Kompetenzaufgaben (Hypothese 3) im Vergleich zu einem niedrigen Bildungsniveau negativer für ein mittleres (alltagsmathematische Kompetenz:  $\beta = -0.27$ ,  $p = .011$ ; Lesekompetenz:  $\beta = -0.24$ ,  $p = .041$ ) und höheres Bildungsniveau (alltagsmathematische Kompetenz:  $\beta = -0.27$ ,  $p = .015$ ; Lesekompetenz:  $\beta = -0.39$ ,  $p = .001$ ). Für kognitive Grundfähigkeit zeigte sich für ein mittleres ( $\beta = 0.05$ ,  $p = .789$ ) oder höheres Bildungsniveau ( $\beta = -0.25$ ,  $p = .165$ ) dagegen kein negativerer Zusammenhang im Vergleich zu einem niedrigeren Bildungsniveau.

#### Fazit

Für kognitive Grundfähigkeit war wie erwartet die Verarbeitungsgeschwindigkeit und für die PIAAC Kompetenzen das Bildungsniveau eine bedeutsame Personenvariable, die den Zusammenhang zwischen Bearbeitungszeit und Erfolg beeinflusste. Der Aufgabenbearbeitungsprozess wird also bei Kompetenzaufgaben auch durch Bildungsfaktoren beeinflusst. Dies untermauert, dass die PIAAC Kompetenzaufgaben auch Ergebnisse individueller Lerngelegenheiten abbilden. Der entgegen der Erwartung zusätzliche negative Zusammenhang für hohe Verarbeitungsgeschwindigkeit bei Lesekompetenz kann durch eine hohe Anforderung an Verarbeitungsprozesse bei Leseaufgaben erklärt werden.

### **Äpfel und Birnen? Vergleich der Messung von Computer Literacy im NEPS und in ICILS 2013**

**Jan Marten Ihme, Martin Senkbeil**

IPN, Kiel; [ihme@ipn.uni-kiel.de](mailto:ihme@ipn.uni-kiel.de)

#### Theoretischer Hintergrund

Ein zentrales Ziel des Nationalen Bildungspanels (National Educational Panel Study; NEPS) ist, ausgewählte Kompetenzbereiche über die gesamte Lebensspanne zu erheben (Blossfeld, Schneider & Doll, 2009). Dabei wird zur gemeinsamen Interpretation die Anschlussfähigkeit an nationale und internationale Large Scale Assessments (LSA) angestrebt (Blossfeld, 2008). Konkret soll in diesem Beitrag der Zusammenhang zwischen der in NEPS erhobenen computerbezogenen Kompetenz (ICT Literacy) und dem Konstrukt Computer and Information Literacy (CIL) der International Computer and Information Literacy Study (ICILS 2013) untersucht werden.



Im NEPS wird ICT Literacy mit dem Test zur Erfassung technologischer und informationsbezogener Literacy (TILT; Senkbeil, Ihme & Wittwer, 2013) für die neunte Jahrgangsstufe erfasst. Es handelt sich dabei um einen Papier-Bleistift-Test mit 36 Multiple-Choice- und Multiple-True-False-Aufgaben. In ICILS 2013 wird CIL mit einem computerbasierten Test gemessen, der aus vier thematischen Testmodulen besteht, von denen jede Testperson zwei Module bearbeitet (Fraillon, Schulz & Ainley, 2013). Kennzeichnend für den CIL-Test ist die Nutzung dreier verschiedener Aufgabenformate, die sich hinsichtlich Interaktivität, Komplexität und Authentizität unterscheiden (Ihme, Senkbeil, Goldhammer & Gerick, 2017): nicht interaktive Aufgaben (z.B. MC-Aufgaben), kleine Simulationen und komplexe Autorenaufgaben.

Vor dem Hintergrund ähnlicher Konstruktbeschreibungen, aber offensichtlicher Unterschiede in den Testkonstruktionen wird der Frage nachgegangen, inwieweit die Testergebnisse von NEPS und ICILS 2013 gemeinsam interpretiert werden können. Hierfür sind nach Kolen und Brennan (2004) eine hinreichende Ähnlichkeit hinsichtlich der (1) Schlussfolgerungen, (2) der Merkmale und Umstände der Messung, (3) der Zielpopulationen sowie (4) der operationalisierten Konstrukte notwendig. Die Ähnlichkeit der Konstrukte kann weiterhin in (4.1) konzeptionelle, (4.2) dimensionale und (4.3) Skalenäquivalenz ausdifferenziert werden (van de Vijver, 1998).

#### Fragestellungen

1. Wird mit beiden Tests das gleiche Konstrukt gemessen oder unterschiedliche Konstrukte? Wird mit den nicht interaktiven Aufgaben aus dem ICILS-Test das gleiche Konstrukt gemessen wie mit dem NEPS-Test?
2. Inwieweit stellt CIL durch die Messung mit verschiedenen Aufgabentypen in ICILS 2013 ein mehrdimensionales Konstrukt dar? Inwieweit ergeben sich dadurch differentielle Zusammenhänge mit dem NEPS-Test?
3. Spielt der Modus der Messung (mit Papier und Bleistift oder computerbasiert) eine bedeutsame Rolle bei der Messung?

Zur Beantwortung der Forschungsfragen werden verschiedene Faktorenmodelle geprüft, in denen (1) die NEPS- und ICILS-2013-Aufgaben auf gleichen oder auf unterschiedlichen Faktoren laden, (2) die Aufgaben nach Typus drei Faktoren zugeordnet werden oder ein g-Faktor postuliert wird, (3) ein Methodenfaktor berücksichtigt wird oder nicht. Auf diese Weise ergeben sich  $2 \times 2 \times 2 = 8$  Modelle, die vergleichend evaluiert werden.

#### Methode

In ICILS 2013 wurden in Deutschland  $N = 2225$  Schülerinnen und Schüler der achten Klasse untersucht. Jede Testperson bearbeitete zwei der vier ICILS-Testmodule. Im Rahmen einer sich unmittelbar anschließenden nationalen Zusatzstudie wurde zusätzlich der TILT vorgegeben.

#### Ergebnisse

Der Vergleich der Faktorenmodelle zeigt, dass (1) Modelle, die zwei unterschiedliche gemessene Konstrukte in ICILS 2013 und NEPS annehmen, nicht bedeutsam besser abschneiden als die entsprechenden Modelle, die nur ein gemeinsames Konstrukt annehmen. Weiterhin (2) passen Modelle, die eine Faktorenstruktur nach Aufgabentypus annehmen, besser als g-Faktor-Modelle. (3) Die Einführung eines Methodenfaktors führt jeweils nur zu einer sehr geringfügigen Verbesserung der Modellpassung, so dass kaum von relevanten Testmoduseffekten auszugehen ist. Insgesamt ergibt sich zwischen den beiden Gesamttests eine Korrelation von  $r = .82$ , zwischen NEPS-Gesamttest und den nicht interaktiven ICILS-Aufgaben sogar von  $r = .89$ .

Aus diesen Ergebnissen kann gefolgert werden, dass die beiden Tests auf globaler Ebene unterschiedlich sind, mit den nicht interaktiven Aufgaben aus beiden Studien aber die gleiche Facette von CIL gemessen wird. In ICILS 2013 wird CIL umfassender gemessen, mit dem NEPS-Test dagegen nur ein Teilkonstrukt.

## Verbesserung interkultureller Vergleichbarkeit nicht-kognitiver Selbstberichte durch Ankervignetten

**Tamara Marksteiner<sup>1</sup>, Susanne Kuger<sup>2</sup>, Eckhard Klieme<sup>2</sup>**

<sup>1</sup>Universität Mannheim, Deutschland; <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland;  
[t.marksteiner@uni-mannheim.de](mailto:t.marksteiner@uni-mannheim.de)

Selbstberichte in internationalen Vergleichsstudien wie dem Programme for International Student Assessment (PISA) weisen häufig kulturspezifische Antworttendenzen auf. Diese Antworttendenzen erschweren den Vergleich von Forschungsergebnissen zwischen verschiedenen Kulturen oder Ländern. Nichtsdestotrotz ist die valide Interpretation psychologischer Daten durch Forscher von zentraler Wichtigkeit für die Forschung selbst, Praktiker sowie Politiker. Zur Abschwächung von kulturspezifischen Antworttendenzen wurden Ankervignetten erfolgreich eingesetzt (King & Wand, 2007; Kyllonen & Bertling, 2014). Die Resultate früherer Studien zeigen, dass Ankervignetten das sogenannte „attitude achievement paradox“ reduzieren können (e.g., Mickelson, 1990). Dieses Paradoxon beschreibt eine erwartete (positive) Korrelation zwischen, beispielsweise, der Einstellung zur Schule und der Leistung auf der individuellen (Schüler-) Ebene, sowie eine unerwartete (negative) Korrelation auf Länderebene (siehe Kyllonen & Bertling, 2014). Zur Testung, ob Ankervignetten tatsächlich kulturspezifische Antworttendenzen abschwächen, können Messinvarianzanalysen sowohl für verankerte als auch nicht-verankerte Selbstberichte durchgeführt werden. Die Wichtigkeit der Prüfung der Messinvarianz im interkulturellen Forschungskontext (Marsh, Hau, Artelt, Baumert, & Peschar, 2006) und die Abmilderung kulturspezifischer Antworttendenzen durch Ankervignetten (He, Buchholz & Klieme, 2017; King & Wand, 2007; auch Bolt, Lu, & Kim, 2014) wurde bereits mehrfach hervorgehoben. Dennoch befasste sich bislang erst eine Studie (He et al., 2017) mit der erhöhten Validität und verbesserten interkulturellen Vergleichbarkeit verankerter Selbstberichte verglichen mit nichtverankerten Selbstberichten mittels Messinvarianzanalysen und Strukturgleichungsmodellen. He und Kollegen verwendeten den PISA 2012 Datensatz, um Messinvarianz von lehrerbezogenen, nicht-kognitiven Faktoren Teacher Support und Classroom Management zu testen.

Da das „attitude achievement paradox“ jedoch vor allem im Zusammenhang mit schülerbezogenen Selbstberichten kritisch diskutiert wird und neben der Validität eines Indikators auch seine Messgenauigkeit sichergestellt sein sollte, ist es das Ziel der vorliegenden Studie, die Forschungsliteratur bezüglich der Möglichkeit der Verbesserung der Datenqualität mithilfe von Ankervignetten zu erweitern. Wir testeten, ob Ankervignetten die Reliabilität von Berichtsindikatoren erhöhen, indem verschiedene Reliabilitätskriterien angewandt wurden (z.B. Varianzverteilung auf den Ebenen, interne Konsistenz, Faktorstruktur und Messinvarianz). Verglichen mit der Studie von He et al. (2017) liegt das Hauptaugenmerk dieser Studie auf schülerbezogenen nicht-kognitiven Faktoren (z.B. Organisation und Leistungsmotivation). Es wird angenommen, dass schülerbezogene Faktoren prädiktiver für Schülerverhalten und -leistung sind als lehrerbezogene Faktoren. Die Analysen beschränken sich dabei nicht auf die Kompatibilität von verankerten und nicht-verankerten Faktoren zwischen verschiedenen Kulturen, sondern beziehen möglicherweise ursächliche Faktoren wie z. B. die Erhebungssprache mit ein. Dies ist wichtig, da

die Bedeutung der Konstrukte nicht nur zwischen Ländern oder Kulturen variieren könnte, sondern auch Sprachnuancen eine relevante Varianzquelle darstellen können (Grisay, de Jong, Gebhardt, Berezner, & Halleux-Monseur, 2007).

Die Daten des PISA 2015 Feldtests (PISA2015FT) beinhalten 15-jähriger Schüler aus 72 verschiedenen Ländern. Während des PISA2015FT wurden zwei verschiedene Sets von Ankervignetten (Organisation und Leistungsmotivation) verwendet, um die Selbstberichte der Schüler zu verankern. Die Vignetten zu Organisation wurden verwendet, um Selbstberichte über Prokrastination, Geschäftigkeit und Planung und Organisation zu verankern. Die Vignetten zu Leistungsmotivation wurden verwendet, um die Selbstberichte über Leistungsmotivation und akademische Selbstwirksamkeit zu verankern. Die Analysen basieren auf Daten von nSTU = 23.229 Schülerinnen und Schülern in nCNT = 57 Ländern.

Die Ergebnisse sind über verschiedene Skalen und Vignetten hinweg nicht eindeutig. Die durchgeführten exploratorischen und konfirmatorischen Faktorenanalysen deuten darauf hin, dass die Datenqualität mancher Konstrukte tatsächlich durch Verankerung verbessert werden kann. Es wurde jedoch bei anderen Konstrukten ein deutlich schwächerer Effekt der Ankervignetten beobachtet. Das Einbeziehen der Testsprache scheint die Messinvarianz deutlich zu erhöhen. Abschließend werden die Ergebnisse mit Blick auf die sich ergebenden Konsequenzen für die Forschungsmethodik diskutiert.

## **Der Einfluss von Stichproben auf Skaleneigenschaften**

**Christine Aichele, Johannes Hartig**

DIPF, Deutschland; [Christine.Aichele@dipf.de](mailto:Christine.Aichele@dipf.de)

Dieser Beitrag thematisiert die allgemeine Problematik, dass psychometrische Kennwerte von Wissens- oder Kompetenztests schlecht ausfallen können, wenn die untersuchte Stichprobe keine oder nur geringe Lerngelegenheiten für das gemessene Konstrukt gehabt hat.

Im Prozess der Testentwicklung dient die Pilotierungsphase zur Erprobung neu entwickelter Items, oftmals auch der Itemselektion. Dafür ist eine angemessene Stichprobe erforderlich. Nach Rost zielt „[d]as Prinzip der Passung von Personenstichprobe und Itemstichprobe [...] in erster Linie auf die Maximierung der Varianz der Antwortvariablen ab.“ (1996, S. 57). Items, bei denen alle Testteilnehmer dieselbe Antwort geben, also keine Varianz aufweisen, sind für den Rückschluss von Antwortverhalten auf das interessierende Merkmal wertlos. Ohne Lerngelegenheiten bestehen nur geringe systematische interindividuelle Unterschiede zwischen den Testteilnehmern, was zu geringen Trennschärfen und interner Konsistenz sowie hohen Itemschwierigkeiten führen kann. Das Phänomen wird an Teilstichproben und ausgewählten Skalen des Projekts Ko-NaMa illustriert.

Die Pilotierung im Ko-NaMa-Projekt diente unter anderem der Erprobung von Items zu BWL-Wissen ( $I = 129$ ) und der Kürzung dieser Skala für die Haupterhebung. Im Vortrag wird die Skalierung von zwei Teilstichproben vorgestellt, die sich anhand ihrer universitären Lerngelegenheiten, operationalisiert über die Anzahl der Fachsemester, hinsichtlich BWL-Themen unterscheiden. Stichprobe 1 ( $N_1 = 173$ ) umfasst wirtschaftswissenschaftliche Studierende aus den ersten beiden Fachsemestern, Stichprobe 2 Studierende aus den Fachsemestern 3 bis 6 ( $N_2 = 426$ ).

Die Modellschätzung mit dem Rasch-Modell für Stichprobe 1 weist eine niedrigere Reliabilität auf ( $EAP.PV_1 = .642$ ) als die Modellschätzung für Stichprobe 2 ( $EAP.PV_2 = .744$ ). Die Gesamtskala kann hinsichtlich der Korrelation von Itemantwort und Testwert (rp.b.WLE) optimiert werden. Dadurch lassen sich die Reliabilitäten beider Modellschätzungen verbessern auf  $EAP.PV_1 = .704$  beziehungsweise  $EAP.PV_2 = .757$ . Damit erreicht die Reliabilität der Modellschätzung in beiden Stichproben ein akzeptables Niveau. (Im Vortrag soll die Skalenoptimierung ergänzt um Itemfit-Maße vorgestellt werden.) Die Anzahl der dafür ausgeschlossenen Items unterscheidet sich jedoch. In Stichprobe 1 werden 59 Items aufgrund ihrer Kennwerte ausgeschlossen, in Stichprobe 2 sind es 20, darunter 13 übereinstimmende Items. In der Stichprobe mit weniger Lerngelegenheiten werden also nicht nur mehr Items aufgrund psychometrischer Kennwerte ausgeschlossen, sondern auch andere Items als in der Stichprobe mit mehr Lerngelegenheiten.

Fazit: Ziel der Pilotierung im Ko-NaMa-Projekt war die Erprobung neu entwickelter Items und Kürzung der Skala für die Haupterhebung. Nutzt man zunächst psychometrische Kennwerte als Anhaltspunkt für die Itemselektion der Skala zu BWL-Wissen, ergeben sich unterschiedliche finale Itemsets in den Stichproben in Abhängigkeit ihrer Lerngelegenheiten. Wird zur Testerprobung eine Stichprobe ohne vorige Lerngelegenheiten untersucht, besteht die Gefahr, Items zu verwerfen, die bei einer Messung von Lernergebnissen, also in einer Stichprobe mit vorherigen Lerngelegenheiten, zufriedenstellende Eigenschaften aufweisen würden. Zudem ist der Itempool in Stichprobe 1 nach der Selektion mithilfe psychometrischer Kriterien deutlich kleiner als in Stichprobe 2. Dadurch bleibt für die finale Itemauswahl nach theoretischen und inhaltlichen Überlegungen wenig Spielraum, was sich negativ auf die Repräsentativität der Iteminhalte auswirken kann.

Im Vortrag werden Möglichkeiten zum allgemeinen Umgang mit dieser Problematik bei der Messung von Lernzuwächsen diskutiert.

## D09

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Bernoullianum 32, H223

### Zum Einfluss der Lebenslage auf die Motivation des Weiterbildungspersonals

**Andreas Martin**

Deutsches Institut für Erwachsenenbildung, Deutschland; [martin@die-bonn.de](mailto:martin@die-bonn.de)

#### Theoretischer Hintergrund

Wie in anderen Bildungsbereichen bildet Professionalität auch in der Weiterbildung ein zentrales Thema. Die Weiterbildungsforschung sieht sich jedoch einer ganzen Reihe von spezifischen Gegebenheiten gegenüber. Ein besonders prägender, aber noch immer wenig diskutierter Unterschied zu den anderen Bildungsbereichen besteht dabei vor allem darin, dass Weiterbildung nicht nur im Kontext des Staates und im öffentlichen Auftrag, sondern zum weit überwiegenden Teil in betrieblichen oder marktförmigen Kontexten stattfindet. Anders als in den öffentlichen Bildungsbereichen dominieren zweckrationale Handlungsorientierungen. Die maßgeblichen Leistungserwartungen unterscheiden sich darum in weiten Teilen kaum von denen kommerzieller Dienstleistungen (Schrader 2011, S.122). Die Erfassung und Messung professioneller Kompetenzen dienen aus dieser Perspektive nicht nur der Professionalisierung des Tätigkeitsfeldes sondern vor allem auch der Bewertung des Arbeits- und Leistungsvermögens (Schrader & Goeze 2017 i.V).

Sowohl aus professionstheoretischer als auch aus arbeitssoziologischer Perspektive erweisen sich dabei die motivationalen Dimensionen professioneller Kompetenzen als zentral.

Aus professionstheoretischer Perspektive besteht eine Bedingung für die Qualität von Weiterbildung darin, dass Lehrende ihr fachliches, fachdidaktisches und pädagogisches Wissen tatsächlich in der Praxis anwenden und kontinuierlich weiterentwickeln. Dies ist jedoch wesentlich von motivationalen Orientierungen abhängig (Kunter et al. 2009, S. 157).

Aus arbeitssoziologischer Perspektive spielt die Motivation von Erwerbstätigen im Weiterbildungsbereich besonders hinsichtlich der Rentabilität interaktiver Dienstleistungen eine maßgebliche Rolle. Renditen sind nur dann realisierbar, wenn die Arbeitskraft über das für die eigene Reproduktion notwendige Maß hinaus verausgabt wird (Braverman 1980, S. 53). Dies ist im Dienstleistungsbereich jedoch problematisch. Zum einen liegt die Kontrolle über den Arbeitsprozess nicht wie im Bereich der einfachen Industriearbeit beim Management. Stattdessen muss sich der Arbeitgeber darauf verlassen, dass die Arbeitskräfte aus eigenem Antrieb Interaktionen erfolgreich gestalten und im Interesse der Organisation handeln (Falk 2006). Zum anderen sind die Ergebnisse von Dienstleistungsarbeit nur schwer objektivierbar (Hoffmann et al. 2012). Aus diesem Moral Hazard leitet die Ökonomie das „Shirking“-Problem ab (Shapiro und Stiglitz 1984).

Sowohl für die Produktivität als auch die Qualität der Dienstleistungsarbeit in der Weiterbildung sind die motivationalen Aspekte professioneller Kompetenzen entscheidend. Dabei wird in der Ökonomie und der Soziologie die Entlohnung als entscheidender Prädiktor für die Motivation angesehen (Becker & Stiegler 1974, Akerloff & Yellen 1990; Fehr et al. 1997). Auch in der Weiterbildungsforschung richtet sich die Argumentation zur Professionalisierung vor allem auf die Lohnhöhe (Dobischat et al 2010).

#### Fragestellung

Diese Argumentationen reduzieren die Perspektive jedoch auf die unmittelbare Wirkung des Lohnes auf die Motivation und übersehen dabei, dass der Wert des Lohnes wesentlich von dessen Bedeutung für die Lebenshaltung des jeweiligen Haushaltes abhängt. Je prekärer die finanzielle Situation eines Haushaltes, umso höher muss der Wert jedes zusätzlichen Euro angenommen werden. Bei gleichen Löhnen und sonst gleichen sozioökonomischen und demographischen Merkmalen sollte deswegen die Motivation umso höher sein, je geringer das Nettoäquivalenzeinkommen ist.

#### Daten, Methoden und Ergebnisse

Um dies zu überprüfen, wurde auf der Grundlage der erweiterten Zusatzbefragung des wb-personalmonitor (Autorengruppe wb-personalmonitor 2017) mit n=836 hauptberuflichen Erwerbstätigen der Einfluss von Nettoäquivalenzeinkommen auf die Verausgabungsbereitschaft des Weiterbildungspersonals regressiert. Die Nettoäquivalenzeinkommen wurden auf der Grundlage der Haushaltsstruktur, der Haushaltseinkommen und den regionalen Bedarfsgewichten berechnet. Die Verausgabungsbereitschaft ist eine der vier im wb-personalmonitor erhobenen AVEM-Dimensionen (ebd.) und konnte valide und reliabel operationalisiert werden. Um eine möglichst restriktive Schätzung zu gewährleisten, wurde neben den normalen OLS auch Dose-Response-Modelle mit General-Propensity-Score geschätzt. Im Propensity-Score wurden die Löhne, die Wochenarbeitszeiten, der Bildungsstand, die Stellung im Beruf, Alter, Geschlecht sowie der jeweilige Einrichtungstyp berücksichtigt.

Die Ergebnisse bestätigen, dass bei vergleichbaren Löhnen und sonst gleichen sozioökonomischen und demographischen Merkmalen die Nettoäquivalenzeinkommen die Verausgabungsbereitschaft wie erwartet beeinflusst.

### Überprüfung des GI/E-Modells für die Zusammenhänge zwischen empfundener Diskriminierung durch die Lehrkraft und Schülermotivation

**Friederike Helm<sup>1</sup>, A. Katrin Arens<sup>2</sup>, Jens Möller<sup>1</sup>**

<sup>1</sup>Christian-Albrechts-Universität zu Kiel, Deutschland; <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt, Deutschland; [helm@psychologie.uni-kiel.de](mailto:helm@psychologie.uni-kiel.de)

Das generalisierte Internal/External Frame of Reference Modell (GI/E-Modell; Möller, Müller-Kalthoff, Helm, Nagy & Marsh, 2016) baut auf dem I/E-Modell (Marsh, 1986) auf. Dieses beschreibt den pfadanalytischen Befund, dass verbale und mathematische Schülerfähigkeitsselbstkonzepte geringer korrelieren als die Fachleistungen und sich positive Effekte der Fachleistungen auf die korrespondierenden Selbstkonzepte sowie negative Effekte der Fachleistungen auf die nicht-korrespondierenden Selbstkonzepte zeigen (z.B. Möller, Pohlmann, Köller & Marsh, 2009). Dieses Zusammenhangsmuster wird mit dem Zusammenwirken externaler (sozialer) Vergleiche (Vergleiche eigener Leistungen in einer Domäne mit Leistungen von Mitschülern in dieser Domäne) und internaler (dimensionaler) Vergleiche (Vergleiche eigener Leistungen in einer Domäne mit eigenen Leistungen in einer anderen Domäne) begründet: Soziale Vergleiche führen zu positiven Effekten der Leistungen auf die korrespondierenden Selbstkonzepte, dimensionale Vergleiche begründen negative Effekte der Leistungen auf die nicht-korrespondierenden Selbstkonzepte.

Das GI/E-Modell erweitert das Verständnis dimensionaler Vergleiche, indem es sie als Vergleiche zwischen jeglichen Merkmalen von zwei Domänen definiert (Möller & Marsh, 2013; Möller et al., 2016), die sich auch auf andere Variablen außer

auf Selbsteinschätzungen auswirken können. In der vorliegenden Studie wurden die Annahmen des GI/E-Modells für die Zusammenhänge zwischen wahrgenommener Diskriminierung durch die Lehrkraft und Schülermotivation überprüft. Ergänzend wurde untersucht, ob die Zusammenhänge zwischen wahrgenommener Diskriminierung durch die Lehrkraft und dem Fachinteresse sowie der wahrgenommenen Wichtigkeit des Faches innerhalb und zwischen Domänen durch das Schüler selbstkonzept vermittelt werden.

Die analysierten Daten wurden in der Längsschnittstudie BiKS (Faust, 2013) erhoben, Daten von N=1685 Sechstklässlern wurden ausgewertet. Es wurden Strukturgleichungsmodelle spezifiziert, in denen die wahrgenommene Diskriminierung, die fachspezifischen Selbstkonzepte, das Interesse sowie die empfundene Wichtigkeit des Faches als latente Faktoren modelliert wurden. Selbstkonzept, Interesse und Wichtigkeit in Deutsch und Mathematik wurden in drei Modellen auf die wahrgenommene Diskriminierung in beiden Fächern regrediert, dabei wurde für die Leistungen in den beiden Fächern kontrolliert. Zur Überprüfung der Mediationsfragestellung wurden in zwei Modellen Interesse und Wichtigkeit auf die empfundene Diskriminierung in beiden Fächern zurückgeführt und zudem indirekte Effekte der Diskriminierung in beiden Fächern auf Interesse und Wichtigkeit über die Selbstkonzepte spezifiziert. Die hierarchische Struktur der Daten wurde durch Anwendung eines entsprechenden Schätzverfahrens berücksichtigt.

Die Ergebnisse zur ersten Fragestellung zeigen hypothesenkonform negative Effekte der wahrgenommenen Diskriminierung auf Selbstkonzept (Mathematik: -.26; Deutsch: -.19, beide  $p < .05$ ), Interesse (Mathematik: -.36; Deutsch: -.31, beide  $p < .05$ ) und Fachwichtigkeit (Mathematik: -.31; Deutsch: -.28, beide  $p < .05$ ) innerhalb der Fächer sowie positive Effekte zwischen den Fächern (Mathematik-Diskriminierung  $\rightarrow$  Deutsch-Selbstkonzept/Interesse/Wichtigkeit: .10,  $p < .05$ .11,  $p < .05$ .03, n.s.; Deutsch-Diskriminierung  $\rightarrow$  Mathematik-Selbstkonzept/Interesse/Wichtigkeit: .10/.15/.07, alle  $p < .05$ ). Die Annahmen des GI/E-Modells können damit für die Zusammenhänge zwischen wahrgenommener Diskriminierung und Selbstkonzept, Interesse sowie Wichtigkeit der Fächer Mathematik und Deutsch als bestätigt gelten.

Die Ergebnisse der Überprüfung der Mediationsfragestellung zeigen neben den negativen Effekten der Diskriminierung in beiden Fächern auf die entsprechenden Selbstkonzepte sowie den positiven Effekten auf die nicht entsprechenden Selbstkonzepte positive Effekte der Selbstkonzepte auf Interesse/Wichtigkeit im entsprechenden Fach (Mathematik: .66/.53; Deutsch: .44/.41, alle  $p < .05$ ). Zudem zeigen sich im Vergleich zum Modell ohne Mediator geminderte direkte Effekte der empfundenen Diskriminierung auf Interesse/Wichtigkeit innerhalb der Fächer (Mathematik: -.18/-.18; Deutsch: -.23/-.20, alle  $p < .05$ ) und zwischen den Fächern (Mathematik-Diskriminierung  $\rightarrow$  Deutsch-Interesse/Wichtigkeit: .10,  $p < .05$ -.01, n.s.; Deutsch-Diskriminierung  $\rightarrow$  Mathematik-Interesse/Wichtigkeit: .09/.03, beide  $p < .05$ ). Sämtliche indirekten Effekte fallen signifikant aus (Mathematik-Diskriminierung  $\rightarrow$  Mathematik-Selbstkonzept  $\rightarrow$  Mathematik-Interesse/Wichtigkeit: -.17/-.13; Deutsch-Diskriminierung  $\rightarrow$  Deutsch-Selbstkonzept  $\rightarrow$  Deutsch-Interesse/Wichtigkeit: -.08/-.08; Mathematik-Diskriminierung  $\rightarrow$  Deutsch-Selbstkonzept  $\rightarrow$  Deutsch-Interesse/Wichtigkeit: .04/.04; Deutsch-Diskriminierung  $\rightarrow$  Mathematik-Selbstkonzept  $\rightarrow$  Mathematik-Interesse/Wichtigkeit: .07/-.05; alle  $p < .05$ ). Damit zeigen sich partielle Mediationen der Effekte der empfundenen Diskriminierung auf Interesse und Wichtigkeit desselben Faches durch das Selbstkonzept sowie vollständige Mediationen der entsprechenden Effekte zwischen den Fächern.

## **Inwieweit verändern sich Überzeugungen zu Mehrsprachigkeit in Schule und Unterricht durch Lerngelegenheiten zum sprachsensiblen Unterrichten? Ein Vergleich von Lehramtsstudierenden und Lehrkräften**

**Nele Fischer, Timo Ehmke**

Leuphana Universität Lüneburg, Deutschland; [nele.fischer@leuphana.de](mailto:nele.fischer@leuphana.de)

Aktuell ist die Lehrerbildung gefordert, Studierende aller Fächer auf das sprachensible Unterrichten vorzubereiten. Hierzu gehört, angelehnt an das Modell professioneller Kompetenz von Lehrkräften (Baumert & Kunter 2006), nicht nur die Vermittlung von Wissen, um im Unterricht einen professionellen Umgang mit Mehrsprachigkeit zu ermöglichen, sondern auch die Reflexion bestehender Überzeugungen. Inwiefern diese stabil oder veränderlich sind, ist bisher nicht abschließend geklärt (Trautwein 2013). Hierbei ist zum einen interessant, ob Überzeugungen durch entsprechende Lerngelegenheiten beeinflusst werden und zum anderen, ob sich Überzeugungen durch fortschreitende Berufspraxis verändern. Im Forschungsdiskurs dominiert die Ansicht, dass Überzeugungen eher stabil und veränderungsresistent sind (Pajares 1992, Kane et al. 2002). Dennoch wird immer wieder betont, dass geeignete Lerngelegenheiten eine Veränderung initiieren können (u.a. Trautwein 2013, Hachfeld et al. 2012).

Inwieweit Überzeugungen zu Mehrsprachigkeit in Schule und Unterricht veränderbar sind, wurde anhand einer längsschnittlichen Studie mit N = 368 Teilnehmenden untersucht, von denen N = 333 Lehramtsstudierende und N = 35 bereits ausgebildete Lehrkräfte waren. Die Gruppe der Lehramtsstudierenden durchlief an drei verschiedenen Standorten unterschiedliche Lehrangebote zu sprachsensiblen Unterrichten. Zusätzlich wurde eine Kontrollgruppe von N = 34 Studierenden gebildet, die keinerlei Lehrangebot mit entsprechenden Inhalten besuchten. Die teilnehmenden Lehrkräfte nahmen ebenfalls an einer Fortbildung zum genannten Thema teil. Die Überzeugungen zu Mehrsprachigkeit in Schule und Unterricht wurden anhand eines empirisch erprobten Instruments erfasst. Dieses umfasst insgesamt 21 Items ( $\alpha_1 = .85$ ,  $\alpha_2 = .86$ ), bei denen die Zustimmung zu Aussagen auf einer vierstufigen Likert-Skala abgefragt wird. Zusätzlich wurden Personenmerkmale (u. a. Alter, Geschlecht, Unterrichtsfächer und Schulform) und Lerngelegenheiten zum sprachsensiblen Unterrichten erfasst.

Anhand der erhobenen Daten wurde folgenden Forschungsfragen nachgegangen:

1. Lassen sich Überzeugungen zu Mehrsprachigkeit in Schule und Unterricht mit geeigneten Lerngelegenheiten positiv beeinflussen?
2. Welche Zusammenhänge lassen sich zwischen den Überzeugungen und Personenmerkmalen sowie den wahrgenommenen Lerngelegenheiten feststellen?
3. Gibt es Unterschiede zwischen den Überzeugungen von Lehramtsstudierenden und Lehrkräften?

Die Studie konnte zeigen, dass Überzeugungen gegenüber Mehrsprachigkeit in Schule und Unterricht sowohl von Studierenden als auch bei aktiven Lehrkräften positiv beeinflussbar sind. Ein Vergleich mit einer Kontrollgruppe belegt, dass die Veränderung abhängig ist von Lerngelegenheiten, die Wissen zu sprachsensiblen Unterrichten vermitteln. In Hinblick auf bedeutsame Personenmerkmale von Lehramtsstudierenden konnte herausgearbeitet werden, dass vor allem die steigende Semesterzahl, das Alter und Geschlecht im Zusammenhang zu Überzeugungen hinsichtlich Mehrsprachigkeit in Schule und Unterricht stehen. Deren Veränderung über die Zeit stand in einem signifikanten Zusammenhang mit dem Unterrichtsfach Mathematik. Bei der Gruppe der bereits ausgebildeten Lehrkräfte stellte sich das Alter als besonders interessanter Faktor dar –

zu Anfang der Intervention stand es in einem negativen Zusammenhang mit den Überzeugungen, stand dann aber in einem positiven Zusammenhang mit den Veränderungen der Überzeugungen (je älter die Lehrkräfte, desto defizitorientierter die Überzeugungen zum ersten Messzeitpunkt, aber desto stärker die Veränderung zum zweiten Messzeitpunkt). Die Veränderung war gleichzeitig abhängig von den Unterrichtsjahren – je länger eine Lehrkraft im Schuldienst war, desto schwerer ließen sich ihre Überzeugungen verändern. Der Vergleich der Überzeugungen von Lehramtsstudierenden und Lehrkräften zeigte, dass die Lehrkräfte bereits leicht ressourcenorientierter starteten und auch im Laufe der Intervention eine stärkere Entwicklung zeigten als die Studierenden. Gleichzeitig sind die Studierenden mit mehr Vorwissen in die Intervention gestartet und die Veranstaltungen hatten allesamt einen Grundlagencharakter für das Thema des sprachsensiblen (Fach-)unterrichts.

Die vorliegende Studie zeigt, dass Überzeugungen nicht gänzlich veränderungsresistent sind. Welche Einflussfaktoren (Lerngelegenheiten, Personenmerkmale etc.) besonders ausschlaggebend bei der positiven Veränderung sind bzw. welches Wissen eingesetzt werden muss, um besonders stabile Überzeugungen anzufechten, müssen weitere Analysen und auch Studien zeigen. Die gefunden Ergebnisse verdeutlichen aber bereits, dass das Konstrukt der Überzeugungen hohes Potential für die Entwicklung und Weiterentwicklung von Lerngelegenheiten im Bereich des sprachsensiblen Unterrichts für Lehramtsstudierende und bereits ausgebildete Lehrkräfte bietet.

## D10

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus S107

### Zur Bedeutung individueller Schülerprofile für die Wahrnehmung von Feedback im Deutschunterricht der Oberstufe

**Katharina Dreiling, Ariane S. Willems**

Georg-August-Universität Göttingen, Deutschland; [awillem1@gwdg.de](mailto:awillem1@gwdg.de)

Das lern- und leistungsbezogene Feedback, das Schüler\*innen von Lehrkräften erhalten, ist ein bedeutsamer Einflussfaktor schulischen Lernens (Hattie & Timperley, 2007). Im Zentrum aktueller Feedbackforschung steht u.a. die Frage, wie Feedback gestaltet werden sollte, um den Lernprozess von Schüler\*innen zu begünstigen. Aufbauend auf theoretischen Modellen zur Qualität und Wirksamkeit von Feedback (Hattie & Timperley, 2007; Narciss, 2008) zeigen empirische Studien, dass vor allem ein elaboriertes Feedback, welches strategische Hinweise zur Optimierung des Lernprozesses enthält, im Vergleich zu einfachem Feedback, das lediglich Informationen über die Richtigkeit einer Aufgabenbewältigung enthält, positive Wirkungen auf Zielkriterien des Unterrichts hat (Rakoczy et al., 2008; Vollmeyer & Rheinberg, 2005). Allerdings werden die Effekte dabei durch unterschiedliche Merkmale moderiert (Harks et al., 2014; Kluger & DeNisi, 1996). Eine besondere Rolle hat hier die individuelle Wahrnehmung des Feedbacks durch die Schüler\*innen. Erste Befunde zeigen, dass diese positiv mit der Leistungs- und Interessenentwicklung von Schüler\*innen zusammenhängt (Rakoczy et al., 2013; Carvalho et al., 2014). Eine weitgehend offene Frage ist allerdings, inwiefern motivationale und kognitive individuelle Lernvoraussetzungen – wie z.B. Vorwissen, Selbstkonzept, leistungsbezogene Attribution, intrinsische Lernmotivation oder Fachinteresse – in Zusammenhang mit der individuellen Wahrnehmung von Feedback stehen. Dies ist aus methodischer Sicht nicht überraschend, da bisher kaum Studien existieren, in der die Wahrnehmung von Feedback im Unterrichtsgeschehen mittels eines standardisierten Instruments erfasst wird – insbesondere für das Fach Deutsch in der Oberstufe liegen noch keine Studien vor.

Der Beitrag stellt erste Befunde des Projektes FeeHe – Feedback im Kontext von Heterogenität vor. Ziel des längsschnittlich angelegten Projektes ist es, mittels eines neu entwickelten Fragebogens die Schülerwahrnehmung von Feedback zu erfassen und der Frage nachzugehen, wie individuelle Profile der Lernausgangslage von Schüler\*innen die Wahrnehmung und die Wirkung von Feedback beeinflussen.

In der FeeHe-Studie werden N=959 Schüler\*innen in N=46 Kursen des 11. und 12. Jahrgangs (Gymnasien/Gesamtschulen in Niedersachsen) zu Beginn und am Ende eines Schulhalbjahres zu ihrer Wahrnehmung der Qualität des Feedbacks im Deutschunterricht befragt. Darüber hinaus werden individuelle motivationale und kognitive Merkmale der Lernausgangslage der Schüler\*innen erhoben. Zur Erfassung des wahrgenommenen Feedbacks werden in Anlehnung an das theoretische Modell von Hattie & Timperley (2007) vier Subskalen eingesetzt (Dreiling et al., i.E.), die jeweils unterschiedliche Dimensionen eines lernwirksamen Feedbacks abbilden. Neben einer einfachen Komponente (evaluatives Feedback) werden drei Dimensionen des elaborierten Feedbacks unterschieden (prozessorientiert, selbstregulativ, diskursiv). Die Befunde latenter Faktoranalysen zeigen, dass die vier theoretisch postulierten Dimensionen auch empirisch eine zufriedenstellende Anpassung an die Daten aufweisen (CFI=.95; RMSEA=.05) und eine zufriedenstellende bis gute interne Konsistenz erreichen ( $.70 \leq \alpha \leq .86$ ). Zur Erfassung der individuellen motivationalen und kognitiven Lernausgangslagen der Schüler\*innen werden neben den Fachnoten als Indikator für das Vorwissen u.a. Skalen zum leistungsbezogenen Selbstkonzept, zur leistungsbezogenen Attribution, zur intrinsischen Lernmotivation sowie zum Fachinteresse eingesetzt. Die einzelnen Messmodelle weisen sehr gute Kennwerte auf (CFI $\geq$ .95) und die internen Konsistenzen der Skalen sind gut bis sehr gut ( $.75 \leq \alpha \leq .95$ ). Basierend auf den Items zur Beschreibung der motivationalen und kognitiven Lernausgangslage der Schüler\*innen werden mittels Multilevel-Latent-Profilen Analysen zunächst individuelle Schülerprofile bestimmt. Im Anschluss werden varianzanalytisch Unterschiede zwischen den Profilen für die vier Feedbackdimensionen untersucht.

Die Ergebnisse der Profilanalysen weisen auf vier unterschiedliche Schülerprofile hin. Das erste Schülerprofil ist durch eine überdurchschnittlich hohe Ausprägung sowohl auf motivationalen als auch auf kognitiven Merkmalen gekennzeichnet. Demgegenüber lässt sich eine Gruppe identifizieren, die durch eine niedrige motivationale und kognitive Lernausgangslage gekennzeichnet ist. Zwei weitere Profile sind Mischtypen, mit entweder einer eher hohen motivationalen oder einer eher hohen kognitiven Lernausgangslage. Erste Analysen zeigen zudem, dass die Wahrnehmung des Feedbacks durch die Schüler\*innen systematisch zwischen den Schülerprofilen variiert. Die vorliegenden Befunde legen daher den Schluss nahe, dass die Wahrnehmung unterrichtsbezogener Prozesse – insbesondere des Feedbacks – substanzial durch unterschiedliche Konfigurationen in der Lernausgangslage von Schüler\*innen beeinflusst wird.

### Wie rezipieren Lehrkräfte (quantitative) Rückmeldungen ihrer Unterrichtsqualität?

**Samuel Merk**

Universität Tübingen, Deutschland; [samuel.merk@uni-tuebingen.de](mailto:samuel.merk@uni-tuebingen.de)

Tagtäglich werden Daten generiert, welche Unterrichtsqualität beschreiben: Etwa beurteilen Fremdevaluator\*innen beobachtete Unterrichtssequenzen, Forscher\*innen melden Kovariaten von Studien zurück, Schüler\*innen geben ihren Lehrer\*innen Feedback in Form von Fragebögen.

Den so gewonnenen Daten wird in prominenten Modellen der Unterrichts- und Schulentwicklung großes Potential zugeschrieben (Visscher & Coe, 2003; Helmke, 2005). So wird in diesen Modellen meist postuliert, dass eine erfolgreiche Auseinandersetzung mit dem Datenmaterial zu einer vertieften Reflektion des eigenen unterrichtlichen Handelns und damit potentiell zu einer Steigerung der Unterrichtsqualität führt.

Eine notwendige Bedingung dafür, stellt jedoch die korrekte und zielgerichtete Rezeption der vorhandenen Rückmeldedaten dar, wobei es diesbezüglich als fraglich gilt, inwiefern Lehrerinnen und Lehrer über eine dafür ausreichende Datenkompetenz verfügen (Schneewind, 2007), welche Aufbereitung der Daten verständnisfördernd ist und welche didaktischen Begleitmaßnahmen eine vertiefte Rezeption fördern.

In der vorliegenden Studie wurden Rückmeldungen (SEEQ, Marsh 1982; N = 10.227) über eine neu entwickelte interaktive Webapplikation distribuiert, in der sowohl offene Antworten, als auch sog. full stacked bar graphs der Einzelitemantworten und Darstellungen der Skalenwerte des SEEQ angefordert werden konnten.

Die Darstellung der Dimensionsscores wurden experimentell variiert. Während die Hälfte der Lehrkräfte frei entscheiden konnte, welche Dimensionen (Lernerfolg, Enthusiasmus, Beziehung, Gruppeninteraktion), welche statistischen Größen (einzelne Skalenwerte, arithmetische Mittel, Quartile, Konfidenzintervalle) und welche Skalierungen (rohe Skalierung des

SEEQ, grand mean Standardisierung, ipsative Standardisierung) graphisch dargestellt wurden, konnte die andere Hälfte zwischen Zielvorgaben (z.B. Z1: "Ich will mich mit dem Gesamtergebnis vergleichen") wählen, die dann in entsprechenden Graphiken resultierten. Dabei wurde die Informationsgleichheit beider Gruppen gewährleistet.

Von allen Lehrerinnen und Lehrern wurden Logdaten ihres Rezeptionsverhaltens auf der Webapplikation aufgezeichnet, zudem wurden Sie im Anschluss an die Rezeption zu wahrgenommener Informativität und Interpretationssicherheit der einzelnen Elemente der Webapplikation befragt.

Die Ergebnisse weisen zunächst darauf hin, dass Lehrerinnen und Lehrer die Darstellung von offenen Antworten, Einzelantworten und Skalenwerten nicht differentiell bzgl. deren Informativität und der eigenen Interpretationssicherheit beurteilen.

Weiterhin ergaben sich keine signifikanten Unterschiede zwischen der freien und zielgeleiteten Version der Darstellung der Skalenwerte bzgl. Informativität und Interpretationssicherheit, jedoch hochsignifikante und in ihrer Größe bedeutsame Unterschiede zwischen einzelnen statistischen Merkmalen und Skalierungen (z.B. zu Ungunsten der Konfidenzintervalle)

Der Vortrag schildert das Material und die Analysen im Detail und diskutiert mögliche Implikationen für die Rückmeldepraxis.

## **Nutzung von Schüler-Lehrer-Feedback. Eine Lehrkräfte-Typologie.**

**Sebastian Röhl, Wolfram Rollett**

Pädagogische Hochschule Freiburg, Deutschland; [sebastian.roehl@ph-freiburg.de](mailto:sebastian.roehl@ph-freiburg.de)

Unterrichtsfeedback von Schülerinnen und Schülern wird als ein wichtiges Instrument zur Unterrichtsentwicklung angesehen (Bastian, Combe & Langer, 2007; Berger, Granzer, Looss & Waack, 2013; Scheunpflug, 2014). Allerdings bestehen gegenüber dem Einsatz von Schülerfeedback durchaus Vorbehalte von Lehrkräften (Döring, 2014). Für die Nutzung durch Lehrkräfte in der Praxis konnten u.a. das Erleben von positiver Bestätigung des eigenen pädagogischen Handelns sowie eine den Schülerrückmeldungen gegenüber offene Haltung nachgewiesen werden (Gaertner, 2014). Dagegen verringert eine als hoch empfundene Arbeitsbelastung die Nutzung (Ditton & Arnold, 2004). Darüber hinaus liegen Befunde klassifizierender Analysen für die Nutzung von Schülerfeedback vor, in denen ein offen-eingestelltes Muster mit Nutzung der Rückmeldungen sowie ein ablehnendes Muster ohne Nutzung identifiziert werden konnten (Ditton & Arnold, 2004).

In der hier vorgestellten Studie wird untersucht, ob sich diese Nutzungstypen in einer weiteren Stichprobe identifizieren lassen und ob bzw. inwieweit diese ggfs. in einer Beziehung zu Merkmalen der arbeitsbezogenen Erlebens- und Verhaltensmuster (AVEM, Schaarschmidt & Fischer, 2008) stehen, wie dies der Befund von Ditton und Arnold (2004) andeutet.

Datengrundlage bilden zwei Mitarbeiterbefragungen (n=98, Rücklaufquote >90%) an mehreren weiterführenden Schulen eines Schulverbands, in dem – im Gegensatz zu den oben genannten Studien – das Einholen von Schülerfeedback für Lehrkräfte verpflichtend war. Für die Erfassung der Nutzung und des Erlebens des Schülerfeedbacks wurden drei Kurzskalen verwendet: Nutzung zur Unterrichtsentwicklung (3 Items, McDonalds Omega: .637), Zugeschriebene Relevanz und Aussagekraft (2 Items, Omega: .821) und Positiv-bestätigendes Erleben der Rückmeldung (2 Items, Omega: .830). Darüber hinaus wurde der AVEM-44 (Schaarschmidt & Fischer, 2008) eingesetzt.

Zur Bestimmung von Nutzungstypen wurde eine hierarchische Clusteranalyse mit anschließender Clusterzentrenanalyse der drei Kurzskalen zum Schülerfeedback vorgenommen (vgl. Bortz & Schuster, 2010, 462ff.). Der Verlauf des Heterogenitätskoeffizienten legte eine Fünf-Clusterlösung nahe. Wie zu erwarten unterscheiden sich die Cluster über alle drei Skalen hinweg signifikant ( $.60 < \eta^2 < .78$ ). Reklassifizierende und kreuzvalidierende Diskriminanzanalysen bestätigten die Güte und Stabilität der Klassifizierung (Anteil richtiger Zuordnungen: 98% bzw. 95%, vgl. Bacher, Pöge & Wenzig, 2010).

Es ergaben sich folgende Typen:

1. Kritische Nutzer (18,4%): Geringe zugeschriebene Aussagekraft bei mittlerer Nutzung.
2. Positiv eingestellte Nichtnutzer (30,6%): Hohe zugeschriebene Aussagekraft bei einer geringen Nutzung.
3. Negativ resümierende Nutzer (17,3%): Niedrigste Ausprägung beim positiv-bestätigenden Erleben sowie mittlerer Relevanzzuschreibung und Nutzung.
4. Positiv eingestellte Nutzer (21,4%): Höchste Ausprägungen bei allen drei Skalen.
5. Negativ eingestellte Nichtnutzer (12,2%): Geringste Werte bei der Nutzung sowie der Relevanz. Die Rückmeldungen werden etwas negativer erlebt.

Weiterhin wurden einfaktorielle Varianzanalysen der weiteren Merkmale mit anschließenden PostHoc-Tests bezüglich der Typenzugehörigkeit durchgeführt. Bezüglich der Skalen des AVEM zeigten diese Analysen signifikante Unterschiede bei drei von vier Skalen aus dem Bereich Widerstandsfähigkeit gegenüber Belastungen (Distanzierungsfähigkeit, Resignationstendenz, Innere Ruhe), wobei die größten Kontraste jeweils zwischen Typ 4 mit günstigeren Ausprägungen in den AVEM Skalen und Typ 5 auftraten. Ein Vergleich der hier ermittelten Typologie mit der AVEM-Typologie ergab eine deutliche Überschneidung zwischen den positiv eingestellten Nutzern und dem Schonungsmuster.

Zwei der fünf identifizierten Typen erinnern an die von Ditton und Arnold (2004) berichteten Muster. Mit den übrigen drei Typen differenziert sich das Bild. Nutzung verbindet sich mit geringer Zuschreibung von Aussagekraft bzw. mit relativ negativem Erleben, aber auch geringe Nutzung mit einem relativ positiveren Erleben. Der Vergleich über die AVEM-Skalen deutet an, dass der inneren Widerstandsfähigkeit der Lehrpersonen eine Bedeutung für den Nutzen bzw. die Bewertung von Schülerfeedback zukommt.

Die vorgestellten Befunde illustrieren, wie unterschiedlich Lehrkräfte Schülerfeedback nutzen bzw. bewerten und liefern Hinweise für die Einführung und Begleitung von Schüler-Lehrer-Feedbackverfahren an Schulen. So könnte an die sich hier abzeichnenden differenten Haltungen argumentativ angeknüpft werden, indem beispielsweise sinnvolle Nutzungsmöglichkeiten sowie die Aussagekraft der Schülerrückmeldungen mit den Lehrerkollegien bearbeitet werden. Zudem deutet sich an, dass auch Defizite im Bereich der beruflichen Resilienz lohnende Anknüpfungspunkte sein könnten.

## **Erhebung und Analyse von Scaffolding in unterrichtlichen Interaktionen**

**Hanna Mach, Hermkes Rico, Minnameier Gerhard**

Goethe-Universität, Deutschland; [Mach@em.uni-frankfurt.de](mailto:Mach@em.uni-frankfurt.de)

Theoretischer Hintergrund

Scaffolding stellt eine Form der maßgeschneiderten Unterstützung (tailored support) dar (Greenfield, 1984; van de Pol & Elbers, 2013, Hermkes et al., 2017). Maßgeschneidertheit bezieht sich dabei auf die situative Passung zwischen der Unterstützung der Lehrperson und dem aktuell vorliegenden Verständnis der SuS. Black & William (2009) sprechen von "moments of contingency" (S. 10). Der Einfluss von Scaffoldingqualität auf Lernoutcomes konnte bislang nur im Rahmen von elterlich-tutorieller Unterstützung (Wood, 1980; Pino-Pasternak & Whitebread, 2010) oder in experimentellen Settings gezeigt werden (Murphy & Messer, 2000). Entsprechende Befunde im Unterrichtsgeschehen stehen noch aus. Dies ist im Wesentlichen auf methodische Defizite bei der Erhebung von Scaffolding zurückzuführen. Gemäß Hardy (2012) sollten interaktionsspezifische Analysen im Unterricht auf der Basis von „reasoning units“ erfolgen, was der o.g. Definition von Scaffolding entspricht. Dies setzt wiederum eine Rekonstruktion der schülerseitig stattfindenden Wissenskonstruktionsprozesse voraus (vgl. Minnameier et al., 2015).

#### Fragestellungen

Im Zentrum der durchgeführten Studie stehen drei Fragestellungen:

(1) Kann Scaffolding in Lehrer-Schüler-Interaktionen auf der Basis von inhaltlich-kognitiven Analyse-Einheiten als Prozessgröße valide und reliabel erfasst werden?

(2) Unterscheiden sich Scaffolding-Charakteristika (Dauer, Passung) bei verschiedenen Lehrpersonen, Schülergruppen, Aufgabentypen und Unterrichtsphasen?

Entsprechend der zwei Fragestellungen verfolgt der Beitrag zwei Ziele. Zum einen soll eine kürzlich entwickelte Methodik zur Erhebung von interaktionsbasiertem Scaffolding vorgestellt werden. Zum anderen werden aktuelle Ergebnisse der empirischen Videostudie zur Scaffoldingqualität im Unterricht präsentiert.

#### Methode

Insgesamt nahmen 63 SuS der 8./9. Klassenstufe (Fach „Politik und Wirtschaft“) aus 3 Gymnasien an der Studie teil. Der Unterricht umfasste 8 Schulstunden und fand in Kleingruppen à drei SuS statt. Jede Kleingruppe wurde durch eine Lehrkraft unterrichtet. Lehrkräfte waren Masterstudierende der Wirtschaftspädagogik. Es fand ein problemorientierter Unterricht statt, in dem die SuS ein Modellunternehmen führen und dabei Grundlagen des Rechnungswesens lernen sollten. Zur Erhebung der Schülervariable "Level of Attainment" wurden neben Videoaufnahmen des Unterrichts zusätzlich die Tabletaufzeichnungen der SuS herangezogen. Die Variable „Strength of Intervention“ bildet die Lehrerunterstützung ab. Aus den beiden Variablen wurden Interaktionsmuster gebildet, die mit Hilfe des von Wood et al. (1976) eingeführten Contingent-Shift-Prinzips hinsichtlich ihrer Scaffoldingqualität (Kontingenz) bewertet wurden.

#### Ergebnisse

##### Fragestellung 1

Ein Interaktionsmuster beginnt mit Auftreten einer Lernschwierigkeit und endet dem Erreichen eines höheren Level of Attainment. Ein Scaffold enthält insofern die Unterstützung, die die Lehrkraft bezogen auf eine Lernschwierigkeit bis zu deren Überwindung gibt. Auf diese Weise bildet ein Muster ein vollständiges Scaffold valide ab. Anhand der generierten Interaktionsmuster lassen sich Fehler im Lehrerverhalten präzise lokalisieren.

Die Interrater-Reliabilitäten für die Kodierung der Schüler- und Lehrer-Turns liegen zwischen  $\kappa = .771$  (N=693) und  $.897$  (N=821). Für die Schülervariable liegt die Übereinstimmung zwischen 731 und 1.00 (Spearman's Rho). Auch für die Lehrervariable konnten zufriedenstellende bis sehr gute Übereinstimmungen erzielt werden (zwischen  $\rho = .611$  und  $.849$ ). Lediglich zwischen zwei Ratern lag die Übereinstimmung bei  $\rho = .384$ .

##### Fragestellung 2

Von den bisher 281 analysierten Scaffolds waren 232 (82%) kontingent und 49 (18%) nicht-kontingent.

Der Anteil kontingenten Scaffoldings über dem gesamten Unterrichtszeitraum variiert zwischen den einzelnen Schülergruppen zwischen 66.7% und 93.1%. Für die jeweiligen Aufgaben betrug sie zwischen 60 und 100%. Der Anteil von nicht-kontingenten Scaffolding nimmt im Unterrichtsverlauf zu.

Es zeigt sich, dass Scaffolds sehr stark in ihrem Umfang variieren. Die durchschnittliche Anzahl an Lehrer- und Schüler-Turns, die ein Scaffold umfasst, liegt bei  $M=17,2$ , ( $SD=22,9$ ;  $Md=10$ ). Ein Faktor, der die Dauer eines Scaffolds beeinflusst, scheint im Aufgabentyp zu liegen. So weisen Scaffolds in Aufgaben zur kreativen Lösungsfindung durchschnittlich eine größere Turnanzahl auf ( $M=44,2$  Turns,  $SD=45,4$ ) als bei Probleminduktionsaufgaben, die der kognitiven Aktivierung der SuS dienen ( $M=18,2$  Turns,  $SD=22,7$ ). Auch gelingt es den Lehrkräften bei ersteren besser, kontingent zu unterstützen (durchschnittlicher Anteil kontingenten Scaffoldings von 91,9%) als bei Probleminduktionsaufgaben (74,3%).



## D11

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus, Aula 33

### **Kita-Förderung, Kita-Qualität und Sprachkompetenzen von Kindern – erste Ergebnisse der SEIKA-NRW-Studie**

Chair(s): **Kristine Blatter** (Deutsches Jugendinstitut e. V.), **Alexandru Agache** (Ruhr-Universität Bochum), **Katharina Kohl** (Ruhr-Universität Bochum, Deutschland)

DiskutantIn(nen): **Marcus Hasselhorn** (DIPF Frankfurt)

In den vergangenen Jahren ist die Feststellung und Förderung kindlicher Sprachkompetenzen im Elementarbereich in den Fokus von Forschung, Praxis und Politik gerückt. Ein wesentliches Ziel neuinitiiertener Fördermaßnahmen auf Bundes- und Landesebenen ist es, rechtzeitig vor Schuleintritt bessere Ausgangsbedingungen für Kinder mit Sprachförderbedarf zu schaffen. Zahlreiche Programme wurden auf den Weg gebracht, doch nur wenige sind auf ihre Effektivität hin überprüft. In Nordrhein-Westfalen erfolgte im Jahr 2014 eine Revision des Kinderbildungsgesetzes. Seitdem steht die alltagsintegrierte Sprachbildung in Kindertageseinrichtungen im Vordergrund, d.h. die Sprachförderung ist in den pädagogischen Alltag eingebunden und soll nicht mehr in zeitlich begrenzten, isolierten Einheiten stattfinden. Zur Unterstützung dieser alltagsintegrierten Sprachbildung stellt das Land Nordrhein-Westfalen zusätzliche finanzielle Fördermittel für Kitas mit einem hohen Anteil an Kindern mit besonderem Unterstützungsbedarf und/oder zusätzlichem Sprachförderbedarf zur Verfügung, welche hauptsächlich für pädagogisches Fachpersonal einzusetzen sind. Insgesamt wurden rund 100 Millionen € pro Kindergartenjahr für die Förderung bereitgestellt. Diese Fördermaßnahme wird im Projekt Sprachbildung und -entwicklung im Kita-Alltag (SEIKA-NRW) systematisch evaluiert. Die Studie untersucht, wie Ressourcen verteilt werden und welche Ressourcen in welchem Maße dazu beitragen, dass sich die sprachlichen Kompetenzen der Kinder bis zum Übergang in die Schule ausreichend verbessern. Untersucht wird dies auf unterschiedlichen Ebenen, dabei werden in mehrfacher Hinsicht methodisch neue Wege gegangen. Das Design umfasst sowohl qualitative als auch quantitative Methoden. Neben der Auswertung von vorhandenen Datensätzen und Dokumentenanalysen werden umfassende Interviews mit Jugendämtern und Kita-Leitungen durchgeführt. Außerdem werden in einem Längsschnitt rund 1000 zufällig ausgewählte Kinder untersucht. Diese stammen aus 95 Kitas, welche anhand eines aufwändigen Samplings von statistischen Drillings mittels Propensity Score Matching ausgewählt wurden.

Im Symposium werden vier Teilergebnisse der ersten Phase der aktuell noch laufenden Studie vorgestellt, die die Bedingungen der oben geschilderten finanziellen Fördermaßnahme auf unterschiedlichen Ebenen (Jugendamts-, Kita- und Kind-Ebene) betrachten.

Im ersten Beitrag werden die Kriterien der für die Verteilung zuständigen 186 Jugendämter anhand einer Dokumentenanalyse betrachtet. Hier werden typische Kriterien (z.B. ökonomische Situation der Familien, Sprachbedarf der Kinder) identifiziert, nach denen die geförderten Einrichtungen ausgewählt wurden. Zudem deckt die Studie Defizite in der Verfügbarkeit der für die Jugendämter notwendigen Indikatoren auf.

Der zweite Beitrag beschäftigt sich auf Kita-Ebene mit dem Einsatz der Fördermittel in der Praxis. Erste Ergebnisse zeigen strukturelle Unterschiede zwischen geförderten und nicht geförderten Einrichtungen (z.B. höherer Anteil von Kindern mit nicht-deutscher Familiensprache in geförderten Einrichtungen). Befunde zur Umsetzung alltagsintegrierter sprachlicher Bildung und zu Anforderungen an die Qualifikationen des zusätzlichen Personals fallen heterogen aus.

Im dritten Beitrag werden reziproke Einflüsse von unterschiedlichen Dimensionen der Kita-Qualität mit netzwerkanalytischen Verfahren untersucht. Die Ergebnisse zeigen Defizite im Bereich der sprachlich-kognitiv anregenden Interaktionen und identifizieren als wichtige Prädiktoren dafür das Involvement der Fachkräfte, die Strukturierung lernfördernder Aktivitäten und die emotionale Sensitivität. Dabei zeigt sich auch, dass der Personal-Kind-Schlüssel eine untergeordnete Rolle spielt.

Der vierte Beitrag untersucht in Mehrebenenmodellen die Zusammenhänge zwischen Quantität und Qualität der Kita-Betreuung und Sprachkompetenzen von Kindern mit deutscher und nicht-deutscher Familiensprache. Hier zeigt sich, dass eine hohe Kita-Qualität im Bereich der Lernunterstützung mit einer Steigerung der morphologischen Kompetenzen und des passiven Wortschatzes von Kindern mit nicht-deutscher Familiensprache einhergeht.

*Beiträge des Symposiums*

### **Evaluation zur Verteilung der Landeszuschüsse für plusKITA-Einrichtungen und Kindertageseinrichtungen mit zusätzlichem Sprachförderbedarf durch die Jugendämter in Nordrhein-Westfalen**

**Tijen Atkaya<sup>1</sup>, Nora Jehles<sup>2</sup>**

<sup>1</sup>Forschungsverbund DJI/ Technische Universität Dortmund, <sup>2</sup>Familiengerechte Kommune e.V.

Hintergrund

Mit der am 01. August 2014 in Kraft getretenen Revision des Kinderbildungsgesetzes (KiBiz) wurde die Stärkung der Sprachbildung und die Verbesserung der Bildungsgerechtigkeit und der Bildungschancen von Kindern als gesetzliche Regelaufgabe festgelegt. Dieser Revision liegt die Erkenntnis zugrunde, dass der Bildungserfolg maßgeblich vom sozialen Hintergrund der Familien geprägt wird (Ehmke & Jude, 2010) und dabei der frühkindlichen Bildung eine zentrale Funktion bei der Kompensation sozialer Ungleichheit zukommt (Flöter, Egert, Lee, & Tietze, 2013; Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). Für die Umsetzung der Gesetzesgrundlage stellt das Land Nordrhein-Westfalen finanzielle Fördermittel für Kindertageseinrichtungen zur Verfügung. Zu diesem Zweck wurden zusätzliche Landeszuschüsse für Kindertageseinrichtungen („plusKITA“) mit einem hohen Anteil von Kindern mit besonderem Unterstützungsbedarf im Bildungsprozess und Landeszuschüsse für zusätzlichen Sprachförderbedarf gesetzlich festgelegt. In einem ersten Schritt wurden die zusätzlichen Finanzmittel vom Landesministerium an die Jugendämter in Nordrhein-Westfalen verteilt, die diese in einem zweiten Schritt an die zu fördernden Kindertageseinrichtungen verteilten.

Fragestellung

Welche Kriterien wurden durch die Jugendämter zur Verteilung der beiden Landeszuschüsse an die Kindertageseinrichtungen ausgewählt?

Methoden/Daten

Um die Frage nach den verwendeten Kriterien zur Verteilung der Zuschüsse zu beantworten, wurden Beschlussdokumente von 186 Jugendämtern in Nordrhein-Westfalen in den Ratsinformationssystemen recherchiert und mittels Qualitativer Inhaltsanalyse (Mayring, 2015) ausgewertet.

#### Ergebnisse

Für die Verteilung der zusätzlichen Landeszuschüsse werden verschiedene Indikatoren verwendet, die sich den Themenfeldern der ökonomischen Situation, Migration, dem Sprachförderbedarf sowie der familiären Situation zuordnen lassen. Je nach Landeszuschuss bilden andere Indikatoren den Schwerpunkt: Die Landeszuschüsse für plusKITA-Einrichtungen werden von den meisten Jugendämtern auf der Basis von Kriterien verteilt, die die ökonomische Situation der Familien beschreiben. Als Kriterium werden am häufigsten die Elternbeiträge genutzt. Bei der Verteilung der zusätzlichen Sprachfördermittel zeigt sich das Spannungsfeld, in dem sich die örtliche Jugendhilfeplanung befindet: Mangels alternativer Daten wurde am häufigsten auf die Ergebnisse der stark kritisierten (Neugebauer & Becker-Mrotzek, 2013) und letztlich abgeschafften Sprachstandserhebung Delfin 4 (MGFFI, 2007) zurückgegriffen. Im Hinblick auf die Verfügbarkeit von Daten zeigen die Beschlussdokumente, dass ein möglichst kleinräumiges Monitoringsystem die Jugendämter in die Lage versetzt, Ressourcen, wie die zusätzlichen Landesmittel, bedarfsgerecht zu steuern und Fehlallokationen zu vermeiden. Für wissenschaftliche Forschung stellen die kommunalen Ratsinformationssysteme eine Quelle für qualitatives Datenmaterial dar, die bislang kaum Beachtung findet.

#### Schlussfolgerung

Die Analyse zeigt Diskussionsbedarf über das Spannungsfeld, in dem sich die Jugendämter befinden: Einerseits sollen die Kriterien vor dem Hintergrund des Forschungsstandes geeignet sein, um Einrichtungen mit einem hohen Anteil von Kindern mit besonderem Unterstützungsbedarf im Bildungsprozess und zusätzlichem Sprachförderbedarf auszuwählen. Andererseits müssen dafür die entsprechenden Daten in den bzw. für die Jugendämter verfügbar sein.

### **Sprachbildung im Elementarbereich in NRW: Ein Vergleich von Kitas mit unterschiedlicher finanzieller Förderung**

**Kristine Blatter, Katarina Groth, Veronika Eichmann, Margarita Stolarova**

Deutsches Jugendinstitut e. V.

Untersuchungen zeigen, dass ein relativ großer Anteil an Kindern zu Schulbeginn nicht über entsprechende Sprachleistungen im Deutschen verfügt (Autorengruppe Bildungsberichterstattung, 2016) und somit von Beginn der eigenen Bildungskarriere an benachteiligt ist (Artelt et al., 2001; Baumert & Schümer, 2002; Gogolin & Krüger-Potratz, 2006). Der Förderung kindlicher Sprachkompetenzen im Elementarbereich wird daher eine zunehmend wichtige Rolle zugeschrieben (Schneider et. al, 2012). Dabei erfolgt die Umsetzung sprachlicher Bildung in Kitas bundesweit sehr heterogen (Lisker, 2011, 2013). In Nordrhein-Westfalen soll seit August 2014 die Sprachförderung in Kitas nicht mehr in zeitlich begrenzten und isolierten Einheiten stattfinden, sondern als Sprachbildung den gesamten pädagogischen Alltag durchdringen (MFKJKS, 2014). Damit Kitas mit einem hohen Anteil an Kindern mit besonderem Unterstützungsbedarf und/oder Sprachförderbedarf diese alltagsintegrierte Sprachbildung besser umsetzen können, stellt das Land Nordrhein-Westfalen zusätzliche finanzielle Fördermittel bereit. Diese sind für pädagogisches Personal einzusetzen, was wiederum den erfolgreichen Erwerb der deutschen Sprache der Kinder fördern soll.

Vor dem Hintergrund dieser Fördermaßnahme wird im vorliegenden Beitrag folgenden Forschungsfragen nachgegangen:

1. Wie wird Sprachbildung in Kitas in NRW umgesetzt?
2. Gibt es hierbei Unterschiede zwischen Kitas mit unterschiedlicher finanzieller Förderung?

Um den Einsatz der finanziellen Fördermittel sowie die Umsetzung von Sprachbildung in Kitas in NRW genauer zu untersuchen, wurden im Sommer 2016 und 2017 leitfadenzentrierte Interviews mit insgesamt 27 Kita-Leitungen geführt. Für die Rekrutierung dieser Stichprobe wurden ausschließlich Kitas berücksichtigt, die zusätzliche finanzielle Fördermittel erhalten und in Jugendamtsbezirken in NRW liegen, deren Vertreter zuvor an einem Interview zu diesem Themenbereich teilgenommen hatten. Zusätzlich fand im Frühjahr 2017 eine standardisierte Online-Befragung statt. An dieser nahmen die Leitungen von 1679 der insgesamt ca. 9800 Kitas in NRW teil. Beide Befragungen untersuchten unter anderem, wie die zusätzlichen finanziellen Fördermittel verwendet werden, d.h. welches Personal mit welcher Qualifikation für wie viele Stunden in den Kitas beschäftigt wird und welche Aufgaben von diesen Fachkräften übernommen werden. Weiter sollte geklärt werden, wie Sprachbildung in den Kitas umgesetzt wird und ob sich hierbei Unterschiede zwischen Kitas in Abhängigkeit vom Erhalt zusätzlicher finanzieller Fördermittel zeigen. So wurde beispielsweise erfragt, wie die Sprachbildung erfolgt, welche Personen mit welchen Qualifikationen diese durchführen und welche Fortbildungen in den letzten 12 Monaten im Bereich „Sprache“ besucht wurden. Zusätzlich wurden Einstellungen zur Sprachbildung in der Kita abgefragt. Um zu klären, ob sich Kitas mit und ohne zusätzliche finanzielle Förderung in ihren Strukturmerkmalen voneinander unterscheiden, wurden Angaben zu Lage, Trägerschaft, Gruppenstruktur, Anzahl der Gruppen und Kinder, Anzahl der Kinder mit nicht-deutscher Herkunftssprache etc. erhoben.

Im vorliegenden Beitrag werden ausgewählte Ergebnisse dieser Befragungen präsentiert. Dabei wird insbesondere auf den Vergleich zwischen Kitas mit unterschiedlicher finanzieller Förderung eingegangen (Kitas ohne zusätzliche finanzielle Förderung vs. plusKITAs vs. Kitas mit zusätzlichem Sprachförderbedarf vs. Kitas, die beide Förderlinien erhalten). Analysen der Strukturdaten bestätigen die erwarteten Gruppenunterschiede: Unter anderem liegen Kitas, die keine finanziellen Fördermittel erhalten, häufiger in Landgemeinden oder Kleinstädten, sind seltener in öffentlicher Trägerschaft und haben eine geringere Anzahl an Betreuungsplätzen sowie an Kindern mit nicht-deutscher Herkunftssprache. Erste inhaltsanalytische Auswertungen der qualitativen Interviews zeigen trotz eines einheitlichen Verständnisses von (alltagsintegrierter) Sprachbildung eine heterogene Umsetzung in der Praxis und unterschiedliche Ansprüche an die Qualifikation der pädagogischen Fachkräfte im Bereich „Sprache“. Diese Ergebnisse werden von ersten Analysen der standardisierten quantitativen Online-Befragung mit N=1679 Kita-Leitungen untermauert. So lassen sich varianzanalytisch Gruppenunterschiede zwischen Kitas mit und ohne zusätzliche finanzielle Förderung bei der Umsetzung von Sprachbildung, der Qualifikation der Personen, die die Sprachbildung umsetzen sowie bei den Einstellungen zur Sprachbildung in der Kita feststellen.

Inwieweit diese und weitere Unterschiede auf die finanzielle Fördermaßnahme des Landes NRW zurückzuführen sind oder bereits im Vorfeld bestanden, bleibt eine offene Frage. Diese und weitere Ergebnisse werden im Hinblick auf Forschungs- und Handlungsimplicationen für den Elementarbereich diskutiert.

## **Kita-Qualität: Neue Einblicke in ein dynamisch-multikausales System aus Sicht der Netzwerkanalyse**

**Alexandru Agache, Katharina Kohl, Lilly Bihler, Jessica Willard, Birgit Leyendecker**  
Ruhr-Universität Bochum

Einerseits setzen gängige Erklärungsmodelle voraus, dass die Kita-Qualität ein komplexes Phänomen ist, jedoch wird sie typischerweise in simplen unidirektionalen X→Y-Modellen untersucht. So wird angenommen, dass strukturelle Rahmenbedingungen (z.B. Personal-Kind Schlüssel) Aspekte der Fachkraft-Kind Interaktionen prägen können (Phillipsen et al., 1997; Vermeer et al., 2016). Zur Konzeptualisierung der Interaktionsqualität selbst werden ebenfalls komplexe Beobachtungsverfahren eingesetzt, die aber letztendlich in der Evaluationspraxis auf wenige Mittelwertescores simplifiziert werden. So wird mit dem CLASS Pre-K Beobachtungsinstrument (Hamre et al., 2013) die Qualität der Fachkraft-Kind Interaktionen anhand von hunderten von behavioralen Markern in 10 Ratingdimensionen gewichtet (z.B. Positives Klima, Feinfühligkeit, Verhaltensmanagement, kognitive Anregung und Sprachbildung), die wiederum in drei Domänen unterteilt werden: die emotionale Unterstützung, die Organisation des Kita-Alltags und die Lernunterstützung. Das Drei-Domänen-Modell wurde in der bisherigen Forschung mit faktorenanalytischen Verfahren untersucht. Nach diesem Ansatz determinieren latente Variablen die beobachtete Varianz in den einzelnen CLASS-Dimensionen. Eine alternative Sicht bietet die Netzwerktheorie von komplex-dynamischen Systemen (Boresboom, 2017; Koller & Friedman, 2009). Demnach können die übergeordneten Domänen und deren Zusammenhänge mit strukturellen Gruppenvariablen aus einer dynamischen Verkettung der einzelnen Dimensionen (z.B. Feinfühligkeit→ Positives Klima→ Verhaltensmanagement→ Sprachbildung) resultieren. Die Wirkrichtungen können auch bidirektional sein (z.B. dadurch, dass eine Erzieherin, die häufiger in Konversationen mit den Kindern tritt, auch sensibler für deren Bedürfnisse wird).

### Fragestellungen

Hauptziel ist es, die reziproken Zusammenhänge zwischen den einzelnen Qualitäts-Dimensionen und deren Dynamik über einen typischen Vormittag mit einem netzwerkanalytischen Verfahren explorativ zu untersuchen.

Zusätzlich werden die Zusammenhänge zwischen der Interaktionsqualität und situationsbedingten Tagesabläufen (z.B. Essen/Außenbereich), fördernden Aktivitäten (z.B. Vorlesen/Singen), Personal-Kind Schlüssel und Gruppenkomposition (Anteil an Kindern mit Zuwanderungshintergrund/niedrigem Einkommen) untersucht.

### Methoden

Die Interaktionsqualität wurden mit CLASS-Pre-K an 4 wiederholten Messungen zusammen mit Beobachtungen zum Involvement der Fachkräfte und sprachfördernden Aktivitäten in 177 vorwiegend altersgemischten Kita-Gruppen aus NRW erfasst. Zusätzlich wurden strukturelle Gruppenmerkmale und sozio-demographische Daten zu allen Kindern erhoben.

In den Netzwerkanalysen wurden die multikausalen Zusammenhänge innerhalb (within-Analysen: fixed effects-Modelle; Bollen & Brand, 2010) und zwischen den Kita-Gruppen (between-Analysen: partielle Korrelationen mit graphischen LASSO; Epskamp & Fried, 2016) untersucht. Während in den within-Analysen zeitinvariante, ungemessene Einflüsse (z.B. Persönlichkeitsmerkmale) kontrolliert wurden, konnten in den between-Analysen strukturelle Gruppenvariablen berücksichtigt werden.

### Ergebnisse

In Übereinstimmung mit früheren Befunden fanden wir hohe Werte für die Domäne emotionale Unterstützung und Werte im niedrigen Bereich für Lernunterstützung. Autoregressiv- und Wachstumskurvenmodelle replizieren ebenfalls (von Suchodoletz et al., 2014) eine moderat-hohe zeitliche Stabilität, höhere Werte in den Morgenstunden und einen leicht abnehmenden Verlauf der Qualitätsdimensionen über den Vormittag.

Die Ergebnisse zur multivariaten Netzwerk-Struktur zeigen bisher ununtersuchte bidirektionale Effekte zwischen den Qualitätsdimensionen auf. Insbesondere bestehen starke reziproke Zusammenhänge innerhalb der Clusters zur emotionalen Unterstützung und zur Lernunterstützung (z.B. jeweils positive Effekte für: Sensitivität - positives Klima; kognitive Anregung - Feedbackqualität). Die Dimensionen der Lernarrangements (z.B. aktive Teilnahme der Kinder in Interaktionen/Lernaktivitäten; Vielfalt der Lernmaterialien und gute Organisation der Fachkräfte), das Stattfinden von sprachfördernden Aktivitäten und das Involvement der Erzieherinnen wirken dabei als vermittelnde „Brücke“ zwischen den Variablen des emotionalen Klimas und der Lernunterstützung. Die Essenszeiten und der Aufenthalt im Außenbereich werden hingegen nur sehr selten für kognitiv-sprachfördernde Aktivitäten genutzt, was vor allem durch ein geringes Involvement der Fachkräfte in dyadischen Interaktionen erklärbar ist. Zusätzlich zeigen die between-Analysen, dass die Gruppen mit geringerem Personalschlüssel und höherer Feinfühligkeit in Interaktionen auch höhere Werte im Bereich der Klassenorganisation und der Lernunterstützung aufweisen. Der Befund, dass der Personalschlüssel nur im Zwischenvergleich der Kita-Gruppen prädiktiv ist, deutet darauf hin, dass dies zwar eine wichtige, aber nicht entscheidende Voraussetzung für die Interaktionsqualität ist. Zusammenhänge zwischen Interaktionsqualität und Gruppenkomposition wurden nicht gefunden.

Zusammenfassend deuten Zentralitätsanalysen der Netzwerkstrukturen darauf hin, dass die wichtigsten Kandidaten für zukünftige Interventionen das Involvement, die sprachlich-kognitive Anregung und die Feinfühligkeit der Fachkräfte sind.

## **Effekte der Quantität und Qualität der Kita-Betreuung auf die Sprachkenntnisse von Kindern**

**Lilly Bihler, Alexandru Agache, Katharina Kohl, Jessica Willard, Birgit Leyendecker**  
Ruhr-Universität Bochum

### Hintergrund

Die Befundlage zu Effekten der Quantität und Qualität der Kita-Betreuung auf kindliche Sprachkenntnisse ist uneinheitlich. Mehrere Studien deuten darauf hin, dass eine höhere Qualität der Lernunterstützung einen signifikant positiven Effekt auf den rezeptiven Wortschatz der Kinder hat (Keys et al., 2013; Mashburn et al., 2008), während dieser Effekt in einer jüngsten Metaanalyse von Perlman et al. (2016) nicht bestätigt werden konnte. Weiterhin wurden mögliche Moderationseffekte mit der Familiensprache bisher nur vereinzelt untersucht (Biedinger, Becker, & Rohling, 2008; Klein & Becker, 2017). Gerade eine solche differenzierte Betrachtung ist jedoch wichtig, um die Sprachkenntnisse von Kindern angemessen zu fördern.

### Fragestellung

Die vorliegende Studie untersucht, wie Quantität und Qualität der Kita-Betreuung mit den Sprachkenntnissen von Kindern zusammenhängen. Ebenso werden mögliche Moderationseffekte der Familiensprache betrachtet.

### Methoden

Die analysierten Daten stammen aus der Studie SEIKA-NRW., in der die Sprachkenntnisse von N = 2231 Kinder im Alter von 2 ½ - 6 Jahren untersucht wurden. Die Kinder besuchten 177 Kita-Gruppen in 95 Kitas. Als Indikator für die Sprachkenntnisse wurde der rezeptive Wortschatz mit dem deutschsprachigen PPVT (Lenhard, Lenhard, Segerer, & Suggate, 2015) erfasst. Für eine Teilstichprobe wurden weiterhin zwei Subskalen des Sprachentwicklungstests SETK 3-5 (Grimm, 2015), morphologische Regelbildung (n = 802) und phonologisches Arbeitsgedächtnis (n = 712), herangezogen.

Die Quantität der Kita-Betreuung wurde über das Eintrittsalter in die Kita operationalisiert. Zur Erfassung der Betreuungsqualität wurde das Classroom Assessment Scoring Systems (CLASS Pre-K; Pianta, La Paro, & Hamre, 2015) mit den Domänen emotionale Unterstützung, Klassenorganisation und Lernunterstützung eingesetzt.

#### Ergebnisse

Für den rezeptiven Wortschatz zeigte sich ein Interaktionseffekt zwischen der Familiensprache und dem Eintrittsalter: Mehrsprachige Kinder, die mit einem Jahr in die Kita kamen, schnitten um den Normbereich ab, mit steigendem Eintrittsalter nahmen die Wortschatzwerte ab. Für einsprachige Kinder dagegen zeigte sich kein signifikanter Effekt des Eintrittsalters. Für die Domäne Lernunterstützung der Kita-Qualität zeigte sich ein ähnlicher Effekt: Mehrsprachige Kinder, die eine Einrichtung mit hoher Qualität der Lernunterstützung besuchten, erreichten deutlich höhere Wortschatzwerte als mehrsprachige Kinder, die eine Einrichtung mit niedriger Qualität besuchten. Für einsprachige Kinder zeigte sich wiederum kein Effekt. Dieser Interaktionseffekt zwischen der Qualität der Lernunterstützung und der Familiensprache zeigte sich auch für morphologische Regelbildung. Für das phonologische Arbeitsgedächtnis fanden sich erwartungskonform weder Effekte der Quantität noch der Qualität der Kita-Betreuung.

#### Schlussfolgerungen

Die Ergebnisse verdeutlichen die Bedeutung einer differenzierten Betrachtung der Effekte von Kita-Quantität und -Qualität für ein- und mehrsprachige Kinder und unterstreichen die Bedeutung der Kita für die Sprachentwicklung mehrsprachiger Kinder, welche von einer hoher Qualität und einem frühen Eintrittsalter profitieren.

## D12

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H119

### Kompetenz durch Kohärenz - Der Freiburger Weg der Lehrerbildung

Chair(s): **Katharina Hellmann** (Pädagogische Hochschule Freiburg, Deutschland)

DiskutantIn(nen): **Andreas Borowski** (Universität Potsdam)

Der Lernerfolg von Schülerinnen und Schülern setzt eine kompetente Lehrkraft und folglich ein qualitativvolles, d.h. kompetenz- und professionsorientiertes, Hochschulstudium voraus (z.B. Beck, 2008; Hattie, 2010; Krauss et al., 2008). Der Professionserwerb von Lehrkräften vollzieht sich in der Hochschulphase dabei v.a. durch den Erwerb zentraler Wissenskomponenten (z.B. Baumert & Kunter, 2006; Loewenberg-Ball, Thames & Phelbs, 2008). Diese Wissenskomponenten werden traditionell durch die Säulen Fachwissenschaft, Fachdidaktik und Bildungswissenschaften vermittelt. Es wird jedoch kritisiert, dass diese Trennung fragmentiertes Wissen produziert, welches von den Studierenden nur mühsam integriert oder adaptiv in der Unterrichtspraxis genutzt werden kann; den Anforderungen eines zunehmenden komplexeren Lehrberufs werde man auf diese Weise nicht gerecht (z.B. Darling-Hammond, 2013; Kotthoff & Terhart, 2013). Daher scheint es ratsam, die Verknüpfung der verschiedenen Wissensbereiche bereits frühzeitig im Hochschulstudium zu stärken und strukturell sowie curricular in der Lehrerbildung zu implementieren. Mit diesem Ziel werden in Freiburg an der Pädagogischen Hochschule und Universität derzeit neue Studienstrukturen und Lehr-Lern-Konzepte in Pilotprojekten entwickelt und beforscht.

Das Konzept der Kohärenz („Kompetenz durch Kohärenz“) dient hierbei als zentrale theoretische Annäherung an die Problematik. Kohärenz wird als die Generierung von Lerngelegenheiten verstanden, bei welchen in vertikaler (über den Studienverlauf; Theorie-Praxis-Bezug) und horizontaler Richtung (über die Säulen der Lehrerbildung) Bezüge geschaffen werden, die es den Studierenden ermöglichen, ihr Studium über verschiedene Lehrveranstaltungen hinweg sowohl strukturell als auch inhaltlich als zusammenhängend und sinnhaft zu erleben (z.B. Biggs, 2003; Buchman & Floden, 1991; Canrinus, Bergem, Klette, & Hammerness, 2015; Hammerness, 2006). Das Erfahren dieser Zusammenhänge sollte sich auf die nachfolgende Lehrtätigkeit übertragen und zu professionellem und kompetentem Handeln im Lehrberuf befähigen. In verschiedenen Lehrentwicklungsprojekten werden derzeit Strukturen aufgebaut sowie Konzepte ausgearbeitet und erprobt, welche einer Professionsorientierung und einem Kompetenzzuwachs bei Studierenden durch Kohärenzförderung Rechnung tragen wollen.

Im Symposium werden empirische Befunde verschiedener Disziplinen (Erziehungswissenschaft, Physik-, Mathematik-, Romanistikdidaktik) vorgestellt, welche sich der Bestimmung von Kohärenz sowie entsprechenden Zusammenhängen des Kohärenzerlebens mit Merkmalen auf Studierendenseite widmen. Im ersten Beitrag (Hellmann & Henning-Kahmann; Erziehungswissenschaft) werden Ergebnisse einer Studierendenbefragung präsentiert, welche Einblick darin geben, inwieweit und unter welchen Umständen Studierende verschiedener Studiengänge und unterschiedlicher Semester Kohärenz im Studienverlauf auf struktureller und curricularer Ebene wahrnehmen. Der zweite Beitrag (Schwichow, Physik) widmet sich über einen Inverted-Classroom-Ansatz der Kohärenz von Theorie und Unterrichtspraxis, sowie Effekten auf studentische Lernmotivation. Der dritte Beitrag (Wessel, Mathematik) zeigt exemplarisch für die Analysisdidaktik einen Vorschlag zur Schaffung von Theorie-Praxis-Kohärenz mit dem 4C-ID Model auf und beleuchtet situative Wirkungen dieses Ansatzes auf Kompetenzentwicklung. Im vierten Beitrag (Zaki, Romanistik) werden anhand qualitativer Lernportfolioanalysen Effekte einer kohärenten Lernveranstaltungsgestaltung auf die Studierendenleistung dargestellt.

Abschließend sollen Fragen einer „effektiven“ bzw. „guten“ Hochschullehre diskutiert werden. Welche Auswirkungen hat die Implementation kohärenter Studienstrukturen für Studierende? Wie kann diese Implementierung effektiv erfolgen? Welche Herausforderungen und Konsequenzen ergeben sich hieraus für die Hochschulentwicklung und -lehre? Wie können Effekte dieser Maßnahmen (empirisch) bestimmt werden? Und schließlich: Inwieweit kann durch solche Maßnahmen den aktuellen Herausforderungen einer professionsorientierten Lehrerbildung angemessen begegnet werden?

#### Beiträge des Symposiums

### Kohärenzwahrnehmung von Lehramtsstudierenden - Ebenen und Einflussfaktoren

**Katharina Hellmann, Jan Henning-Kahmann**

Pädagogische Hochschule Freiburg, Deutschland

#### Theoretischer Hintergrund und Fragestellung

Die Schaffung kohärenter Studienstrukturen sowie Lehr-Lern-Gelegenheiten sollte strukturell und inhaltlich vernetztes Wissen generieren und so der Forderung nach einer stärkeren Professionsorientierung der Lehramtsausbildung Rechnung tragen (z.B. Darling-Hammond, 2013; Kotthoff & Terhart, 2013). Generell sind hierbei verschiedene Kohärenzebenen unterscheidbar. Beispielsweise kann auf Hochschulebene ein kohärentes Leitbild („vision“) zum Lehrberuf vermittelt werden. Dieses entwickelt sich idealerweise über das Gesamtstudium und wird von allen Akteuren der Lehrerbildung gleichermaßen geteilt. Vertikale Kohärenz beschreibt differenzierter strukturelle und curriculare Verknüpfungen, die über den Studienverlauf hinweg geschaffen werden und mit fortschreitendem Studium anwachsen sollten. Horizontale Kohärenz liegt vor, wenn Verknüpfungen zwischen den Säulen Fachwissenschaft, Fachdidaktik und Bildungswissenschaften geschaffen werden (Hammerness & Klette, 2015; Hammerness, Klette & Bergem, 2014; Laschke & Blömeke, 2014).

Es wird untersucht, inwieweit Lehramtsstudierende am Hochschulstandort Freiburg Kohärenz auf diesen Ebenen wahrnehmen. Das Kohärenzerleben sollte sich dabei durch ein von der Hochschule vermitteltes Leitbild aufzeigen lassen, sowie weiterhin durch eine wahrgenommene Verknüpfung innerhalb und zwischen den Säulen der Lehrerbildung. Auch der Studienfortschritt soll als Einflussvariable untersucht werden. Die gewonnenen Ergebnisse sollten erste Aufschlüsse über mögliche Einflussfaktoren bezüglich Kohärenzerleben geben und Hinweise auf potentielle Anknüpfungspunkte für die Studiengestaltung liefern.

#### Methode

Zur Erfassung des Kohärenzerlebens wurde in Anlehnung an die oben genannten Arbeiten ein Fragebogen entwickelt, welcher u.a. das wahrgenommene Leitbild (4 Items,  $\alpha=,94$ , Beispielitem: „Die Hochschule hat im Verlauf meines gesamten Lehramtsstudiums ein Leitbild zum Lehrberuf stimmig vermittelt“), die Wahrnehmung des Gesamtstudiums („Ich fühle mich auf die vielfältigen und komplexen Herausforderungen des Lehrberufs durch mein Studium gut vorbereitet“), die Vernetzung innerhalb („Inhaltlich bauten spätere Lehrveranstaltungen der Fachwissenschaft auf früheren auf“) sowie zwischen den Säulen („Es gab inhaltliche Verknüpfungen zwischen den jeweiligen Lehrveranstaltungen aus der Fachdidaktik und den

Bildungswissenschaften“) auf einer fünfstufigen Likert-Skala, sowie wichtige Hintergrundvariablen erfasst. Studierende verschiedener Studiengänge und Fachbereiche (N=359; Studiendauer: M=5.31, SD=2.69) bearbeiteten den Fragebogen zum Ende des Sommersemesters 2017. Daten wurden deskriptivstatistisch sowie korrelativ ausgewertet.

#### Ergebnisse

Das von der Hochschule vermittelte Leitbild zum Lehrberuf wurde von den Studierenden im Durchschnitt als nur teilweise vorhanden eingeschätzt (M=2.45, SD=1.16). Dabei konnte kein signifikanter Zusammenhang des Studienfortschritts mit der Beurteilung dieses leitbildbezogenen Kohärenzerlebens festgestellt werden ( $r=-.10$ , ns). Gleichwohl lässt sich zeigen, dass mit einer steigenden Anzahl besuchter fachwissenschaftlicher Lehrveranstaltungen eine signifikant geringe Wahrnehmung eines kohärenten Leitbildes einhergeht ( $r=-.28$ ,  $p<.001$ ). Die Anzahl der besuchten fachdidaktischen ( $r=.22$ ,  $p=.001$ ) und bildungswissenschaftlichen ( $r=.18$ ,  $p=.006$ ) Lehrveranstaltungen steht hingegen in einem signifikant positiven Zusammenhang mit dem leitbildbezogenen Kohärenzerleben.

Das Gefühl, auf die komplexen Herausforderungen des Lehrberufs gut vorbereitet zu sein, war bei den Studierenden nur mittelmäßig ausgeprägt (M=2.93, SD=1.04) und wies ebenfalls keinen signifikanten Zusammenhang mit dem Studienfortschritt auf ( $r=.01$ , ns). Es zeigte sich jedoch, dass Studierende, die ein Leitbild zum Lehrberuf wahrnahmen, sich tendenziell auch besser auf diese Herausforderungen vorbereitet fühlten ( $r=.38$ ,  $p<.001$ ).

Mehrheitlich gaben die Studierenden zudem an, einen klaren Sinn darin zu sehen, fachwissenschaftliche, fachdidaktische und bildungswissenschaftliche Studienanteile eng miteinander zu verknüpfen (M=4.24, SD=1.00). Die vorhandenen Verknüpfungen innerhalb der einzelnen Säulen (FW: M=3.54, SD=1.15; FD: M=3.39, SD=1.16; BW: M=3.05, SD=1.12) und insbesondere zwischen den Säulen (FW-FD: M=3.28, SD=1.16; FW-BW: M=2.49, SD=1.17; FD-BW: M=3.27, SD=1.09) wurden von den Studierenden jedoch nur als mittelmäßig eingeschätzt.

#### Implikationen

Die Ergebnisse implizieren zunächst, dass eine selbstständige Verknüpfung von Studieninhalten über den Studienverlauf nur bedingt stattfindet. Angehende Lehrkräfte sollten daher von Hochschuleseite durch geeignete strukturelle und curriculare Vernetzungsmöglichkeiten darin unterstützt werden. Weiterhin weisen die meist hohen Standardabweichungen der Ergebnisse auf eine Heterogenität innerhalb der Stichprobe hin. Es müssen daher weiterführende Faktoren herausgearbeitet werden, die das „Wann“ und „Wie“ studentischen Kohärenzerlebens erklären und differenziert aufzeigen, welche Studierenden („Bei wem?“) bei welchen („Wo“?) Lehr-Lern-Gelegenheiten Kohärenz wahrnehmen sowie welche Zusammenhänge dieses Kohärenzerlebens zu studentischer Motivation und Kompetenzerwerb hat.

### **Wirkung einer fachdidaktischen Lehrveranstaltung im Inverted-Classroom Ansatz auf die Motivation der Lernenden**

**Martin Schwichow, Silke Mikelskis-Seifert**

Pädagogische Hochschule Freiburg, Deutschland

Zum Aufbau eines fundierten fachdidaktischen Wissens, das angehende Lehrkräfte bei der Planung, Durchführung und Reflexion von Unterricht anwenden können, sollten sich Didaktikveranstaltungen sowohl an Forschungsbefunden als auch an der Unterrichtspraxis orientieren. Fachdidaktische Lehrveranstaltungen stehen somit in einem Spannungsfeld zwischen fundierter Einführung in die Fachdidaktik als empirische Bildungsforschung und der Bearbeitung praxisnaher Aufgabenstellungen. In traditionellen Vorlesungen, die nur auf die Vermittlung von Konzepten und Wissen ausgerichtet sind, kann eine solche Kohärenz zwischen Theorie und Unterrichtspraxis kaum verwirklicht werden (Borko & Putnam, 1996). Um die genannte Ausgangssituation zu verbessern, wurde das Konzept des Inverted-Classroom Ansatz nach Spannagel (2014) auf die Physikdidaktikausbildung übertragen. Er sieht eine Umkehrung der üblichen Aktivitäten innerhalb und außerhalb des Hörsaals vor. Den Studierenden werden Inhalte nicht mehr vor Ort vermittelt, sondern sie erarbeiten sich diese, selbstgesteuert und im eigenen Lerntempo, anhand von digitalen Lernmaterialien. Die Präsenzzeiten dienen der gemeinsamen Vertiefung über Anwendungsaufgaben. Lehrveranstaltungen nach dem Inverted-Classroom Konzept gliedern sich entsprechend in Selbstlern-, Präsenz- und Diagnosephase (Fischer & Spannagel, 2012). In den Selbstlernphasen erarbeiten sich die Studierenden fachdidaktische Inhalte anhand von digitalen Lernmaterialien. Zur Selbstkontrolle erhalten sie Verständnisaufgaben und werden aufgefordert, Fragen zu formulieren, an denen in der Präsenzphase gearbeitet werden kann. In den Präsenzphasen erfolgt die explizite Theorie-Praxis-Verzahnung, indem das fachdidaktische Wissen in für die Studierenden sinnstiftenden Aufgaben angewendet wird.

Lehrveranstaltungen im Inverted-Classroom Ansatz sollten sich nach der Selbstbestimmungstheorie von Deci und Ryan (1993) positiv auf die Motivation der Studierenden auswirken. So ermöglichen die Selbstlernphasen ein Autonomieerleben, die Arbeit an Praxisbeispielen ein Kompetenzerleben und kooperative Lernformen in den Präsenzphasen das Erleben sozialer Eingebundenheit. Da es sich bei fachdidaktischem Wissen um ein Amalgam aus Fachwissen und erziehungswissenschaftlichem Wissen handelt (Shulman, 1987), sollte die Motivation allerdings vom Vorwissen der Studierenden in beiden Bereichen abhängen. Darüber hinaus unterscheiden sich Lehrveranstaltungen zwischen unterschiedlichen Studiengängen (Primar-, Sekundar- und gymnasiales Lehramt) stark hinsichtlich der Fachinhalte und methodischen Gestaltung. Zur Evaluation der Lehrveranstaltung haben wir folgende Forschungsfragen untersucht: 1) Wie schätzen Studierende unterschiedlicher Studiengänge die Berufsorientierung, die soziale Eingebundenheit sowie das Autonomie- und Kompetenzerleben während der Veranstaltung ein? 2) Inwiefern hängt die Motivation der Studierenden von ihrem erziehungswissenschaftlichem und fachlichen Wissen ab?

#### METHODEN

Zur Evaluation des Lehrkonzepts wurden das Fachwissen Mechanik (16 Items,  $\alpha=.93$ ), das erziehungswissenschaftliche Wissen (Ewi, 7 Items,  $\alpha=.86$ ) vor, sowie die Motivation nach der Lehrveranstaltung (15 Items,  $\alpha=.80$ ) erhoben. Es wurden Daten von 44 Studierenden aus drei Lehrveranstaltungen zur Einführung in die Physikdidaktik erhoben. Der Einfluss des erziehungswissenschaftlichem und fachlichen Wissen auf die Motivation wurde mittels Regressionsanalyse, und die Unterschiede in der Wahrnehmung der Lehrveranstaltung mittels Varianzanalyse untersucht.

#### ERGEBNISSE & DISKUSSION

Die Studierenden unterschieden sich hinsichtlich der wahrgenommenen sozialen Eingebundenheit ( $F(2,40)=5.87$ ,  $p<.01$ ,  $M_{\text{Primar}}=0.46$ ,  $M_{\text{Sekundar}}=0.25$ ,  $M_{\text{Gym}}=-0.57$ ) und des Autonomieerlebens ( $F(2,40)=7.25$ ,  $p<.01$ ,  $M_{\text{Primar}}=-0.55$ ,  $M_{\text{Sekundar}}=0.80$ ,  $M_{\text{Gym}}=0.44$ ). Die Unterschiede können vermutlich auf unterschiedliche Lernerfahrungen zurückgeführt werden. Während im gymnasialen Lehramt klassische Vorlesungen überwiegen, überwiegen im Primarbereich Seminare mit

kooperativen Lernformen. Entsprechend erleben Studierende des Grundschullehramts eventuell eine geringere Autonomie im Vergleich zu üblichen Lehrveranstaltungen ihres Studiengangs. Entgegen unserer Erwartung zeigen die Ergebnisse der Regressionsanalyse (Konstante=0.00,  $\beta_{\text{Fachwissen}}=0.36$  SE=0.15  $p<0.05$   $R^2=0.13$ ), dass die Motivation nur vom Fachwissen und nicht vom erziehungswissenschaftlichem Wissen abhängt.

Die Ergebnisse verdeutlichen, dass die Wirkung von Lehrveranstaltungen im Inverted-Classroom Ansatz die vermutete Wirkung auf die Motivation hat. Allerdings unterscheidet sich diese Wirkung entsprechend der Lernerfahrungen der Studierenden. Um die motivierende Wirkung für alle Studiengänge zu erhalten, ist folglich eine Anpassung des Veranstaltungskonzepts an einzelne Studiengänge erforderlich. Unsere Befunde unterstreichen die Bedeutung von Wissen über studiengangsspezifische Lehrveranstaltungen zur Erhöhung der Kohärenz. Dies kann u.a. durch kollegiale Hospitationen oder überfachliche Lehr tandems erlangt werden.

### **Kohärenz in der Mathematiklehrerausbildung– Verbindung von Theorie und Praxis durch 4CID am Beispiel Analysisdidaktik**

Lena Wessel

Pädagogische Hochschule Freiburg, Deutschland

Theoretischer Hintergrund: Bezüglich der Struktur von Lehrerkompetenzen hat sich seit Shulman (1987) ein Facettenmodell herausgeschält, das pädagogische, fachwissenschaftliche und fachdidaktische Kompetenzen unterscheidet und empirisch erfassbar macht. In der Schulpraxis werden diese Kompetenzen in der Regel jedoch nie einzeln, sondern in enger Verbindung miteinander und in berufsspezifischen Kontexten aktiviert (Ball et al. 2008). Auch wenn die Frage, in welcher Weise zukünftige Lehrkräfte im Studium die für die spätere Praxis notwendigen Kompetenzen erwerben bislang nur unzureichend erforscht ist (Grossman 2008), gibt es Argumente, diese Kompetenzen in Vernetzung der später im Berufsleben notwendigen Kompetenzfacetten zu lehren (ebd.), u.a. um der Gefahr der Ausbildung trägen Wissens zu entgegenen (Renkl et al. 1996).

Die fachdidaktischen Studienanteile im Lehramtsstudium Mathematik bewegen sich potentiell in einer weiten Spanne zwischen systematischem theoretischem Wissen über Theorien und Prinzipien zum Lehren und Lernen von Mathematik und konkreten Beispielen und Tätigkeiten aus der täglichen Unterrichtspraxis. Für die Didaktik der Analysis als ein zentraler Gegenstand im gymnasialen Lehramtsstudium ist das theoretische Wissen (als Teil des topic-specific pedagogical content knowledge (PCK), Hadfield et al. 1998) bereits gut konsolidiert (z.B. Danckwerts & Vogel 2006). Die Autoren beschreiben, welche Gegenstände der Didaktik der Analysis in der Regel zugerechnet werden. Wie Lehramtsstudierende Fähigkeiten der Anwendung des Wissens erwerben ist hieraus jedoch nicht zu ersehen.

Einen Ansatz für die Konzeption von Hochschullehre, die den Erwerb komplexer Fähigkeiten für das Handeln in authentischen Situationen fördert, liefert als konkretes Instruktionsmodell das 4CID Modell (Four Component Instructional Design) nach Merriënboer & Kirschner (2007). Dieses enthält vier spezifische Komponenten (Lernaufgaben, Wissen als unterstützende Informationen, Hilfen zu Prozeduren, Übungen von Teilfähigkeiten), die als Design-Elemente und Design-Prinzipien zur Erhöhung der Theorie-Praxis-Kohärenz in Lehrveranstaltungen im Mathematik-Lehramt vielversprechend scheinen. Bedarf besteht jedoch bzgl. der konkreten Entwicklung und sorgfältigen Dokumentation entsprechender Lehrveranstaltungen sowie der Beforschung ihrer Wirksamkeit und Wirkungen auf die Kompetenzentwicklung der Studierenden (Überblick über PCK-Forschung im Bereich Mathematik geben Depape, Verschaffel & Kelchtermans 2013).

Fragestellungen: Entsprechend der genannten Desiderata werden die folgenden Fragestellungen auf Entwicklungs- und Forschungsebene bearbeitet:

1. Wie kann eine Didaktik der Analysis Veranstaltung für Gymnasiallehramtsstudierende entsprechend des 4CID Modells so konzipiert werden, dass die fachdidaktischen Kompetenzen in enger Verbindung zum Kontext ihrer Anwendung erworben werden können?
2. Welche situativen Wirkungen der konzipierten Veranstaltung können bzgl. der Entwicklung fachdidaktischer Kompetenzen der Studierenden rekonstruiert werden?

Methode: Das Projekt verschränkt die Theorie zum 4CID Instruktionsmodell mit dem Ansatz Fachdidaktischer Entwicklungsforschung (Prediger et al. 2012) für Hochschullehrentwicklung. In drei Entwicklungszyklen (SS16, WS16/17, SS17) wurden 14 Seminarsitzungen entsprechend der Design-Prinzipien nach 4CID im kollegialen Austausch unter den (Mathematik-)didaktikerInnen an der PH Freiburg konzipiert, mit Lehramtsstudierenden der Universität Freiburg in Design-Experimenten erprobt (n=34) und sukzessive weiterentwickelt. Zur Beantwortung der zweiten Fragestellung wurden über zwei Semester alle schriftlichen Studierendenbearbeitungen der Lernaufgaben und Übungen von Teilfähigkeiten, die jeweils auf die Ausbildung komplexer Handlungsfähigkeiten zielen, gesammelt. Zusätzlich wurden für die Handlungssituationen „Lernendenendenken anhand von authentischen Aufgabenlösungen analysieren“ und „Unterrichtliche Zugänge zur Analysis bewerten und vergleichen“ die schriftlichen Bearbeitungen von n=7 Studierenden im Prä-Post-Vergleich mittels qualitativer Inhaltsanalyse nach Mayring (2000) hinsichtlich Unterschieden und Entwicklungen analysiert.

Ergebnisse. Das Seminarkonzept mit theoriegeleiteter Ausarbeitung typischer authentischer Handlungssituationen und zugehörigen, aufeinander aufbauenden Lernaufgaben leistet einen wesentlichen Beitrag zur Theorie-Praxis-vernetzenden Spezifizierung von Kompetenzfacetten und unterstützenden Prozeduren in der Analysisdidaktik, die im Vortrag jeweils aufgefaltet werden. Der Prä-Post-Vergleich von Studierendenbearbeitungen zeigt für beide untersuchten Handlungssituationen ein Kontinuum auf, in das die Analysen und Bewertungen bzgl. ihrer fachdidaktischen Fundierung von geringer zu hoher Fundierung eingeordnet werden können. Dieses Kontinuum soll vor dem Hintergrund des implementierten 4CID Modells diskutiert werden.

### **Kohärenz durch E-Portfolioarbeit - am Beispiel eines Lehrentwicklungsprojekts zur Vernetzung von PCK und CK in der Fremdsprachenlehrer\*innenausbildung**

Katja Zaki

Pädagogische Hochschule Freiburg, Deutschland

Theoretischer Hintergrund

Kompetenzorientierte Reformen der Fremdsprachenlehrer\*innenausbildung rückten in den vergangenen Jahren zunehmend fachdidaktische und bildungswissenschaftliche Aspekte in den Vordergrund, um die traditionell meist stark fachwissenschaftlich geprägten Studiengänge professionsorientierter zu gestalten (vgl. de Florio-Hansen 2015, Legutke 2016). Auf der Grundlage der Modellierung unterschiedlicher Wissensbereiche bzw. -facetten professioneller Handlungskompetenz (vgl. Shulman 1987,

Baumert & Kunter 2006) stellt sich allerdings nicht nur die Frage nach der Gewichtung, sondern auch nach deren (handlungs-)kompetenzorientierter Verknüpfung im Sinne einer „kohärenten“ Kompetenzentwicklung bzw. Ausbildung.

#### Problemstellung

Eine umfassende professionelle Lehrer\*innenkompetenz, die sich in beruflicher Praxis wie in komplexen Unterrichtssituation als handlungsrelevant und abrufbar erweist, erfordert von Fremdsprachenlehrkräften schließlich nicht nur fundiertes fachwissenschaftliches, sprachpraktisches, fachdidaktisches und bildungswissenschaftliches Wissen, sondern die Fähigkeit, Verknüpfungen zwischen unterschiedlichen Komponenten professioneller Kompetenz herzustellen bzw. diese situations- und themenadäquat abzurufen (Renkl, Mandl & Gruber, 1996). Hierfür sollte bereits im Studium auf eine stärkere Vernetzung der bislang oft fragmentiert wahrgenommenen Ausbildungsanteile und -phasen hingewirkt werden (vgl. Canrinus et al. 2015).

Eine kohärente Professionalisierung von Fremdsprachenlehrkräften erfordert folglich eine „ganzheitliche Architektur“ (vgl. Kirschner/van Merriënboer 2007). Erforderlich sind evidenzbasierte Ansätze, die curriculare Strukturen, modulare Zielvorgaben, Lehrveranstaltungskonzepte sowie die aktive Rolle der Studierenden integrativ in den Blick nehmen (vgl. Legutke 2016). Die romanistischen Teilprojekte im FACE setzen vor diesem Hintergrund auf eine kollegiale Lehrentwicklung zur Steigerung horizontaler (Vernetzung von Fachwissenschaft/Sprachpraxis, Fachdidaktik und Bildungswissenschaften) sowie vertikaler (Theorie-Praxis-Verzahnung) Kohärenz. Einen ersten Knotenpunkt bildet die Einführung in die Fachdidaktik der Romanischen Sprachen im „Lehramtsoptionsbereich“ des polyvalenten Bachelorstudiengangs, die in Kooperation mit Kolleg\*innen der Bildungswissenschaften und der Hochschuldidaktik konzipiert wurde. (Die Vorlesung ist hochschulübergreifend angelegt und wird von Studierenden der Schulfremdsprachen Französisch, Italienisch und Spanisch besucht; N=80 im ersten Zyklus, Wintersemester 2016/17.)

#### Forschungsfragen

Angesichts der angestrebten Professions- und Kohärenzorientierung stellte sich für das fokussierte Lehrentwicklungsprojekt die Frage, wie sich fachdidaktisches und bildungswissenschaftliches Wissen auf Lehrveranstaltungsebene vernetzt vermitteln lässt und insbesondere ein kursübergreifendes e-Portfolio (mit Lernaufgaben und leitfragengestützten Lerntagebüchern) zur aktiven Wissensintegration und Reflexion beiträgt:

1) Wie kann ein kurs-/phasenübergreifendes ePortfolio die aktive (selbstregulierte) Wissensintegration und -vernetzung (PCK / CK) auf Studierendenseite unterstützen und fördern? Zeigt das gepromptete Aufgabendesign mit Lernaufgaben und leitfragegestützten Lerntagebüchern hier Vorteile gegenüber dem freien Schreiben?

2) Wie wirksam sind die durchgeführten Maßnahmen hinsichtlich

a) der Kohärenzwahrnehmung unter Studierenden?

b) der aktiven Wissensintegration (von fachdidaktischem und bildungswissenschaftlichem bzw. vernetztem Wissen)?

#### Methodisches Vorgehen

Den forschungstheoretischen Rahmen des skizzierten Lehrentwicklungsprojekts bildet ein Design Based Research-Ansatz. Die Entwicklung und Evaluation der Maßnahmen erfolgte im ersten Zyklus durch eine Fragebogenerhebung (Prä-Post-Test im paper pencil-Format zu Semesterbeginn und -ende, Items u.a. zur fachdidaktischen / bildungswissenschaftlichen Konzeptbildung und Kohärenzwahrnehmung) sowie durch die Auswertung der e-Portfolios (leitfragengestützte Lerntagebücher ergänzt um fünf Lernaufgaben) unter Verwendung qualitativer Inhalts- (Mayring 1980) und textlinguistischer Analyseansätze (vgl. u.a. Brinker 2010). Die Evaluationsergebnisse bilden die Basis der Weiterentwicklung der Veranstaltung im zweiten Zyklus.

#### Ergebnisse und Diskussion

Die Auswertung einer ersten Stichprobe an Prä-Post-Fragebögen (n: 17) und ausgewählten Portfolio-Aufgaben zeigt neben dem Ausbau fachdidaktischer Theorie-Konzeptbildung eine signifikante Zunahme der wahrgenommenen Kohärenz zwischen Fachdidaktik und Bildungswissenschaften. (Die Nennungen der wahrgenommenen Vernetzungen und die Häufigkeit exemplarisch ausgewählter Themenbereiche zeigt dabei, dass insbesondere explizite gemachte Verknüpfungen von Bildungswissenschaften und Fachdidaktik in Erinnerung blieben). Wichtige Knotenpunkte der „Vernetzung“ waren in der Einschätzung der Studierenden die Lernaufgaben selbst, (gefolgt von digitalen Verbindungen der Kursbereiche bzw. Hyperlinks auf frühere Foliensätze). Als erstaunlich wenig hilfreich wurden hingegen die Einträge ins Lerntagebuch bewertet, die im Rahmen der textlinguistisch unterfütterten qualitativen Inhaltsanalyse andererseits die detailliertesten Einblicke in die die Entwicklung einer nicht nur fachdidaktischen, sondern vernetzenden Reflexionskompetenz zuließen. Eine vergleichende Analyse von freien und leitfragengestützten Aufgaben zeigt in diesem Zusammenhang zudem die Wichtigkeit ausgewählter Prompts zur Unterstützung des gewünschten selbstregulierten Vorgehens (zumindest für diese Zielgruppe bzw. in der Studiengangphase).



## D13

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H118

### „Beliefs = Beliefs?“ – Überzeugungen von Lehrkräften aus der Perspektive verschiedener Fachdidaktiken

Chair(s): **Hendrik Lohse-Bossenz** (Pädagogische Hochschule Heidelberg, Deutschland), **Markus Rehm** (Pädagogische Hochschule Heidelberg, Deutschland), **Tobias Dörfler** (Pädagogische Hochschule Heidelberg, Deutschland)

DiskutantIn(nen): **Thilo Kleickmann** (Christian-Albrechts-Universität Kiel)

Überzeugungen („Beliefs“) von Lehrkräften werden im Modell von Baumert und Kunter (2006) neben Professionswissen, motivationalen Orientierungen und selbstregulativen Fähigkeiten als ein Aspekt professioneller Kompetenz konzeptualisiert. Sie stellen jedoch im Unterschied zu anderen Kompetenzaspekten des Modells kein einheitliches, klar definiertes Konstrukt dar. Vielmehr handelt es sich um einen Oberbegriff, der in seinen Teilaspekten spezifisch zu konzeptualisieren und operationalisieren ist.

Sowohl die theoretische Konzeptualisierung wie die empirische Operationalisierung kann in diesem Zusammenhang entweder fachspezifisch (z. B. Überzeugungen zur Nature-of-Science in der Naturwissenschaftsdidaktik) oder fachunabhängig (z. B. lehr-lern-theoretische Überzeugungen) erfolgen. Während fachspezifische Überzeugungen per definitionem ausschließlich innerhalb der jeweiligen Domänen untersucht werden, finden sich fachunabhängige Überzeugungen in verschiedenen Studien, welche Lehrkräfte unterschiedlicher Fächer untersuchen.

Das Ziel des Symposiums besteht darin, aus der Perspektive verschiedener Fächer (spezifisch Naturwissenschaften, Deutsch und Geschichte) Befunde zu professionellen Überzeugungen von Lehrkräften darzustellen und hierbei herauszuarbeiten, an welchen Stellen Gemeinsamkeiten und Unterschiede existieren. Fokussiert werden hierbei die verwendeten Konstrukte, Methoden und Ausprägungen.

Im ersten Beitrag wird das Wissenschaftsverständnis (in Anlehnung an das Nature-of-Science-Konzept) von angehenden Naturwissenschaftslehrkräften als Aspekt fachdidaktischen Wissens konzeptualisiert und mittels Videovignetten operationalisiert. Die Ergebnisse deuten auf die reliable Erfassung des Konstrukts hin. Überdies lässt sich die postulierte faktorielle Struktur empirisch aufzeigen. Der zweite Beitrag thematisiert für das Fach Deutsch die Überzeugungen zur Bedeutung von Fachwissenschaft und Fachdidaktik im Lehramtsstudium. Neben der Analyse mittels der dokumentarischen Methode wird explizit die Übereinstimmung von Lehrenden und Studierenden thematisiert und mit Blick auf deren Passung innerhalb des Lehramtsstudiums thematisiert. Im dritten Beitrag werden lehr-lern-theoretische Überzeugungen („transmissiv“ und „konstruktivistisch“) für das Fach Geschichte ausdifferenziert und mittels eines Fragebogens operationalisiert. Die Ergebnisse zeigen eine reliable Erfassung und eine mit den theoretischen Annahmen übereinstimmende empirische Struktur. Darüber hinaus werden Zusammenhänge zu Merkmalen des Lehramtsstudiums hergestellt. Im Gegensatz zu den anderen Beiträgen, werden im vierten Beitrag Lehrkräfte unterschiedlicher Fächer hinsichtlich ihrer lehr-lern-theoretischen Überzeugungen und ihrer Überzeugungen zur eigenen Lehrerrolle direkt miteinander verglichen. Die Ergebnisse zeigen dabei – auch unter Kontrolle verschiedener Hintergrundvariablen – systematische Niveauunterschiede zwischen Lehrkräften mit mathematisch-naturwissenschaftlichen und solchen mit sprach-gesellschaftswissenschaftlichen Fächern.

Neben den Diskussionen im Anschluss an die jeweiligen Beiträge findet eine übergreifende Betrachtung durch den Diskutanten statt. Damit soll eine interdisziplinäre Diskussion zu theoretischen Konzeptualisierungen und empirischen Operationalisierungen angestoßen werden, um weitere Forschungsdesiderata in den Blick zu nehmen und damit Forschungsfragen zu identifizieren. Hierzu zählen beispielsweise die Fragen, wie sich theoretische Konzeptionen von Überzeugungen durch eine fachspezifische Differenzierung verändern und inwiefern Vergleichbarkeit zwischen Domänen zielführend und möglich ist.

*Beiträge des Symposiums*

### Wissenschaftsverständnis im naturwissenschaftlichen Unterricht

**Tim Kramer, Markus Rehm, Tobias Dörfler, Lohse-Bossenz Hendrik**  
Pädagogische Hochschule Heidelberg, Deutschland

#### Theoretischer Hintergrund

Professionswissen zum Wissenschaftsverständnis wird in den naturwissenschaftlichen Fachdidaktiken unter der Bezeichnung Nature of Science (NOS) diskutiert und gelehrt (z. B. Lederman, 2007; Höttecke & Henke, 2010; McComas, 2014). Es beinhaltet ein reflektiertes Verständnis der Rolle der Naturwissenschaften im gesellschaftlichen Kontext sowie Eigenschaften naturwissenschaftlicher Erkenntnis. Dazu gehören u.a. folgende Facetten: (a) Wissen über die Differenzierung zwischen Beobachtung und Schlussfolgerung, (b) Wissen über die Veränderung wissenschaftlicher Erkenntnisse im Laufe der Zeit und (c) Wissen über die Bedeutung von Subjektivität und Kreativität für wissenschaftliche Erkenntnis.

Dieses fachdidaktische Wissen in professionelles Unterrichtshandeln umzusetzen, ist ein anspruchsvoller Schritt zur Umsetzung eines lernwirksamen Unterrichts. Aspekte von NOS werden in der Unterrichtspraxis jedoch häufig ignoriert (z. B. Bartos & Lederman, 2014). Angenommen wird daher, dass die Wirksamkeit des Unterrichts durch die explizite Ausbildung von Lehrkräften in diesem Bereich gesteigert werden kann.

Professionswissen von angehenden Lehrkräften kann durch Verzahnung von fachwissenschaftlichen, fachdidaktischen und pädagogischen Lerngelegenheiten gefördert werden (z. B. Korthagen, 2010). Grundsätzlich sollte eine solche Integration insbesondere in der Ausbildungsphase des Referendariats stattfinden. Untersuchungen, die Veränderungen zum fachdidaktischen Professionswissen zu NOS im Referendariat zeigen, stellen bisher ein Desiderat dar.

Um lernwirksames Professionswissen einer empirischen Überprüfbarkeit zugänglich zu machen, legen aktuelle Forschungsansätze nahe, kontextualisierte und situationsspezifische Testaufgaben zu verwenden (z.B. König, 2015). Hierfür bieten sich schriftliche Unterrichtsvignetten, d.h. kurze, authentische Beschreibungen von Unterrichtssituationen als Testaufgaben an, wodurch die Messung in einem möglichst realitätsnahen Bezugsrahmen stattfindet (vgl. Shavelson, 2013).

#### Fragestellung

Das Projekt untersucht longitudinal die Veränderung des Professionswissens angehender Lehrkräfte zu NOS. Untersucht werden insbesondere die oben genannten Facetten (a) bis (c), ihre Modellstruktur sowie Zusammenhänge zu Kovariaten. Die Untersuchung soll dazu dienen, Kompetenzentwicklungen angehender Lehrkräfte zu modellieren sowie Empfehlungen für die Lehrerbildung abzuleiten.

## Methode

Zur Erhebung des Professionswissens zu NOS wird ein vignettenbasierendes Instrumentarium (vgl. Stecher et al., 2006; Brovelli et al., 2014) erarbeitet. In neun Vignetten (drei pro NOS-Facette) werden den Proband\*innen Unterrichtssituationen präsentiert. Pro Vignette werden den Proband\*innen drei bis fünf Handlungsalternativen angeboten, deren situationspezifische Angemessenheit auf einer Skala von 1 bis 6 zu bewerten ist. Zur Testwertberechnung wurde eine Musterlösung entwickelt, an der sich Naturwissenschaftsdidaktiker\*innen (N = 8) beteiligten, die zu NOS forschen. Mit dem Modalwert dieser Musterlösung werden die Antworten der Proband\*innen verglichen und mit einem Partial-Credit-Verfahren kodiert. Die Itemscores der einzelnen Vignetten werden zu einem Vignettenscore aggregiert. Die Vignettenscores einer Facette werden summiert. Pro NOS-Facette ergibt sich ein theoretisch möglicher Maximalwert von sechs Punkten.

Das Instrument wird bei Sekundarstufe I-Lehrkräften der naturwissenschaftlichen Fächer (Biologie, Chemie, Physik) im Rahmen einer Längsschnitterhebung im Referendariat eingesetzt (N = 292, davon 72.1% weiblich). Die Referendar\*innen sind zwischen 22 und 49 Jahre alt (M = 26.64 Jahre, SD = 3.39 Jahre). Messzeitpunkt für die vorliegenden Ergebnisse ist der Referendariatsbeginn im Februar 2017. Eine zweite Messung erfolgt im Dezember 2017.

## Ergebnisse

Durchgeführt wurden die Analysen mit R und den Paketen lavaan (Rosseel, 2012) sowie psych (Revelle, 2012). Die Modellstruktur mit drei NOS-Facetten lässt sich durch eine konfirmatorische Faktorenanalyse bestätigen, (CFI = .934, RMSEA = .042, alternative Modellstrukturen werden berichtet).

Mittlere Ergebnisse erzielen die Referendar\*innen in den Facetten zum Wissen über die „Differenzierung zwischen Beobachtung und Schlussfolgerung“ (M = 2.95, SD = 2.33, Min = 0.75, Max = 5.17) und „Veränderung wissenschaftlicher Erkenntnis im Laufe der Zeit“ (M = 2.63, SD = 2.00, Min = 0.25, Max = 5.50). Geringere Ergebnisse erzielen die Referendar\*innen in der Facette zur „Bedeutung von Subjektivität und Kreativität“ (M = 1.79, SD = 1.08, Min = 0.00, Max = 5.50). Varianzanalytisch zeigen sich multivariat keine signifikanten Unterschiede zwischen Referendar\*innen in Bezug auf verschiedene naturwissenschaftliche Fächer, Hochschulstudienorte, Geschlecht und Muttersprache (alle  $p > .15$ ).

## **Überzeugungen von Lehrenden und Studierenden des Lehramts Deutsch zu Fachdidaktik und Fachwissenschaft**

**Felix Zühlsdorf, Iris Winkler**  
Friedrich-Schiller-Universität Jena

Lernprozesse in der universitären Lehrerbildung können in einem Angebot-Nutzungs-Modell gefasst werden (Kunter et al., 2011). Diese Perspektive nimmt sowohl die Angebote selbst wie auch die Voraussetzungen der Lehrenden und Studierenden in den Blick. Eine Gelingensbedingung für Lehrerbildung ist die „Passung“ der Lernangebote mit Blick auf die Studierendenüberzeugungen (Hascher, 2014). Überzeugungen der Hochschullehrenden werden dabei oft nicht mitgedacht. Inwieweit Überzeugungen von Lehrenden und Studierenden zusammenhängen, ist kaum geklärt.

Unter (Lehrer)Überzeugungen werden wertende und für richtig erachtete Vorstellungen verstanden, die sich auf Lernen, Lehren sowie die eigene Identität der Lehrenden beziehen. Diese können sowohl explizit als auch implizit sein (Reusser/Pauli, 2014). Fungieren Überzeugungen als Informationsfilter (Fives/Buehl, 2012), wird die Relevanz der Überzeugungen für das Gelingen von Lehrerbildung deutlich.

Der Beitrag präsentiert Befunde zu Überzeugungen von Lehrenden und Lernenden im Lehramtsstudium Deutsch zum Stellenwert und dem Verhältnis von Fachwissenschaft und Fachdidaktik. Anlass und Zielpunkt der Untersuchung ist die Gestaltung von Kooperationsseminaren aus Literaturwissenschaft und Fachdidaktik im Lehramtsstudium Deutsch an der Universität Jena. In diesen Lernangeboten zeigen Lehrende beider Domänen als Team auf, wie fachwissenschaftliche und fachdidaktische Perspektiven auf literarische Texte bezogen werden können. (Zühlsdorf et al. i. Dr.) Dies kann die Studierenden anregen, über die Bedeutung der einzelnen Domänen sowie deren Verhältnis nachzudenken. Aber auch für die Lehrenden dient die Kooperation als Motor für Reflexion über die Spezifika ihrer Fachdisziplin sowie über deren Relevanz und Aufgabe in der Lehrerbildung. Vor diesem Hintergrund stellt sich die Frage, ob und wie sich in diesen Überzeugungssystemen Gemeinsamkeiten bzw. Unterschiede finden lassen.

Qualitative Studien zu Überzeugungen von Lehramtsstudierenden beziehen sich oft auf Lehrerbildung insgesamt und sind fächerübergreifend (Schröter, 2014; Taibi 2012). Um zu möglichst konkreten Aussagen zu kommen, ist die vorgestellte Studie fachspezifisch auf das Lehramtsstudium Deutsch ausgerichtet. In der Deutschdidaktik werden Überzeugungen meist mit der dokumentarischen Methode untersucht (Ballis et al. 2014), um sowohl explizite als auch implizite Wissensbestände in den Blick zu bekommen.

Erste Befunde zu Lehrendenüberzeugungen zum Stellenwert und Verhältnis von Fachwissenschaft und Fachdidaktik im Lehramtsstudium Deutsch liegen aus Projekten der Qualitätsoffensive Lehrerbildung vor (Zühlsdorf et al. i. Dr.; Wieser & Winkler i. Dr.).

Der Beitrag fragt danach, welche Überzeugungen sich bei Studierenden zu Aufgabe und Relevanz von Fachwissenschaft und Fachdidaktik sowie deren Vernetzung in der Lehramtsausbildung Deutsch zeigen. Ergänzend wird gefragt, ob und wie sich Überzeugungen von Hochschullehrenden in der Lehrerbildung dazu in Beziehung setzen lassen.

Zur Erhebung der Studierendenüberzeugungen wurden problemzentrierte Interviews (Witzel/Reiter 2012) durchgeführt. Das Sample setzt sich aus dreizehn Studierenden zusammen, die im Rahmen zweier Kooperationsseminare aus Literaturwissenschaft und Fachdidaktik Deutsch für die Studie gewonnen werden konnten. Die Interviews werden mit Hilfe der dokumentarischen Methode (Nohl 2017) ausgewertet, um daraus sinngenetische Typen herauszuarbeiten. Die Ergebnisse werden vergleichend in Beziehung zu den Erfahrungen und Überzeugungen der Lehrenden (Zühlsdorf et al., i. Dr.; Wieser & Winkler, i. Dr.) gesetzt.

Bei der Auswertung der Studierendeninterviews kristallisieren sich drei Typen heraus: Der „praxisorientierte Typ“, der Fachwissenschaft und Fachdidaktik in der Pflicht sieht, praxisnah schulrelevante Inhalte zu vermitteln, der „wissenschaftsorientierte Typ“, der an wissenschaftlichen Erkenntnissen interessiert ist und der „Reflexionstyp“, der das erworbene Wissen im Studium eher als Folie für Reflexionen über Schulerfahrungen sieht. Letztere beiden Typen spiegeln sich auch in den Befunden zu den Lehrendenüberzeugungen. Es zeigt sich aber auch, dass dort die Überzeugungen zu Fachdidaktik und Fachwissenschaft stärker differenzieren und somit unterschiedliche Konstellationen von Überzeugungen in universitären Lernangeboten aufeinandertreffen können, die eine „Passung“ mit Blick auf die Studierenden kaum möglich erscheinen lassen.

## **Geschichtstheoretische und -didaktische Überzeugungen von Lehrpersonen - Konstruktvalidierung, Zusammenhänge und Kontextbezüge**

**Martin Nitsche**

Pädagogische Hochschule der Fachhochschule Nordwestschweiz

### Theoretischer Hintergrund

Individuelle Epistemologische Konstrukte werden domänenübergreifend entweder auf verschiedenen Stufen zwischen ‚naiv‘ und ‚reflektiert‘ unterschieden, in unterschiedliche Dimensionen etwa zur ‚Erkenntnisweise‘ und ‚Struktur des Wissens‘ differenziert (z.B. Hofer & Pintrich, 1997) oder hinsichtlich ihrer Kontexteinbindung erforscht (z.B. Hofer, 2016). Lehr- lerntheoretischen Überzeugungen von Lehrpersonen werden in quantitativer Hinsicht häufig als ‚transmissive‘ und ‚konstruktivistische‘ Beliefs konzipiert (z.B. Voss, Kleickmann, Kunter & Hachfeld, 2011), während qualitative Studien auch auf ‚sozialkonstruktivistische‘ Aspekte ausgerichtet sind (z.B. Fives, Lacatena & Gerard, 2015). Alger (2009) zeigt ausserdem, dass die Ausprägungen vom Kontext, in dem sich die Befragten bewegen (Hochschule vs. Schule), beeinflusst sind.

Für das Fach Geschichte liegen im internationalen Rahmen wenige Arbeiten zu epistemologischen Überzeugungen vor (z.B. VanSledright & Maggioni, 2016; Stoel, Logtenberg, Wansink, Huijgen, van Boxel, & van Drie, 2017). Wansink, Akkerman, Haenen, Vermunt & Wubbels (2017) haben zudem die kontextuelle Einbindung epistemologischer Konzepte bei Studierenden untersucht. Deutschsprachige Untersuchungen und internationale, die beide Gegenstände adressieren, fehlen.

Für die theoretische Modellierung der epistemologischen Konstrukte bin ich von Maggionis (2010) Stufenkonzept und Hofer & Pintrichs (1997) dimensionaler Perspektive ausgegangen. Geschichtstheoretisch (z.B. Lorenz, 1998) werden die drei Positionen ‚Positivismus‘, ‚Skeptizismus‘ und ‚Narrativer Konstruktivismus‘ unterschieden. Geschichtsdidaktisch (z.B. van Norden, 2012) können die lehr- lerntheoretischen Überzeugungen in die drei Überzeugungspositionen ‚Transmission‘, ‚Individueller Konstruktivismus‘ und ‚Sozialer Konstruktivismus‘ differenziert werden (Nitsche, 2016).

### Fragestellung

Ausgehend vom Forschungsstand wurde angenommen,

- a) dass sich für beide Konstrukte drei Positionen quantitativ valide und reliabel erheben lassen (H 1),
- b) dass die konstruktübergreifenden Positionen theoretisch plausibel (z.B. Positivismus und Transmission) zusammenhängen (H 2: Kriteriumsvalidität),
- c) dass kontextuelle Aspekte (z.B. Anzahl besuchter Lehrveranstaltungen) kontrolliert über individuelle Merkmale (z.B. Alter, Lehrerfahrung) die Ausprägung beider Konstrukte beeinflussen.

### Methode

Ausgehend von drei Pilotstudien wurde der „Epistemological Beliefs Questionnaire in History“ (EBHQ) und der „Teaching and Learning Beliefs Questionnaire in History“ (TLBHQ) mit vierstufiger Likert-Skala (stimmt nicht bis stimmt genau) entwickelt (Nitsche & Waldis, 2017). Nach Ausschluss nicht interpretierbarer Items basieren die Berechnungen zum ersten Konstrukt auf 12 von 22 Items, zum zweiten auf 13 von 18 Fragen. Ausserdem wurden individuelle und kontextuelle Merkmale erhoben. Das aktuelle Sample besteht aus 186 angehenden Geschichtslehrpersonen von sechs Hochschulen und 12 erfahren Lehrkräften der Deutschschweiz. Davon sind 102 Personen (54.8%) weiblich. Die Befragten waren im Mittel 28.79 Jahre alt (SD = 7.79) und hatten etwa 6 Lehrveranstaltungen in Geschichte (M = 6.38, SD = 8.37) sowie zwei Kurse in Geschichtsdidaktik (M = 2.39, SD = 1.32) besucht.

Zur Überprüfung von H 1 und 2 wurden Konfirmatorische Faktorenanalysen (CFA), zur Testung der dritten Hypothese Strukturregressionsmodelle (SR) geschätzt (Kline, 2016).

### Ergebnisse

Die CFA legen eine valide (EBHQ:  $\chi^2=51.49$ ,  $df=51$ ,  $p=.45$ ; CFI=1.00; TLI=1.00; RMSEA=.01; SRMR=.06, WRMR=.62; TLBHQ:  $\chi^2=70.00$ ,  $df=57$ ,  $p=.12$ ; CFI=.98; TLI=.97; RMSEA=.04; SRMR=.07, WRMR=.69) und reliable Erhebung (EBHQ:  $\omega = .64$ ,  $\alpha = .65-.73$ ; TLBHQ:  $\omega = .79$ ,  $\alpha = .64-.72$ ) der Konstrukte nahe (H 1). Positivistische geschichtstheoretische Beliefs hängen moderat mit transmissiven geschichtsdidaktischen Überzeugungen zusammen ( $r = .20$ ), während andere Korrelationen H 2 in Frage stellen (z.B. Skeptizismus und Transmission,  $r = .25$ ). Die Ergebnisse der SR-Modellschätzungen (EBHQ:  $\chi^2=252.72$ ,  $df=240$ ,  $p=.27$ ; CFI=.98; TLI=.97; RMSEA=.02; SRMR=.05; TLBHQ:  $\chi^2=324.77$ ,  $df=262$ ,  $p=.01$ ; CFI=.89; TLI=.87; RMSEA=.03; SRMR=.06) lassen sich so interpretieren, dass die Ausprägungen der geschichtstheoretischen Beliefs durch die Anzahl der besuchten Kurse in Geschichte beeinflusst werden (std.  $\beta = .16$  to  $.21$ ). Der Zuspuch zur transmissiven geschichtsdidaktischen Position wird durch die Anzahl der geschichtsdidaktischen Veranstaltungen prädiziert (std.  $\beta = .20$ ).

Letzteres Resultat entspricht nicht den geschichtsdidaktischen Zielstellungen der Geschichtslehrenausbildung (Fenn, 2013). Zur Erklärung sollten Lehrveranstaltungen an Deutschschweizer Hochschulen untersucht werden. Dabei wäre etwa nach dem Zusammenhang von Lehrveranstaltungszielen, methodischer Gestaltung und Überzeugungsausprägungen der Studierenden zu fragen.

## **Überzeugungen zum Lernen von Schülerinnen und Schülern und zur eigenen Rolle – ein Vergleich von angehenden Lehrkräften verschiedener Fächer**

**Hendrik Lohse-Bossenz<sup>1</sup>, Maria Schmidt<sup>2</sup>, Tina Seidel<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Heidelberg, <sup>2</sup>Technische Universität München

### Theoretischer Hintergrund

Baumert und Kunter (2006) formulieren in ihrem Modell neben dem Professionswissen, die motivationalen Orientierungen, die selbstregulativen Fähigkeiten sowie die professionellen Überzeugungen (Beliefs) als zentrale Aspekte professioneller Kompetenz von Lehrkräften. Gleichzeitig stellt sich auch die Frage, wie und wodurch Kompetenzmerkmale aufgebaut werden. Eine zentrale Rolle für die professionelle Entwicklung von Lehrkräften spielt die universitäre erste Phase der Lehrerbildung (Kunter, Kleickmann, Klusmann, & Richter, 2011). Einige Studien zeigen, dass dieser Professionalisierungsprozess u.a. durch die im Rahmen des Lehramtsstudiums studierten Unterrichtsfächer beeinflusst wird (Drake, Spillane, & Hufferd-Ackles, 2001; John & La Baggott Velle, 2004) und zu systematischen Unterschieden zwischen Lehrkräften verschiedener Fächer führt (Bernstein, 1972; Rich & Almozilino, 1999; Yaakobi & Sharan, 1985).

Ausgehend von systematischen Unterschieden zwischen Lehrkräften verschiedener Fächer hinsichtlich des bildungswissenschaftlichen Professionswissen (Lohse-Bossenz, Holzberger, & Kunter, 2016), wird in dieser Studie untersucht,

in welcher Hinsicht sich die Überzeugungen von angehenden Lehrkräften verschiedener Fächer (Naturwissenschaften vs. Sprach-/Gesellschaftswissenschaften) einerseits zum Lernen von Schülerinnen und Schülern (transmissive und konstruktivistische Überzeugungen) und andererseits zu ihrer Rolle im Lehr-Lern-Prozess (Wissensvermittler, Erzieher, Mediator und Helfer) unterscheiden. Durch die Sozialisation in verschiedenen Fächern und die damit verbundene Übernahme von grundlegenden Einstellungen und Überzeugungen werden substantielle Unterschiede zwischen den Fächergruppen erwartet.

#### Methode

Aus den Daten einer Vollerhebung von angehenden Lehrkräften zu Beginn des Vorbereitungsdienstes wurde eine Stichprobe von 492 Personen ausgewählt, von denen Aussagen bzgl. ihrer professionellen Überzeugungen vorlagen und die eindeutig einer der beiden Fächergruppen zugeordnet werden konnten (79 Naturwissenschaften und 413 Sprach-/Gesellschaftswissenschaften). Lehr-lern-theoretische Überzeugungen wurden über adaptierte Skalen von Voss, Kleickmann, Kunter, und Hachfeld (2011) und Überzeugungen zur eigenen Lehrerrolle über ein Verfahren von Kunter et al. (2010) erhoben (Cronbachs Alpha zwischen .74 und .85). Über konfirmatorische Mehrgruppenmodelle kann partiell skalare Messinvarianz gezeigt werden. Somit sind Gruppenvergleiche zulässig (Byrne, Shavelson, & Muthén, 1989; Vandenberg & Lance, 2000). Die Prüfung von Gruppenunterschieden erfolgte getrennt für lehr-lern-theoretische Überzeugungen und Überzeugungen zur Lehrerrolle mittels multivariater Kovarianzanalysen mit den Kovariaten Geschlecht, Muttersprache, Abiturnote und Lehramt.

#### Ergebnisse

Die Ergebnisse der multivariaten Kovarianzanalysen zeigen, dass sich angehende Lehrkräfte verschiedener Fächer unter Kontrolle verschiedener Kovariaten sowohl hinsichtlich lehr-lern-theoretischer Überzeugungen ( $F(2) = 7.62, p=.001$ ) als auch bzgl. ihrer Überzeugungen zur Lehrerrolle ( $F(4) = 6.14, p=.000$ ) unterscheiden. Spezifischer weisen mathematisch-naturwissenschaftliche Lehrkräfte höhere konstruktivistische ( $d=.65$ ) und niedrigere transmissive Überzeugungen ( $d=-.41$ ) zum Lernen auf, während sprach-gesellschaftswissenschaftliche Lehrkräfte sich in ihrer Lehrerrolle stärker als Mediator für Lernprozesse ( $d=-.45$ ), Erzieher ( $d=-.62$ ) und Helfer ( $d=-.29$ ) sehen.

#### Diskussion

Die Ergebnisse bestätigen die bisherigen Befunde zu fachspezifischen Unterschieden in den professionellen Überzeugungen von Lehrkräften. Im Gegensatz zu anderen Studien (z.B. Yaakobi & Sharan, 1985) zeigen jedoch mathematisch-naturwissenschaftliche Lehrkräfte stärkere konstruktivistische Überzeugungen, die sich jedoch nicht in Unterschieden hinsichtlich der Rolle der Lehrkraft als Wissensvermittler niederschlagen. Wie diese Unterschiede entstehen und ob sich diese über den Vorbereitungsdienst hinweg verändern ist Gegenstand der Diskussion.

## D14

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H117

### Chancen und Herausforderungen von Sekundäranalysen

Chair(s): **Jens Möller** (Christian-Albrechts-Universität zu Kiel)

DiskutantIn(nen): **Petra Stanat** (Institut für Qualitätsentwicklung im Bildungswesen)

In der empirischen Bildungsforschung ist in den letzten Jahren und Jahrzehnten eine Vielzahl von ausgesprochen umfangreichen Datensätzen aus large-scale assessments generiert worden. Deren Analysepotential wurde innerhalb der üblichen Projektlaufzeiten oft nicht voll ausgeschöpft. Mittlerweile gewinnt die Nutzung dieser bereits existierenden Datensätze zur Beantwortung methodischer und inhaltlicher Fragestellungen zunehmend an Bedeutung (Stanat, 2014).

Auch die beiden Längsschnittstudien LAU (Aspekte der Lernausgangslage und Lernentwicklung) und KESS (Kompetenzen und Einstellungen von Schülerinnen und Schülern), die seit 1995 von der Freien und Hansestadt Hamburg generiert wurden, bieten noch viele Möglichkeiten für Sekundäranalysen. Vor diesem Hintergrund gründete sich im Sommer 2012 das wissenschaftliche Konsortium MILES (Methodological Issues in Longitudinal Educational Studies), in dessen Rahmen eine Reihe einzelner Projekte auf Basis der gemeinsamen Datengrundlage entwickelt wurde. Das Symposium stellt Befunde aus diesen Projekten vor. Alle Beiträge haben ihren je eigenen inhaltlichen Fokus, gehen aber zugleich darauf ein, wie die sekundäranalytische Herangehensweise ihre Arbeiten beeinflusst. Mit dem Symposium soll aufgezeigt werden, welches thematische Spektrum sich auf Basis der selben Datengrundlage bearbeiten lässt, welche Chancen sich durch Sekundäranalysen bieten, gerade für Fragestellungen, die nicht gleich den Aufwand einer eigenen Studie rechtfertigen und welche Herausforderungen damit verbunden sind, z.B. in Hinblick auf die nicht immer optimale Verfügbarkeit von Variablen für die spezifischen Fragestellungen.

Der erste Beitrag untersucht, welche Faktoren einen schulischen Auslandsaufenthalt zwischen den Klassenstufen 10 und 13 vorhersagen. Dabei werden neben dem sozialen, kulturellen und ethischen Hintergrund der Schülerinnen und Schüler Unterschiede in schulischen Leistungsprofilen (Noten und erstmals Leistungstests), kognitiven Grundfähigkeiten und nicht-kognitiven Variablen (z.B. dem leistungsbezogenen Selbstkonzept) systematisch berücksichtigt. Die Analysen basieren auf Daten der Längsschnittstudie LAU.

Der zweite Beitrag beschäftigt sich mit der Frage, wie sich bilingualer Unterricht in der Sekundarstufe (5. – 10. Klasse) auf die Englischleistungen der Schülerinnen und Schüler auswirkt. Nach Kontrolle der Eingangsselektivität mit Hilfe eines Propensity Score Matchings wurden die Englischleistungen der bilingualen Schülerinnen und Schülern mit denen einer Kontrollgruppe im Längsschnitt verglichen.

Der dritte Beitrag untersucht den Einfluss von Nachhilfeunterricht auf die Schulleistungen. Dafür wurden Sekundäranalysen mit Hamburger LAU-Daten durchgeführt. Mittels Propensity Score Matching und OLS Regressionsanalysen wurde der Effekt von Nachhilfe auf die Noten sowie Testleistungen in den Nachhilfefächern Mathematik, Deutsch, Englisch und Latein analysiert. Unter Berücksichtigung von zentralen Kontrollvariablen, wie dem Vorwissen im jeweiligen Fach, können in der Sekundarstufe keine positiven Effekte von Nachhilfeunterricht beobachtet werden.

Der vierte Beitrag untersucht auf Basis der Daten der Hamburger Längsschnittstudie KESS die Entwicklung und die Konsistenz non-kognitiver Lernresultate auf der Schulebene. Die Ergebnisse der Untersuchung zeigen, dass die Einstellung zum Lesen im Verlauf der Sekundarstufe I im Mittel auf der Schulebene leicht absinkt, das Arbeitsverhalten im Fach Deutsch und das Leseselbstkonzept von SchülerInnen hingegen leicht zunimmt. Außerdem konnten keine Korrelationen zwischen den non-kognitiven Schuleffektivitätskriterien auf der Schulebene gefunden werden.

Abschließend werden die Arbeiten entsprechend der Leitidee des Symposiums in Hinblick auf die Chancen und Herausforderungen von Sekundäranalysen diskutiert.

*Beiträge des Symposiums*

### Ich bin dann mal weg! Zentrale Prädiktoren schulischer Auslandsaufenthalte

**Nicolas Hübner, Ulrich Trautwein, Benjamin Nagengast**

Eberhard Karls Universität Tübingen

Theoretischer Hintergrund. Hohe Kompetenzen in Englisch als „lingua franca“ der Gegenwart gehören zu den zentralen Zielkriterien der schulischen Ausbildung junger Menschen, insbesondere auch im Gymnasium (KMK, 2011). Zudem gelten längere Auslandsaufenthalte während der Schulzeit in einer globalisierten Ökonomie als eine wichtige Gelegenheit zum Erwerb transnationalen Humankapitals (Gerhards, Hans & Carlson, 2014). Angesichts dieser Bedeutung mag es verwundern, dass zu den Wahlmechanismen von Auslandsaufenthalten während der Schulzeit und insbesondere zu deren Effekten insgesamt nur wenig belastbare Zahlen und Befunde existieren. Eine Ausnahme stellt die Studie von Gerhards und Hans (2013) dar, die mit Daten des sozio-ökonomischen Panels (SOEP) zentrale Prädiktoren von Auslandsaufenthalten untersuchte. In ihrer Studie konnten die Autoren zeigen, dass die Wahrscheinlichkeit sich für ein Auslandsschuljahr zu entscheiden insbesondere mit dem sozioökonomischen und kulturellen Kapital der Eltern, dem besuchten Schultyp, der Schulnote in der ersten Fremdsprache sowie der Teilnahme an außerschulischen Aktivitäten zusammenhängt. Bei dem vorliegenden Beitrag handelt es sich um eine konzeptionelle Replikationsstudie, die an Ergebnissen der Untersuchung von Gerhards und Hans (2013) anknüpft und diese systematisch erweitert: Die Daten der Längsschnittstudie Aspekte der Lernausgangslage und der Lernentwicklung (LAU; Behörde für Schule und Berufsbildung, 2012; Lehmann, Peek, Gänsfuß, & Husfeldt, 2002) bieten dafür die besondere Möglichkeit, auch Unterschiede in schulischen Leistungsprofilen (Noten und erstmals Leistungstests), kognitiven Grundfähigkeiten und auf nicht-kognitiven Variablen (z.B. dem leistungsbezogenen Selbstkonzept) hinsichtlich ihrer Rolle bei der Entscheidung für ein schulisches Auslandsjahr zu berücksichtigen.

Fragestellung. Welche Schüler- und Schulmerkmale sagen einen Auslandsaufenthalt vorher? Im Mittelpunkt stehen hier besonders Merkmale des familiären Hintergrunds der Schülerinnen und Schüler sowie Leistungsmerkmale, die über standardisierte Leistungstests erfasst wurden. Die Fragestellung wurde vorab im Open Science Framework präregistriert.

Methode. Um die Fragestellung zu beantworten, wurden die Längsschnittdaten aus den LAU-Studien LAU 7 (1998), LAU 9 (2002) und LAU 13 (2005) verwendet. Berücksichtigt wurden Daten von insgesamt N = 3919 Schülerinnen und Schülern, davon berichteten n = 717, mindestens ein Schulhalbjahr zwischen den Klassenstufen 10 und 13 im Ausland verbracht zu haben. Von diesen gaben n = 529 Schülerinnen und Schüler an, eine Schule im englischsprachigen Ausland besucht zu haben. Der Auslandsaufenthalt wurde mittels multipler logistischer Regressionsmodelle in Mplus (7.3; Muthén & Muthén, 2012)

vorhergesagt. In Anlehnung an Gerhards und Hans (2013) wurden zunächst Prädiktoren zum sozialen, kulturellen und ethischen Hintergrund (HISEI, höchster Bildungsanschluss, häusliche Bildungsgüter und Buchbestand, Migrationshintergrund) sowie die Bildungsaspirationen der Eltern, Unterstützung der schulischen Interessen und Nachhilfe berücksichtigt (1). Anschließend wurde ein Modell (2) mit Prädiktoren zu Lernvoraussetzungen der Schülerinnen und Schüler (z.B. der standardisierten Leistung in Mathematik und Englisch und dem leistungsbezogenen Selbstkonzept) spezifiziert, bevor in einem dritten Modell (3) Prädiktoren aus den Bereichen Freizeitaktivitäten, schulisches und außerschulisches Engagement (z.B. Teilnahme an Schul-AGs oder Jugendgruppen) hinsichtlich ihrer Bedeutung für den Auslandsaufenthalt untersucht wurden. In einem vierten Modell (4) wurden schließlich auch Variablen des Schulkontextes (z.B. mittlerer HISEI der Schule, mittlere Mathematik- und Englischtestleistung der Schule) berücksichtigt. Abschließend wurden alle Prädiktoren in einem gemeinsamen Modell (5) getestet. Die hierarchische Struktur der Daten wurde durch die Schätzung robuster Standardfehler berücksichtigt. Fehlende Werte wurden bei den Analysen mit Full-Information-Maximum-Likelihood (FIML) behandelt.

Ergebnisse. Erste Ergebnisse legen nahe, dass insbesondere der soziale Hintergrund der Familie (HISEI und höchster Bildungsabschluss in der Familie) und die Leistungen der Schülerinnen und Schüler im Zusammenhang mit einem längeren Aufenthalt im Ausland stehen. Bezüglich der Leistungsmerkmale scheinen insbesondere die Englischnote und Klassenwiederholungen im Zusammenhang mit der Entscheidung für bzw. gegen einen Auslandsaufenthalt zu stehen. Ähnlich wie bereits bei Gerhards & Hans (2013) zeigte sich ebenfalls ein positiver Einfluss des außerschulischen Engagements. Zudem fanden sich positive Effekte der Schulkomposition (mittlerer HISEI auf Schulebene) auf den Auslandsaufenthalt. Die Ergebnisse sowie Möglichkeiten und Herausforderungen konzeptueller Replikationsstudien mit Sekundärdaten werden abschließend kritisch diskutiert.

### **Was bringt der bilinguale Sachfachunterricht in der Sekundarstufe wirklich? - Eine längsschnittliche Sekundäranalyse der KESS-Daten mit Propensity Score Matching**

**Rajka Matthießen, Sandra Preusler, Jens Möller**

Christian-Albrechts-Universität zu Kiel

Das Erlernen mindestens einer Fremdsprache ist im allgemeinen Schulsystem der Bundesrepublik Deutschland für alle Schülerinnen und Schüler (SuS) seit den 1970er-Jahren obligatorisch. Der Unterricht in der ersten Fremdsprache (meist Englisch) beginnt in der ersten oder dritten Klasse der Grundschule. Trotz der relativen Erfolge des Fremdsprachenunterrichts, die insbesondere für das Gymnasium belegt wurden (Köller, Knigge & Tesch, 2010; Leucht, Retelsdorf, Pant, Möller & Köller, 2015), liegen die Leistungen deutlich unter einem ersprachlichen Niveau. Daher sind bilinguale Unterrichtsangebote attraktiv, die versprechen, eine Zweitsprache (L2) auf einem Kompetenzniveau zu vermitteln, das dem der Erstsprache (L1) nahekommt, ohne dabei Nachteile in der L1 oder den Sachfächern in Kauf nehmen zu müssen. Von bilingualem Unterricht spricht die Kultusministerkonferenz (KMK) dann, wenn Teile des Sachfachunterrichts in der Fremdsprache unterrichtet werden (KMK, 2006). Auf europäischer Ebene spricht man von Content and Language Integrated Learning (CLIL) (Pérez-Canado, 2012). Der Sachfachunterricht beginnt im Rahmen des bilingualen Unterrichts meistens in der siebten Klasse. Als Vorbereitung auf den fremdsprachlichen Sachfachunterricht erhalten die SuS üblicherweise in der fünften und sechsten Klasse verstärkten Sprachunterricht in der Fremdsprache (1-2 Wochenstunden zusätzlich).

Ziel dieser Studie ist der Vergleich der Englischleistungen der SuS, die ein bilinguales deutsch-englisches Gymnasium besuchten mit SuS aus grundständigen Gymnasien ohne bilingualen Unterricht. Dabei wurde die Fragestellung untersucht, wie sich der deutsch-englische bilinguale Unterricht auf die Entwicklung der Englischleistungen der SuS von der fünften bis zur zehnten Klasse auswirkt.

Die Stichprobe entstand im Rahmen der Vollerhebung in Hamburg „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS) und bestand aus insgesamt 5850 SuS. 274 SuS besuchten ein deutsch-englisches bilinguales Gymnasium, 5576 SuS besuchten ein grundständiges Gymnasium ohne bilingualen Unterricht. Die Englischleistungen wurden mit einem C-Test erfasst. Dieser wird häufig in Schulleistungsstudien zur Messung der allgemeinen Sprachbeherrschung einer Fremdsprache eingesetzt. Der hier verwendete C-Test besteht aus drei Textabschnitten, bei dem lediglich der erste und der letzte Satz vollständig sind. Bei allen anderen Sätzen wurde jedes vierte Wort am Ende zur Hälfte entfernt und musste von den SuS ersetzt werden. Die Summenscores wurden zur längsschnittlichen Vergleichbarkeit nach dem einparametrischen Rasch-Modell skaliert. Als Schätzer für die Personenfähigkeiten wurden Weighted Likelihood Estimates (WLEs) bestimmt. Aufgrund der unterschiedlichen Stichprobengrößen und der Tatsache, dass sich beide Gruppen im Hinblick auf ihre Eingangsvoraussetzungen signifikant unterschieden, wurde ein Propensity Score Matching durchgeführt, um für die differenzielle Eingangsselektivität zu kontrollieren (Morgan & Winship, 2007; Baumert, Becker, Neumann & Nikolova, 2009; Becker, Lüdtke, Trautwein, Köller & Baumert, 2012). Mit diesen, hinsichtlich der Kovariaten parallelisierten Stichproben wurde im Anschluss der Einfluss des bilingualen Unterrichts auf die Englischleistung mittels einfacher Gruppenvergleiche (t-Tests) und einer Varianzanalyse mit Messwiederholung bestimmt.

Die Betrachtung der Mittelwerte zeigt, dass sich die beiden Gruppen zu Beginn der fünften Klasse hinsichtlich ihrer Englischleistungen nicht signifikant unterscheiden. Am Anfang der siebten Klasse sind die Englischleistungen der bilingualen SuS signifikant über denen der Kontrollgruppe. Dieser Leistungsunterschied bleibt bis nach der zehnten Klasse bestehen, ohne sich, wie erwartet, zu vergrößern. Die Ergebnisse der Varianzanalyse unterstützen die deskriptiven Befunde. Es zeigte sich ein signifikanter Leistungszuwachs für beide Gruppen von der fünften bis zur zehnten Klasse. Zusätzlich zeigte sich ein signifikanter Einfluss der Gruppenzugehörigkeit zugunsten der bilingualen SuS. Der Interaktionseffekt zwischen der Gruppenzugehörigkeit und der Länge des Englischunterrichts wurde nicht signifikant.

In Bezug auf unsere Fragestellung deuten die Ergebnisse darauf hin, dass lediglich der verstärkte Englischunterricht in der fünften und sechsten Klasse einen signifikanten Einfluss auf die Leistungsentwicklung der bilingual unterrichteten SuS hat. Der daran anschließende bilinguale Sachfachunterricht hat hingegen offenbar keinen zusätzlichen positiven Einfluss auf die Englischleistungen.

### **Wirksamkeit von Nachhilfeunterricht in der Sekundarstufe I – Sekundäranalysen mit den LAU-Daten**

**Melike Ömeroğulları, Karin Guill, Olaf Köller**

Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Nachhilfeunterricht ist in Deutschland weit verbreitet. Laut dem Bildungsbericht aus dem Jahr 2012 nehmen ca. 21 % aller 13- bis 18-jährigen Schülerinnen und Schüler außerschulische Nachhilfe in Anspruch (Autorengruppe Bildungsberichterstattung, 2012). Nachhilfe wird als eine wirksame Maßnahme zur Verbesserung der Schulleistung erachtet. Das lerntheoretische Modell von Carroll (1963) besagt, dass der Lernerfolg von der aktiv aufgebrauchten Lernzeit sowie benötigten Lernzeit beeinflusst wird.

Der Nachhilfeunterricht bietet Schülerinnen und Schülern zusätzliche Lernzeit und ermöglicht Unterrichtsinhalte zu wiederholen und Wissenslücken zu schließen. Demnach wird im Nachhilfeunterricht durch die effektive Nutzung der zusätzlichen Lernzeit zum schulischen Lernerfolg beigetragen. Darüber hinaus kann sich Nachhilfe positiv auf die Lernmotivation auswirken. Vor allem bei Schülerinnen und Schülern mit schwachen schulischen Leistungen kann sich die Leistungsrückmeldung anhand einer individuellen Bezugsnorm positiv auf das akademische Selbstkonzept und darüber vermittelt auf den Schulerfolg auswirken (Mischo & Kessel, 2005). Andererseits ist umstritten, ob durch Nachhilfe die aufgebrauchte Lernzeit sich in der Summe tatsächlich erhöht oder nur in den Nachhilfeunterricht verschoben wird. Zudem wird die Problematik diskutiert, ob Schülerinnen und Schüler, die Nachhilfe erhalten, möglicherweise im Schulunterricht weniger aufmerksam sind (Kenny & Faunce, 2004).

In Deutschland nehmen vor allem Schülerinnen und Schüler mit schwachen Leistungen Einzel- oder Gruppennachhilfeunterricht in Anspruch, um ihre schulischen Leistungen zu verbessern (Guill, 2012). Jedoch berichten Studien über die Wirksamkeit von Nachhilfe sehr unterschiedliche Befunde. Einige Studien berichten über moderat positive Effekte von Nachhilfe auf standardisierte Testleistungen (Ryu & Kang, 2013; Berberoğlu & Tansel, 2014). Demgegenüber stehen Befunde, die keine oder nur geringe Effekte von Nachhilfe auf schulische Leistungen zeigen (Guill & Bos, 2014; Luplow & Schneider, 2014). Aufgrund dieser heterogenen Befundlage ist es weiterhin notwendig mit Längsschnittdaten und unterschiedlichen Methoden der Fragen nach der Effektivität von Nachhilfeunterricht nachzugehen.

In dieser Arbeit werden Daten der Hamburger Längsschnittuntersuchung LAU (Aspekte der Lernausgangslage und der Lernentwicklung; Behörde für Schule und Berufsbildung, 2011) für Sekundäranalysen genutzt und die Auswirkung von Nachhilfeunterricht auf Schulnoten und Leistungstests untersucht (N = 11026). In LAU 9 wurden Schülerinnen und Schüler der 9. Klasse für verschiedene Schulfächer befragt, ob sie in den letzten zwei Jahren Nachhilfeunterricht erhalten haben. Ca. 22 % der Jugendlichen nahmen gelegentlich oder regelmäßig Nachhilfe in Anspruch. Um systematische Unterschiede zwischen Schülerinnen und Schülern mit und ohne Nachhilfe zu reduzieren, wurde das Matchingverfahren Propensity Score Matching (PSM; Stuart, 2010) durchgeführt (MatchIt). Dadurch konnte die Verteilung der Kovariaten dieser beiden Gruppen angeglichen werden. Zuvor wurden fehlende Werte mittels multiple Imputation in R ergänzt (mice, miceadds). Im Anschluss an das PSM wurden für die jeweiligen Nachhilfefächer Deutsch, Mathematik, Englisch und Latein getrennte OLS Regressionsanalysen durchgeführt. Als abhängige Variable dienen dabei die Schulnoten Ende der 8. Klasse sowie Ergebnisse der Leistungstests in den jeweiligen Nachhilfefächern, die zu Beginn der 9. Klasse durchgeführt wurden. Unter Kontrolle von Geschlecht, Migrationsstatus, Bildungshintergrund der Eltern, Schulform, leistungsbezogenem Selbstkonzept sowie Vorwissen im jeweiligen Fach, konnten keine positiven Effekte von Nachhilfeunterricht beobachtet werden. Mögliche Gründe für das Fehlen von Effekten werden diskutiert.

Die LAU-Daten für Sekundäranalysen zu nutzen, ermöglicht uns den Einfluss von Nachhilfe auf Schulleistungen (Noten und Leistungstests) für verschiedene Schulfächer mittels großer Stichproben zu untersuchen (z.B. Mathematiknachhilfe n = 1536). Jedoch müssen auch mit Einschränkungen umgegangen werden. Beispielsweise ermöglicht die Datenlage nicht zu berücksichtigen, wann genau bzw. für wie lange Nachhilfeunterricht in Anspruch genommen wurde. Bezogen auf die eigene Fragestellung werden Vor- und Nachteile von Sekundäranalysen aufgezeigt.

## **Die Entwicklung und die Konsistenz non-kognitiver Schuleffektivitätskriterien**

**Luisa Grützmaker, Johannes Hartig, Svenja Vieluf**  
Deutsches Institut für Internationale Pädagogische Forschung

### Theoretischer Hintergrund

Im vorliegenden Beitrag werden auf Basis der Daten, die in der Hamburger Längsschnittstudie KESS erhoben und im MILES-Verbundprojekt aufbereitet wurden, die Entwicklung und Konsistenz non-kognitiver Lernergebnisse auf der Schulebene untersucht. Die Bedeutung non-kognitiver Kompetenzen für den Schulerfolg ist unbestritten und wird zunehmend auf Individualebene erforscht (Birenbaum, 1996; Gutman & Schon, 2013). Die Schuleffektivitätsforschung berücksichtigte hingegen lange Zeit nur kognitive Kriterien (Scheerens & Bosker, 1997). Seidel (2008) schlägt statt dieser einseitigen Kriterienauswahl vor, zwischen kognitiven, metakognitiven, motivational-affektiven und behavioralen Schuleffektivitätskriterien zu unterscheiden.

Im vorliegenden Beitrag sollen ein behaviorales (Arbeitsverhalten im Deutschunterricht) und zwei motivational-affektive Kriterien (LeseEinstellung/ Leseselbstkonzept) betrachtet werden. Das Selbstkonzept beschreibt die durch vergangene Erfahrungen und Interaktionen mit der Umwelt herausgebildete (domänenspezifische) Selbstwahrnehmung einer Person (Shavelson, Hubner & Stanton, 1976). Die LeseEinstellung kann als Set von Gefühlen verstanden werden, welches zum Lesen oder zum Vermeiden von Leseaktivitäten bewegt (Alexander & Filler, 1976). Das Arbeitsverhalten kann als behaviorales Schullengagement aufgefasst werden und umfasst Aktivitäten wie das Hausaufgabenverhalten (Fredricks et al., 2004). Metakognitive Kriterien wie Lernstrategien können nicht berücksichtigt werden, da entsprechende Variablen in KESS nicht einheitlich erhoben wurden.

### Forschungsstand

Allgemein ist in der Adoleszenz ein Absinken der Schülermotivation zu verzeichnen (Fredricks & Eccles, 2002). In einer Längsschnittstudie konnten Van de Gaer und KollegInnen (2009) nachweisen, dass Schulen einen starken Effekt auf den Rückgang der Lernmotivation und des Selbstkonzepts haben. Wang und Eccles (2012) beobachteten dagegen zwar auch, dass das schulische Arbeitsverhalten abnahm, in ihrer Studie trug die Schulebene aber nicht zur Erklärung der Abnahme bei. Auch bezüglich der Zusammenhänge zwischen verschiedenen non-kognitiven Schuleffektivitätskriterien sind die Ergebnisse uneinheitlich: In der Studie von Van der Wal und KollegInnen (2007) konnte nach Kontrolle von Klassen- und Schülermerkmalen auf der Schulebene kein signifikanter Zusammenhang zwischen dem Selbstkonzept und dem politischen Engagement festgestellt werden. Weitere Untersuchungen zum Zusammenhang zwischen der Lernmotivation und dem schulischen Selbstkonzept bzw. dem Selbstkonzept, dem Arbeitsverhalten und der Einstellung untersuchten, konnten jedoch schwache Zusammenhänge auf der Schulebene eruieren (Mortimore et al., 1988; Van de Gear et al., 2009).

Bisherige Ergebnisse deuten darauf hin, dass Schulen einen – wenn auch kleinen – Effekt auf die Entwicklung non-kognitiver Kriterien haben, dass verschiedene Kriterien aber nur schwach korreliert sind. Daher werden folgende Hypothesen aufgestellt:

1. Auf Schulebene ist ein Absinken des Leseselbstkonzepts, der LeseEinstellung und des Arbeitsverhaltens im Deutschunterricht im Verlauf der Adoleszenz zu verzeichnen.
2. Schulen haben konsistente Effekte auf die Veränderung im Leseselbstkonzept, der LeseEinstellung und des Arbeitsverhaltens im Deutschunterricht.

### Methoden

Die Stichprobe umfasst 13057 SchülerInnen in 172 Schulen und drei Messzeitpunkte (7., 8., 10/11. Klasse). Die erste Hypothese wird mit drei linearen Wachstumskurvenmodellen, die entsprechend der Datenstruktur auf zwei Ebenen modelliert wurden, getestet (Mplus). Für die zweite Hypothese werden die drei latenten Wachstumskurvenmodelle in ein Modell integriert und der Zusammenhang der Slope-Faktoren ermittelt.

#### Ergebnisse

Die Ergebnisse zeigen, dass die erste Hypothese nur teilweise bestätigt werden kann. Auf der Schulebene verschlechtert sich zwar die Einstellung zum Lesen im Mittel ganz leicht (MSL-Slop =  $-.083$ ), aber das Arbeitsverhalten im Deutschunterricht (MAV-Slop =  $.101$ ) und das Leseselbstkonzept (MSL-Slop =  $.102$ ) steigen im Mittel.

Außerdem zeigte sich, dass die Höhe der Zuwächse auf der Schulebene variiert, aber keine Korrelationen zwischen den non-kognitiven Schuleffektivitätskriterien auf der Schulebene gefunden wurden. Somit kann die zweite Hypothese nicht bestätigt werden.

#### Chancen und Herausforderungen von Sekundäranalysen

Das Arbeiten mit Sekundärdaten kann, wie im Fall der KESS- und LAU-Daten, für ForscherInnen eine gute Möglichkeit sein, das Potential der vorhandenen Daten und der zusätzlichen Informationen, die durch die Verbindung beider Leistungsstudien entstanden ist, auszuschöpfen. Dennoch gibt es immer wieder Herausforderungen zu bewältigen. Eine der Herausforderungen in Bezug auf diesen Beitrag besteht darin, dass aufgrund der vorgegebenen Datenlage nicht alle Schuleffektivitätskriterien berücksichtigt werden können.



## D15

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H116

### **Forschungsimpulse für den Mathematikunterricht: Erziehungswissenschaft und Fachdidaktik im Dialog**

Chair(s): **Tina Hascher** (Universität Bern, Schweiz), **Esther Brunner** (Pädagogische Hochschule Thurgau)

DiskutantIn(nen): **Stefan Ufer** (Ludwig-Maximilians-Universität München)

Die Professionalität von Lehrpersonen ist entscheidend für die Qualität ihres Unterrichts. Wesentliche Impulsgeberin für die (Weiter-)Entwicklung der Kompetenzen von Lehrpersonen ist zum einen die empirische Bildungsforschung und zum anderen eine empirisch orientierte Fachdidaktik. Im Symposium soll diskutiert werden, wie sich deren Erkenntnisse für den Mathematikunterricht und die Aus- und Weiterbildung von Lehrpersonen nutzen lassen. Exemplarisch wird dies an drei empirisch geprüften Interventionsprojekten ausgeleuchtet, die auf verschiedenen Schulstufen aktuelle Herausforderungen des Mathematikunterrichts aufgreifen, nämlich kognitive Aktivierung, Lernmotivation und die Bedeutung neuer Medien:

- Die Förderung des mathematischen Argumentierens im Vorschulalter
- Motivationsförderung für das Mathematiklernen auf der Sekundarstufe I
- Die Rolle neuer Medien für das Mathematiklernen in der Sekundarstufe I am Beispiel des Bruchbegriffs

Die drei Beiträge stellen die (vermuteten bzw. geprüften) Wirkungsmechanismen der durchgeführten Intervention vor, berichten Ergebnisse und diskutieren die Bedeutsamkeit der Befunde für die Professionalisierung von (angehenden) Lehrpersonen und deren Unterrichtsqualität aus erziehungswissenschaftlicher und fachdidaktischer Sicht.

Im Symposium geht es um einen interdisziplinären Dialog und die Frage, welchen Beitrag ein solcher für die Weiterentwicklung der Kompetenzen von Lehrpersonen leistet. Deshalb werden in einer abschliessenden Diskussion Möglichkeiten und Potenziale einer verstärkten interdisziplinären Zusammenarbeit im Hinblick auf eine weitere Professionalisierung von Mathematiklehrpersonen beleuchtet.

*Beiträge des Symposiums*

### **Die Förderung mathematischen Argumentierens im Kindergarten: Erste Ergebnisse zur professionellen Entwicklung der Kindergärtnerinnen und ihres Unterrichts**

**Esther Brunner, Jonas Lampart, Janine Rüdüsüli**  
Pädagogische Hochschule Thurgau

Theoretischer Hintergrund

Mathematisches Argumentieren gewinnt im Zusammenhang mit den neuen Bildungsstandards (D-EDK, 2014; KMK, 2005) an Bedeutung und gilt als wichtige und gleichermaßen anspruchsvolle zu erwerbende prozessbezogene Kompetenz im Mathematikunterricht. Im Sinne eines kumulativen Wissensaufbaus und kohärenten Mathematiklernens über alle Bildungsstufen hinweg (Heinze & Grüssing, 2009) ist es notwendig, diese Kompetenz nicht erst in der Sekundarstufe, sondern auch bereits in der frühen Altersstufe, z.B. im Kindergarten zu fördern. Frühe mathematische Bildung ist aber in vielen Ländern noch ein vergleichsweise junges Gebiet, sowohl in der Forschung wie in der Praxis (Gasteiger, 2015). Dies gilt im Besonderen für frühes mathematisches Argumentieren, das im Vergleich zu frühen numerischen Kompetenzen oder zum späteren Beweisen in der Sekundarstufe noch wenig erforscht ist. Bislang fehlen aber nicht nur altersspezifische theoretische Konzeptualisierungen für das frühe Argumentieren, sondern ebenso sehr erprobte Praxiskonzepte, Lernumgebungen und Materialien zur Förderung von mathematischem Argumentieren im Kindergarten. Lehrpersonen der frühen Bildungsstufen sehen sich deshalb einerseits mit dem Anspruch der Bildungsstandards konfrontiert und vermissen andererseits konkrete Anregungen und Materialien für die Praxisumsetzungen. Es ergibt deshalb Sinn, sich dieser Problemlage im Rahmen einer explorativen Pilotstudie anzunähern und erst in einem zweiten Schritt dadurch gewonnene Erkenntnisse in eine breitere Studie mit Untersuchungs- und Kontrollgruppendesign einfließen zu lassen.

Fragestellung und Methoden

Ziel des von der Stiftung „cogito foundation“ geförderten Pilotprojekts „IvMAiK“ (Intervention zum mathematischen Argumentieren im Kindergarten) war es, im Rahmen einer Intervention mit zehn Kindergärtnerinnen während eines Schuljahres Lernumgebungen zum mathematischen Argumentieren im Kindergarten zu entwickeln, erproben und evaluieren. Dazu erhielten die zehn Lehrpersonen verteilt über ein Schuljahr eine mehrteilige Weiterbildung im Umfang von 20 Stunden, bestehend aus: 1) einem Einführungsworkshop, 2) vier Follow-up-Halbtagen mit 3) jeweils anschliessendem fachspezifisch-pädagogischem Coaching sowie 4) einem abschliessenden Workshop.

Aus Forschungssicht interessieren sowohl die professionelle Entwicklung der Kindergärtnerinnen hinsichtlich ihrer Argumentationspraxis wie auch die differenzielle Leistungsentwicklung der Kindergartenkinder im Verlauf eines Schuljahres (Pre-/Posttest). Der vorliegende Beitrag konzentriert sich auf die professionelle Entwicklung der Kindergärtnerinnen während eines Schuljahres. Diese wird anhand von verschiedenen Datenquellen nachgezeichnet: 1) Pre-/Postbefragung zu mathematikbezogenen Überzeugungen, zu Fachwissen und fachdidaktischem Wissen zum mathematischen Argumentieren sowie zur selbst berichteten Argumentationspraxis im Kindergarten, 2) den ausführlichen Unterrichtsplanungen für vier Einheiten zum Argumentieren mit der eigenen Kindergartenklasse während des Schuljahres sowie 3) zwei videografierten Unterrichtseinheiten zum Argumentieren (Einheit 1 und Einheit 4).

Ergebnisse

Erste Ergebnisse zeichnen ein positives Bild bezüglich der professionellen Entwicklung der beteiligten Lehrpersonen. Sowohl relevante Überzeugungen zu Mathematik und Mathematiklernen wie auch das Fach- und fachdidaktische Wissen zum mathematischen Argumentieren ließen sich über das Schuljahr positiv verändern.

In fast allen Planungsunterlagen der Lehrpersonen konnten explizite Bezüge zur Förderung von Teilkompetenzen des mathematischen Argumentierens sowie zu spezifischen didaktischen Settings festgestellt werden.

Die Auswertungen der Planungsunterlagen und der Videodaten sowie mögliche Zusammenhänge zwischen Fachwissen und Überzeugungen einerseits und realisierter Unterrichtspraxis in den geplanten (und durchgeführten) sowie in den videografierten Sequenzen sind derzeit in Gange.

Die vorgestellte Pilotstudie ermöglicht auch eine Diskussion zum aktiven Einbezug von Praktikerinnen in Forschungsprojekte und zeigt darüber hinaus die eminente Wichtigkeit soliden Praxiswissens der Forschenden im Zusammenhang mit Interventionsstudien auf. Diskutiert werden sollen ferner mögliche Konsequenzen des Pilotprojekts für eine Hauptstudie.

### **Die Bedeutsamkeit der Lehrperson im Hinblick auf die Förderung der Motivation im Mathematikunterricht auf der Sekundarstufe I**

**Nicole Moser, Claudia Brandenberger, Tina Hascher**  
Universität Bern

#### Theoretischer Hintergrund

Ein wiederholt bestätigter Befund in der empirischen Bildungsforschung ist die Abnahme der Lernmotivation und -freude über die Sekundarschulzeit, insbesondere im Fach Mathematik (z.B. Fend, 1997). Erklärungsansätze für diesen Rückgang bieten z.B. die Stage-Environment-Fit-Theorie (Eccles et al., 1993) oder die Theorie der Selbstbestimmung (Deci & Ryan, 2002). Beide Theorien verbindet die Annahme, dass der Rückgang mit einer mangelnden Passung zwischen den psychologischen Bedürfnissen der Jugendlichen (das Erleben von Autonomie, Kompetenz und sozialer Einbindung) und den Anforderungen der Schule zusammenhängt. Entsprechend besteht ein zentraler Ansatzpunkt für Interventionen darin, diese Passung zu verbessern. Konkret bedeutet dies, Schüler/innen im kompetenten Umgang mit den schulischen Anforderungen und den damit einhergehenden emotionalen und motivationalen Herausforderungen zu fördern sowie Lehrpersonen als Schnittstelle zu ebendieser Passung im Hinblick auf autonomie- und kompetenzerwerbunterstützenden Unterricht zu sensibilisieren. Bisher fehlen jedoch Studien, die überprüfen, ob und wie dem o.g. Rückgang entgegengewirkt werden kann und welche Rolle die Lehrperson dabei spielt.

#### Fragestellung und Methode

An dieser Stelle setzt das vom Schweizerischen Nationalfond geförderte Projekt „EMo-Math“ (Nr. 100019\_156710) an, dessen Ziel es ist, verschiedene Möglichkeiten der Prävention und Intervention anhand eines quasi-experimentellen Prä-Post-Follow up-Designs zu überprüfen. Die Intervention ist über einen Zeitraum von 2 Schuljahren (7. Schuljahr 2015/2016 und 8. Schuljahr 2016/2017) angelegt und richtet sich an Schüler/innen, die in Mathematik das niedrigste Anforderungsniveau (Realstufe in der Schweiz; entspricht Hauptschulstufe in Deutschland) besuchen, da diese eine Risikogruppe im Hinblick auf den Motivationsrückgang darstellen. Überprüft werden die Effekte von zwei Interventionsprogrammen, die sich entweder an Schüler/innen (Schüler/innen-Treatment: 8 Klassen) oder an Schüler/innen und ihre Mathematiklehrpersonen (kombiniertes Treatment: 8 Klassen) richten, im Vergleich zur Entwicklung von Lernfreude und Motivation in der Kontrollgruppe (6 Klassen). Weiter sollen differentielle Effekte für diverse Subgruppen (z.B. Geschlecht) untersucht werden. Die beiden Interventionsprogramme (Schüler/innen-Treatment und kombiniertes Treatment) verfolgen das Ziel, die positive Entwicklung von Emotionen und Motivation der Schüler/innen in Mathematik zu unterstützen. Die Schüler/innenworkshops finden im regulären Klassenkontext statt und beinhalten z.B. Motivierungs- und Lernstrategien, Lernziele, Nützlichkeit von Mathematik und der Umgang mit eigenen Emotionen. Inhalte der Workshops für Lehrpersonen sind z.B. Formen der Motivationsunterstützung im Rahmen der Selbstbestimmungstheorie (Autonomieerleben, Kompetenzunterstützung, soziale Einbindung) und der Umgang mit Schüleremotionen. Die Wirksamkeit der Treatments wird anhand von schriftlichen Befragungen der Schüler/innen und Lehrpersonen sowie Leistungstests der Schüler/innen überprüft. Als domänenspezifische, abhängige Variablen dienen die motivationalen Orientierungen, die Emotionen und Mathematikleistungen der Schüler/innen sowie die selbst- und fremdeingeschätzte Unterrichtsqualität der Lehrpersonen.

#### Ergebnisse

Die Ergebnisse des ersten Interventionsjahres (7. Schuljahr, 2015/2016) zeigen, dass Schüler/innen der kombinierten Interventionsgruppe (Schüler/innen und Lehrpersonen) eine positive Entwicklung in der selbstbestimmten Lernmotivation, der Lernfreude und dem Selbstkonzept aufweisen, während es in der Kontrollgruppe keine bedeutsamen Veränderungen von Prä- zu Posterhebung gibt. Entgegen den Erwartungen zeigt sich in der Schüler/innen-Interventionsgruppe eine negative Entwicklung in der selbstbestimmten Motivation und in der Lernfreude. Die Ergebnisse betonen die Bedeutsamkeit der Lehrperson im Hinblick auf die Motivation der Schüler/innen. Im Vortrag werden deshalb mit besonderem Blick auf die Inhalte der Interventionsworkshops für die Lehrpersonen Impulse für die Professionalisierung von (angehenden) Lehrkräften diskutiert.

### **Interaktive Lehrbücher im Bruchrechnenunterricht der Sekundarstufe I**

**Stefan Hoch, Frank Reinhold, Bernhard Werner, Jürgen Richter-Gebert, Kristina Reiss**  
Technische Universität München

#### Theoretischer Hintergrund

Die Digitalisierung der Gesellschaft schreitet immer weiter voran und hält auch in den Klassenzimmern mehr und mehr Einzug. Diese Neuerungen stellen nicht nur die Lehrkräfte vor Herausforderungen: Auch der Schritt von konventionellen zu digitalen Schulbüchern erscheint vor diesem Hintergrund an der Zeit. Allerdings stellt sich die Frage, wie ein gelungenes digitales Schulbuch aussehen kann, da die Möglichkeiten weit über eine bloße Digitalisierung von Text hinausgehen.

Im Projekt ALICE:Bruchrechnen wird ein interaktives Lehrbuch zur Bruchrechnung (Hoch, Reinhold, Werner, Reiss, & Richter-Gebert, 2016) entwickelt und evaluiert. Dabei werden Hands-on-Aktivitäten, die sich beim Erlernen von mathematischen Konzepten als hilfreich erwiesen haben (vgl. Burns & Silbey, 2000) als interaktive Elemente auf das iPad übertragen. Insbesondere bietet das eBook kontinuierlich die Möglichkeit, verschiedene mathematische Darstellungen aktiv zu verändern.

Einige Forschungsergebnisse zur Nutzung von Tablets im Mathematikunterricht sind diesbezüglich vielversprechend. So berichtet Blair (2014) einen höheren Lernerfolg beim Erwerb des Zahlkonzepts in der Grundschule. Auch Black, Segal, Vitale und Fadjó (2012) verzeichnen am iPad einen höheren Lernerfolg beim Erlernen der Additionsregel im Vergleich zur Nutzung derselben Lernumgebung an einem herkömmlichen Computer. Sie führen die Unterschiede auf die natürlichere Eingabemethode auf dem iPad mittels Fingergesten zurück, wobei sie die Embodied Cognition Theory (Gangopadhyay & Kiverstein, 2009) als Basis betrachten.

Darüber hinaus können Schülerinnen und Schüler bei den im eBook angebotenen interaktiven Aufgaben von weiteren, für computerbasierte Lernumgebungen typische Eigenschaften profitieren: Die einzelnen Komponenten des digitalen Lehrbuchs sind so programmiert, dass sie umgehend kurz und präzise formuliertes Feedback geben, das auf den Schülerfehler eingeht. Zudem passen sich die Aufgaben in ihrer Schwierigkeit adaptiv anhand der Lösungsrate der vorherigen Aufgaben an die Nutzerin oder den Nutzer an. Wo dies sinnvoll ist, bietet das eBook zusätzlich abgestufte Lösungshilfen, die zu jeder Zeit der Aufgabenbearbeitung selbstreguliert aufgerufen werden können.

Dadurch bietet sich den Lehrkräften die Möglichkeit, sich vom System unterstützen zu lassen und Freiräume im Unterricht zu schaffen, um Individuen besser zu betreuen.

#### Fragestellung

Welchen Einfluss hat ein interaktives Lehrbuch auf den Kompetenzerwerb der Schülerinnen und Schüler? Ist die Arbeit mit dem Tablet PC geeignet, Lernprozesse zu verbessern?

#### Methode

Eine Evaluation erfolgte im Rahmen einer vierwöchigen Interventionsstudie, an der 19 sechste Klassen von Gymnasien (474 Schülerinnen und Schüler, davon 214 Mädchen) teilnahmen. Eine iPad-Gruppe (6 Klassen, 155 Schülerinnen und Schüler) arbeitete mit dem iBook auf iPads, eine Arbeitsbuchgruppe (7 Klassen, 182 Schülerinnen und Schüler) mit einer gedruckten Version des iBooks als Arbeitsbuch (Experimentalgruppen). Zur Kontrolle von Effekten durch das im Projekt entwickelte Unterrichtsmaterial arbeitete eine Kontrollgruppe (6 Klassen, 137 Schülerinnen und Schüler) mit konventionellen Schulbüchern.

In allen Gruppen wurden ein 15-minütiger Pretest vor der Intervention und ein 55-minütiger Posttest nach der Intervention durchgeführt, der die Bruchrechenkompetenz der Schülerinnen und Schüler in den Bereichen Teil vom Ganzen, Erweitern und Kürzen und Größenvergleich maß.

#### Ergebnisse

Beide Experimentalgruppen erzielten im Posttest signifikant bessere Ergebnisse als die Kontrollgruppe. Jedoch konnten keine signifikanten Unterschiede zwischen den Experimentalgruppen festgestellt werden (siehe auch Reinhold, Hoch, Werner, Richter-Gebert, & Reiss, 2017).

Diese Ergebnisse zeigen, dass Schülerinnen und Schüler von einer Lernumgebung, in der beständig zwischen verschiedenen mathematischen Darstellungen gewechselt werden muss, profitieren können. Dies bestätigt bisherige empirische Befunde zur Didaktik der Bruchrechnung und hebt die Relevanz von Repräsentationswechseln für den Bruchrechenunterricht hervor. Allerdings konnte der erwartete zusätzliche Effekt durch die Verwendung interaktiver und adaptiver Inhalte auf dem iPad nicht nachgewiesen werden.

Die Studie zeigt, wie wichtig die Qualität des Lehr-/Lernmaterials für den Unterricht sein kann, und insbesondere, dass wissenschaftlich fundiertes Material im Unterricht zur Bruchrechnung zu einem höheren Lernerfolg führen kann.

## D16

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H115

### Interkulturelles Lernen in Schule und Unterricht

Chair(s): **Kerstin Göbel** (Universität Duisburg-Essen, Deutschland)

DiskutantIn(nen): **Andrea Hänni-Hoti** (PH Luzern)

Prozesse der Globalisierung in Wirtschaft und Gesellschaft sowie aktuelle und vergangene Zuwanderung erfordern ein Konzept für den Umgang mit kultureller und sprachlicher Heterogenität. Die Schule ist, im Sinne eines Akkulturations- aber auch Sozialisationsagenten in multikulturellen Gesellschaften, ein wichtiges Bindeglied zwischen der aufnehmenden Gesellschaft und den zugewanderten Menschen (Schönpflug, 2008; Vedder & Horenczyk, 2006). Schülerinnen und Schüler, aber auch Lehrerinnen und Lehrer, müssen lernen, mit Menschen unterschiedlicher kultureller Prägungen umzugehen (Banks, 2017; Deardorff, 2004; Göbel & Buchwald, 2017; van Oudenhoven & Benet-Martínez, 2015). Ein konstruktiver Umgang mit sprachlicher und kultureller Heterogenität in der Schule ist daher erforderlich, stellt jedoch auch eine Herausforderung dar. Das vorliegende Symposium wendet sich dem Thema des interkulturellen Lernens in Schule und Unterricht zu und stellt Forschungsansätze zu interkulturellen Lernprozessen vor, die aus dem Kontext der fachdidaktischen Forschung des Fremdsprachenunterrichts, aus der psychologischen Perspektive der Kontaktforschung sowie aus der Moralentwicklungsforschung relevante Aspekte im Lernprozess beleuchten. Im Zentrum der Beiträge stehen Prozessvariablen wie das interkulturelle Klassenklima sowie verschiedene didaktische Ansätze zur Förderung interkultureller Kompetenzen. Dabei interessiert einerseits, wie die Lernbedingungen von den Lernenden selbst wahrgenommen werden und andererseits, wie die didaktischen und klimatischen Bedingungen auf die Entwicklung interkultureller Kompetenzen wirken. Durch die Zusammenstellung von Arbeiten, die sich mit dem Thema der Entwicklung interkultureller Kompetenzen aus unterschiedlichen theoretischen Positionierungen und mit Blick auf unterschiedliche Zielgruppen auseinandersetzen, soll die Identifikation von Gelingensbedingungen interkultureller Lehr-Lern Prozesse in der Schule und im Unterricht ermöglicht werden.

*Beiträge des Symposiums*

### Das Potenzial multikultureller Schulen nutzen: Interkulturelles Klassenklima und interkulturelle Kompetenz

**Miriam Schwarzenthal<sup>1</sup>, Linda Juang<sup>1</sup>, Maja K. Schachner<sup>1</sup>, Fons J. R. van de Vijver<sup>2</sup>**  
<sup>1</sup>Universität Potsdam, <sup>2</sup>Tilburg University

#### Theoretischer Hintergrund

In zunehmend multikulturellen Gesellschaften müssen Schülerinnen und Schüler früh lernen, verschiedene Weltanschauungen zu verstehen und mit Menschen unterschiedlichster kultureller Prägungen zu kooperieren (van Oudenhoven & Benet-Martínez, 2015). Lehrerinnen und Lehrer, Bildungsforscherinnen und Bildungsforscher, Arbeitgeberinnen und Arbeitgeber sowie Politikerinnen und Politiker fordern daher, dass Schülerinnen und Schüler interkulturelle Kompetenz entwickeln sollen (Deardorff, 2004; European Parliament and European Council, 2006; Kultusministerkonferenz, 2013). Interkulturell kompetente Menschen sind motiviert, mit Menschen unterschiedlichster kultureller Hintergründe zu interagieren, sie besitzen Wissen über Normen und Konventionen in verschiedenen Kulturen, sind sich kultureller Einflüsse auf ihr eigenes Denken und Verhalten sowie das Denken und Verhalten anderer bewusst, und verhalten sich angemessen in unterschiedlichen kulturellen Situationen (Van Dyne et al., 2012).

Theorien aus der Sozial- und Entwicklungspsychologie sowie der interkulturellen Psychologie suggerieren, dass interkulturelle Kompetenz erworben werden kann durch interkulturellen Kontakt und die Auseinandersetzung mit verschiedenen kulturellen Perspektiven (Allport, 1954; Piaget, 1977; Tadmor & Tetlock, 2006). Die positiven Lernergebnisse dererlei Erfahrungen wurden bereits in verschiedenen Zielgruppen dokumentiert, so z. B. bei Austauschschülerinnen und Austauschschülern, Studentinnen und Studenten an internationalen Universitäten und Schülerinnen und Schülern, die an interkulturellen Trainings teilnehmen (Busse, Riedesel, & Krause, 2017; Geeraert & Demes, 2012; Gurin, Dey, Hurtado, & Gurin, 2002). Das Potenzial multikultureller Schulen in einer Einwanderungsgesellschaft jedoch wurde in der öffentlichen Debatte und in der psychologischen Forschung bislang vernachlässigt.

Multikulturelle Schulen können, unter bestimmten Bedingungen, eine natürliche Arena für den Erwerb interkultureller Fähigkeiten sein (Sam & Oppedal, 2003). Hierfür ist das Klassenklima von großer Bedeutung (Nesdale & Lawson, 2011; Tropp et al., 2016), vor allem in Bezug darauf, inwiefern es Kontakt und Kooperation zwischen Schülerinnen und Schülern verschiedener Herkunftskulturen begünstigt und in Bezug darauf, ob der Fokus auf Gemeinsamkeiten oder auf eine aktive Auseinandersetzung mit kultureller Variation gelegt wird.

#### Fragestellung

In dieser Studie untersuchen wir, wie drei Arten des interkulturellen Klassenklimas (Kontakt und Kooperation, Farbenblindheit und interkulturelles Lernen) mit vier Subfacetten der interkulturellen Kompetenz (motivational, kognitiv, metakognitiv und behavioral) von Schülerinnen und Schülern zusammenhängen. Da sich Beziehungen zwischen interkulturellen Erfahrungen und interkultureller Kompetenz zwischen Mitgliedern der kulturellen Mehrheit und Minderheit unterscheiden (Denson & Zhang, 2010; Phinney, Jacoby, & Silva, 2007), untersuchen wir zudem, ob die Zusammenhänge zwischen interkulturellem Klassenklima und interkultureller Kompetenz bei Lernenden mit und ohne Migrationshintergrund unterschiedlich sind.

#### Methode

Die Stichprobe bestand aus 1335 Jugendlichen in Deutschland (48% weiblich, MAlter = 14.69 Jahre), von denen 52% einen Migrationshintergrund hatten (d. h. min. ein Elternteil wurde im Ausland geboren). Im Rahmen einer querschnittlichen Fragebogenstudie wurden das interkulturelle Klassenklima (Schachner, Schwarzenthal, Moffitt, Civitillo, & Juang, 2017) und die motivationale, kognitive, metakognitive und behaviorale interkulturelle Kompetenz der Schülerinnen und Schüler (adaptiert von Van Dyne et al., 2012) erfasst.

#### Ergebnisse

Mehrebenenanalysen in Mplus (Muthén & Muthén, 1998-2011) ergaben, dass ein Klima, welches Kontakt und Kooperation zwischen Schülerinnen und Schülern verschiedener kultureller Hintergründe fördert, positiv mit motivationaler, metakognitiver, und behavioraler interkultureller Kompetenz der Lernenden zusammenhängt. Ein stärkerer Fokus auf Gemeinsamkeiten zwischen Menschen verschiedener kultureller Hintergründe (Farbenblindheit) hing positiv mit der motivationalen und

behavioralen interkulturellen Kompetenz der Schülerinnen und Schüler zusammen. Ein stärkerer Fokus auf interkulturelles Lernen hingegen hing mit einer höheren kognitiven, metakognitiven und behavioralen interkulturellen Kompetenz der Schülerinnen und Schüler zusammen. Diese Assoziationen variierten teils abhängig vom Migrationshintergrund.

Wir schließen daraus, dass Kontakt und Kooperation, Farbenblindheit und interkulturelles Lernen in der Schule positive, aber unterschiedliche Beiträge zur Förderung der interkulturellen Kompetenz der Schülerinnen und Schüler leisten. Wenn kulturelle Gemeinsamkeiten in einer Klasse betont werden, sind Schülerinnen und Schüler eher motiviert, mit Menschen unterschiedlicher kultureller Hintergründe zu interagieren. Wenn kulturelle Variationen anerkannt und wertgeschätzt werden, zeigen Schülerinnen und Schüler erhöhtes kulturelles Wissen und Bewusstsein. Alle Ansätze hängen mit interkulturell kompetentem Verhalten zusammen.

### **Interkulturelles Lernen im Englischunterricht interessant und verständlich gestalten: Was zeichnet Unterricht aus, der den Lernenden gefällt?**

**Svenja Vieluf<sup>1</sup>, Kerstin Göbel<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main, <sup>2</sup>Universität Duisburg-Essen

Theoretischer Hintergrund und Fragestellungen

Interkulturelles Lernen stellt ein wichtiges Ziel schulischer Bildung dar, wobei dem Fremdsprachenunterricht eine besondere Bedeutung für die Vermittlung entsprechender Kompetenzen zukommt (z. B. KMK, 2013). Es stellt sich jedoch die Frage, wie dieses Ziel am besten erreicht werden kann. In der Diskussion um didaktische Ansätze zur Vermittlung interkultureller Kompetenzen werden verschiedene Strategien zur Reflexion des Kulturbegriffs sowie die Reflexion kultureller Einflüsse auf das menschliche Wahrnehmen, Denken und Handeln in den Blick genommen (z. B., Blell & Doff, 2014; Busse & Krause, 2016; Byram, 1997; Byram, Gribkova, & Starkey, 2002; Göbel, Lewandowska, & Diehr, 2017; Lange & Paige, 2003; Paige, Jorstad, Siaya, Klein, Colby, 2003). Der Bezugnahme auf die eigene lebensweltliche Erfahrung der Lernenden wird dabei besondere Bedeutung beigemessen (Göbel, 2007). Bislang seltener untersucht wurde jedoch, wie diese didaktischen Ansätze bei den Lernenden selbst ankommen, also ob die Lernenden einen Unterricht, in dem interkulturelle Themen reflektiert und diskutiert werden, interessant und verständlich finden und welche Rolle es dabei spielt, wie gut die Lehrenden an die Erfahrungen und an das Vorwissen der Lernenden anknüpfen, sowie ob sie Themen wie Toleranz, Respekt und Vorurteile im Unterricht thematisieren. Empirische Forschung zur Unterrichtsqualität deutet ferner daraufhin, dass neben der Fachdidaktik, also neben dem Bereitstellen anregender fachlicher Inhalte und Lerngelegenheiten, auch allgemein-didaktische Merkmale des Unterrichts von zentraler Bedeutung sind für erfolgreiches Lernen sind. Hierbei sind vor allem die Gestaltung und Gewährleistung eines geordneten Unterrichtsverlaufs sowie eines respektvollen und unterstützenden sozialen Umgangs während des Unterrichts (z. B., Göbel & Hesse, 2008; Klieme, Pauli, & Reusser, 2009; Pianta & Hamre, 2009; Rakoczy et al., 2007; Seidel & Shavelson, 2007). In der vorliegenden Studie möchten wir daher der Frage nachgehen, ob und wie interkulturell spezifische und allgemeine Dimensionen der Unterrichtsqualität im interkulturellen Englischunterricht realisiert werden und ob die Einschätzung von externen Beurteilern mit der Einschätzung der Lernenden im Hinblick in Zusammenhang steht.

Methode

Für die Beantwortung unserer Forschungsfragen ziehen wir Daten aus der deutschen Schulleistungsstudie DESI (für weitere Informationen siehe Beck & Klieme, 2007; DESI-Konsortium, 2008) heran, im Rahmen derer der interkulturelle Englischunterricht in 93 Klassen videografiert worden ist. Externe Beobachtende beurteilten die Qualität dieser Unterrichtsstunden auf zwölf Skalen, die sich drei Dimensionen der Unterrichtsqualität, nämlich Qualität der interkulturellen Didaktik, Klassenführung, Motivationsunterstützung, zuordnen lassen. Zudem liegen Bewertungen der Interessanztheit und Verständlichkeit der Unterrichtsstunde durch 2.047 Lernende vor, welche direkt im Anschluss an die Stunde mithilfe von Fragebögen erhoben worden waren. Mithilfe von Mehrebenenanalysen haben wir auf der Basis dieser Daten Zusammenhänge zwischen den Ratings externer Beurteilender und der Unterrichtsbeurteilung durch die Lernenden auf der Klassenebene berechnet.

Ergebnisse

In den Analysen zeigten sich auf der Klassenebene mehrere signifikante Zusammenhänge mittlerer Stärke zwischen externen Ratings und den Unterrichtsbeurteilungen durch die Lernenden. Lernende bewerteten Unterrichtsstunden im Mittel als interessanter und verständlicher, in denen Lehrende stärker an ihre Erfahrungen anknüpfen sowie eine positive Schülerorientierung und eine positive Fehlerkultur realisierten. Unterrichtsstunden, in denen die externen Beobachtenden einen respektvolleren, wärmeren und humorvolleren Umgangston sowie eine effektivere Störungsprävention und -intervention feststellten, wurden von den Lernenden ebenfalls als interessanter bewertet. Schließlich scheint das Thematisieren von Vorurteilen zu einer höheren Verständlichkeit des Unterrichts beigetragen zu haben. Es fanden sich dagegen keine Zusammenhänge der Unterrichtsbeurteilungen durch die Lernenden mit der beobachteten Qualität der interkulturellen Didaktik: Unterrichtsstunden, in denen interkulturelle Themen reflektiert und diskutiert wurden und in denen Toleranz, Respekt und Vorurteile thematisiert wurden, wurden weder als interessanter noch als weniger interessant bewertet als andere Unterrichtsstunden. Insgesamt zeigen diese Ergebnisse, wie wichtig es ist, bei der Umsetzung interkultureller Themen im Englischunterricht schülerorientiert vorzugehen, auf einen positiven und respektvollen Umgangston in der Klasse zu achten und Strategien für eine effektive Klassenführung einzusetzen.

### **Interkulturelles Lernen mit VaKE (Values and Knowledge Education)**

**Jean-Luc Patry<sup>1</sup>, Alfred Weinberger<sup>2</sup>, Sieglinde Weyringer<sup>1</sup>**

<sup>1</sup>Universität Salzburg, <sup>2</sup>Private Pädagogische Hochschule der Diözese Linz

Theoretischer Hintergrund

VaKE (Values and Knowledge Education) ist ein konstruktivistisches (i.S. von Piaget, 1976; s. Patry, 2016) Interventionsverfahren, mit dem Werterziehung (basierend auf der Theorie von Kohlberg, 1981, 1984) und Wissenserwerb (Conceptual Change, Vosniadou, 2008) in einem sozialen Austausch i.S. des sozialen Konstruktivismus (Vygotsky, 1978) miteinander kombiniert werden: Ausgehend von der wertbezogenen Diskussion eines speziell konstruierten moralischen Dilemmas in der Tradition von Kohlberg und Blatt (1975) stellen sich die Teilnehmerinnen und Teilnehmer inhaltliche Fragen (Wissen), die sie durch eigene Recherchen zu beantworten suchen (Inquiry-Based Learning, Reitinger et al., 2016). Darauf aufbauend wird die Wertediskussion als Collaborative Learning (Webb, 2009) auf einem höheren Niveau weitergeführt, was zu neuen Fragen und Recherchen führt, etc. Zusätzlich werden Rollenspiele mit Perspektivenwechsel und andere Techniken

eingesetzt. Viele Untersuchungen haben gezeigt, dass VaKE die Wertentwicklung fördert, wobei die Teilnehmerinnen und Teilnehmer jeweils hoch motiviert sind (z.B. Patry et al., 2013).

Im Migrationskontext sind Kurse zur Integration und aktiver Teilnahme an der Gesellschaft so-wohl für Migrantinnen und Migranten wie auch für die gastgebende Gesellschaft von besonderer Bedeutung (Council of Europe, 2006). Demokratie aus interkultureller Perspektive zielt auf ein Gleichgewicht zwischen Einheit und Diversität ab. Demokratie wird hier als Prozess interpretiert, in dem alle Betroffenen auf verantwortungsbewusster Grundlage wichtige Aspekte ihrer eigenen Kultur beibehalten können, dabei verschiedene Standpunkte abwägen, auf Grund von Fakten und bewusst artikulierten Werten argumentieren, jeweils multiple Perspektiven berücksichtigen und kreativ nach Problemlösungen suchen (Banks, 2004, 2017; Weyringer, 2008). Die Hypothese ist, dass VaKE diesbezüglich einen substanziellen Beitrag leisten kann, da genau dies umgesetzt wird (Patry et al., 2017).

#### Fragestellung

Untersucht werden soll erstens, ob VaKE – mit jeweils zielgruppenspezifischen Anpassungen – mit Migrantinnen und Migranten eingesetzt werden kann. Zweitens soll zumindest ansatzweise erfasst werden, ob VaKE ein geeignetes didaktisches Instrument zur Demokratieverziehung für diese Personen sein kann.

#### Methode

In verschiedenen Settings mit Migrantinnen und Migranten wird VaKE eingesetzt (Interventionsstudien). Das oben dargestellte Demokratieverständnis kann dabei nicht auch nur annähernd umfassend erhoben werden, einerseits weil es an sich schon komplex ist, andererseits weil es die praktischen Umstände jeweils nicht zulassen. Insbesondere ist es aus praktischen Gründen nicht möglich, Kontrollgruppen zu untersuchen. Entsprechend werden die Auswirkungen von VaKE je nach Setting und Möglichkeiten unterschiedlich erhoben: mit Fragebögen, teilnehmenden Beobachtungen, Prozessberichten u. a. m.

#### Ergebnisse

In einem Heim für minderjährige unbegleitete Flüchtlinge diskutierten zehn männliche Muslime ein Dilemma, in dem es um Alleinherrschaft oder Demokratie als Regierungsform in einem fernen, neu zu besiedelnden Planeten ging. Obwohl die Teilnehmenden keine Erfahrungen mit Demokratie hatten, entschieden sie sich für diese Regierungsform und gingen sogar so weit, selber eine Wahl durchzuführen (Patry et al., 2016). Parallel dazu wurde in einem Pariser Vorort ein VaKE-Programm für fremdsprachige Schülerinnen und Schüler, die kürzlich in Frankreich angekommen sind, implementiert (Brossard Boerhaug, in Vorb.). Diese Kinder erhalten einen kombinierten Sprach- und VaKE-Kurs zusätzlich zum regulären Unterricht. Die Erfahrungen, erfasst durch Fragebögen, Audio-Aufnahmen und Plenums-Diskussionen, sind durchwegs positiv: Die Teilnehmerinnen und Teilnehmer sind hoch motiviert und die Lehrerinnen und Lehrer berichten Erfolge sowohl im Wissens- als auch im Wertebereich.

#### Diskussion

Insgesamt sind die Erfahrungen ermutigend: Sie zeigen, dass VaKE beim jeweiligen Zielpublikum sehr gut ankommt, dass die Teilnehmerinnen und Teilnehmer engagiert und motiviert mitmachen und zu demokratietheoretisch relevanten Ergebnissen kommen. Darauf aufbauend werden weitere Projekte zum Einsatz von VaKE vorgestellt:

- Eine französische Agentur, die sich mit der Entwicklung und Vermittlung interkultureller Kommunikationskompetenz für Lehrerinnen und Lehrer bzw. anderen Mediatoren befasst, setzt VaKE zur Schulung von Personen ein, die insbesondere Funktionen im Nahen und Mittleren Osten und in Nordafrika haben (SIETAR, 2017).
- Ein auf VaKE basierendes Interventionsprogramm für weibliche Flüchtlinge in Österreich wird demnächst implementiert. Im Symposium werden die ersten Erfahrungen berichtet.

### **Förderung interkultureller Kompetenz im Fremdsprachenunterricht durch problembasierte Lerneinheiten**

**Vera Busse**

Carl von Ossietzky Universität Oldenburg

#### Theoretischer Hintergrund

Die Förderung interkultureller Kompetenz ist eine Querschnittsaufgabe der Schule, zu der der Fremdsprachenunterricht in besonderer Weise beiträgt: So entwickeln Lernende Offenheit gegenüber anderen Deutungsmustern und lernen, kommunikative Situationen aus verschiedenen Perspektiven zu betrachten, ohne diese vorschnell zu bewerten. Sie werden ferner mit unterschiedlichen Strategien und Fertigkeiten zur Bewältigung interkultureller Interaktionssituationen vertraut gemacht, damit sie in der Lage sind, situationsangemessen und adressatengerecht zu handeln (Kultusministerkonferenz 2012). Befunde aus der DESI-Studie (Deutsch-Englisch-Schülerleistungen-International) weisen jedoch darauf hin, dass die interkulturelle Kompetenz deutscher Schülerin-nen und Schüler nicht dem in den Curricula erwarteten Niveau entspricht (Göbel 2007; Hesse & Göbel 2007). Problembasierte Lerneinheiten in Form von Critical Incidents (Göbel & Helmke, 2010) sowie die Reflexion eigener interkultureller Erfahrungen und der Einsatz von affektiv-erfahrungsorientierten Methoden wie Rollen- und Simulationsspiele (z.B. Barrett et al., 2014) erscheinen zur Förderung entsprechender Fähigkeiten vielversprechend, bislang gibt es jedoch kaum Interventionsstudien, die die Wirksamkeit dieser Ansätze empirisch prüfen. Im Vortrag werden eigene Interventionsstudien vorgestellt, in denen die Wirksamkeit problembasierter Lerneinheiten und affektiv-erfahrungsorientierter Vermittlungsmethoden zur Förderung interkultureller Kompetenz untersucht wurden.

#### Fragestellung

Im Zentrum der Studie standen folgende Fragestellungen: 1) Inwieweit fördert eine eigens entwickelte problembasierte Unterrichtseinheit interkulturelle Lernprozesse und inwieweit lässt sich die Effektivität der Lerneinheit durch den Einsatz affektiv-erfahrungsorientierter Methoden steigern? 2) Inwieweit beeinflusst die Unterrichtssprache den Lernerfolg? 3) Welche Reflexionsprozesse werden durch die Unterrichtsansätze angeregt und wie nachhaltig sind diese? 4) Inwieweit halten die Schülerinnen und Schüler die unterschiedlichen Unterrichtsansätze für relevant und motivierend?

#### Methode

Im Fokus des Vortrags steht eine quasi-experimentelle Interventionsstudie mit einem Prä-Post-Design, in denen die Faktoren Unterrichtsansatz (eher affektiv-erfahrungsorientiert/eher analytisch) und Unterrichtssprache (Deutsch/Englisch) verglichen wurden. Zudem wurden einen Monat nach der Intervention Interviews mit Lernenden durchgeführt, um Einblicke in weiterführende Reflexionsprozesse zu erlangen. Die Lerneinheit umfasste 6 x 45 Minuten und wurde im Rahmen des regulären Schulunterrichts in den Sekundarstufen I und II in neun Schulklassen (n = 143) durchgeführt. Als Kontrollgruppe dienten fünf Schulklassen (n = 66), die nicht an der Intervention teilnahmen. Durch den Einsatz von Critical Incidents wurden die

Lernenden zur Perspektivenübernahme und zur Reflexion ihrer eigenen kognitiven und affektiven Prozesse (wie Kategorisierung, Beurteilung und Einfühlung) angeregt. Im Rahmen der Unterrichtseinheit entwickelten die Schülerinnen und Schüler außerdem Strategien zum Umgang mit den dargestellten Situationen. Dazu wurde im affektiv-erfahrungsorientierten Ansatz Rollenspiele und ein Simulationsspiel eingesetzt und im analytischen Ansatz Filmmaterial genutzt. Mit Hilfe von Tests, Fragebögen, Kurzinterviews und Tonbandmitschnitten wurden kognitive, motivationale und einstellungsbezogene Effekte der Intervention untersucht.

#### Ergebnisse

Der Vortrag stellt relevante Ergebnisse vor und vergleicht diese mit Ergebnissen aus der ersten Erhebungsphase (Busse & Krause 2015). Insgesamt legen die Ergebnisse der Studie nahe, dass sich problembasierte Lerneinheiten zur Förderung interkultureller Kompetenz in der Schule eignen. Der affektiv-erfahrungsorientierte Ansatz ist offenbar besonders vielversprechend in Bezug auf die Motivation und die Qualität der Reflexionsprozesse, was die Wichtigkeit der Anregung von Perspektivwechsel und Empathie (vgl. Beelmann & Heinemann 2014) für die interkulturelle Erziehung unterstreicht.

## D17

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H114

### Interne und externe Bedingungsfaktoren für die Wirksamkeit von Trainings zur Förderung des selbstregulierten Lernens

Chair(s): **Ferdinand Stebner** (Ruhr-Universität Bochum, Deutschland), **Wirth Joachim** (Ruhr-Universität Bochum)

DiskutantIn(nen): **Elmar Souvignier** (Universität Münster)

Selbstreguliertes Lernen (SRL) ist eine anspruchsvolle Lernform, bei der Lernende eigenständig Aufgaben analysieren und sich selbst Ziele setzen, aufgabenspezifische Strategien kennen, beherrschen und einsetzen, ihren Lernprozess reflektieren, ihre Fortschritte bewerten und gegebenenfalls korrigierende Maßnahmen zur Optimierung des eigenen Lernens durchführen müssen (Spörer & Brunstein, 2006). Die Forschung hat wiederholt die Bedeutsamkeit von Trainings zur Förderung der Kompetenz des SRL betont (z.B. Artelt et al., 2003). Auch wenn inzwischen eine Vielzahl an Trainings existieren (vgl. Dignath & Büttner, 2008), die zumeist in laborartigen Settings evaluiert wurden, so mangelt es an Forschung zu der Frage, welche Bedingungsfaktoren die Wirksamkeit von Trainings zur Förderung des SRL unter Realbedingungen beeinflussen (Masui & DeCorte, 2005).

Dieses Symposium widmet sich mit vier Beiträgen und einer Diskussion der Frage, welche Bedingungsfaktoren die Wirksamkeit von Trainings zur Förderung des SRL unter Realbedingungen beeinflussen. Das damit verbundene Ziel ist es, Wissenschaftlerinnen und Wissenschaftlern sowie Lehrenden professionelles Handeln mit erfolgreich evaluierten Trainings zu ermöglichen.

Die vier Beiträge unterscheiden sich bezüglich dreier Dimensionen: Trainingsformat (Präsenz vs. webbasiert), Nähe der Bedingungsfaktoren zum zu fördernden Konstrukt SRL (proximal bis distal) und Lernende (Schülerinnen und Schüler vs. Studierende).

Der erste Beitrag fokussiert die Bedeutung des sozioökonomischen Status (als distale Bedingung) für die Wirksamkeit eines Präsenztrainings für Schülerinnen und Schüler der fünften Jahrgangsstufe zum selbstregulierten Lernen durch Experimentieren und aus Sachtexten. Die Ergebnisse zeigen keinen Matthäus-Effekt, sondern je nach Operationalisierung des sozioökonomischen Status sogar eher einen kompensierenden Effekt, nach dem vor allem diejenigen Schülerinnen und Schüler der fünften Jahrgangsstufe von dem Training profitierten, die einen niedrigen sozioökonomischen Status aufwiesen. Im zweiten Beitrag wurden Studierende randomisiert entweder einem Präsenzseminar oder einem webbasierten Kurs mit je fünf inhaltlichen Lektionen zugewiesen. Der Fokus dieser Studie lag u.a. auf Persönlichkeitsvariablen wie Verträglichkeit und Gewissenhaftigkeit (eher proximale Bedingung). Die Ergebnisse zeigen, dass Studierende mit hoher Verträglichkeit stärker von Präsenzveranstaltungen profitierten, wohingegen Studierende mit hoher Gewissenhaftigkeit im webbasierten Kurs erfolgreicher dazulernten. Im dritten Beitrag wird ein siebenwöchiges Präsenztraining mit dem Fokus auf das SRL-Ausgangsniveau (proximale Bedingung) der Studierenden untersucht. Mit Hilfe von Prozessanalysen auf Basis ausgefüllter Lerntagebücher konnte gezeigt werden, dass entgegen der Annahmen kein kompensierender Effekt auftrat, sondern alle Studierenden von dem Training unabhängig von deren SRL-Ausgangsniveau erfolgreich dazulernten. Im vierten Beitrag werden die Bedeutung des SRL-Ausgangsniveaus (proximale Bedingung) und der Erwartungen der Studierenden (eher proximale Bedingung) für das freiwillige Wahrnehmen und die Wirksamkeit eines webbasierten Trainings mit sechs Lektionen untersucht. Die Ergebnisse zeigen hier, dass diejenigen Studierenden freiwillig mehr Lektionen bearbeiten, die zu Beginn eine geringe SRL-Ausgangskompetenz aufwiesen und positive Erwartungen an das webbasierte Training hatten. Zudem zeigt sich, dass die Studierenden, die an dem Training teilnahmen, einen höheren Zuwachs an SRL-Kompetenz zurückmelden und bessere Noten in der Abschlussklausur erzielten als Studierende der Kontrollgruppe, die an keinem Training teilnahmen.

Die abschließende Diskussion wird die Relevanz verschiedener Bedingungsfaktoren für die vier in der Bildungsforschung bewährten Trainingsprogramme herausarbeiten und einen Ausblick geben in Bezug auf die Bedeutung dieser Erkenntnisse für zukünftige Forschungsvorhaben und das professionelle Handeln in der Bildungspraxis.

#### Beiträge des Symposiums

### Kompensierender Effekt oder Matthäus-Effekt? Eine Studie zum Zusammenhang zwischen dem sozioökonomischen Status und der Wirkung eines Trainings zum selbstregulierten Lernen

**Ferdinand Stebner<sup>1</sup>, Corinna Schuster<sup>1</sup>, Yves Karlen<sup>2</sup>, Theresa Dicke<sup>3</sup>, Detlev Leutner<sup>4</sup>, Joachim Wirth<sup>1</sup>**

<sup>1</sup>Ruhr-Universität Bochum, <sup>2</sup>PH FHNW, <sup>3</sup>Australian Catholic University, <sup>4</sup>Universität Duisburg-Essen

Der Zusammenhang zwischen dem sozioökonomischen Status (SES) und dem akademischen Erfolg von Schülerinnen und Schüler (SuS) konnte in vielen Studien empirisch bestätigt werden (z.B. Hansen & Munk, 2012). Geringe elterliche Unterstützung auf Seiten der SuS mit niedrigem SES kann als ein Grund angesehen werden, warum diese SuS durchschnittlich schlechtere Schulleistungen erreichen (Daniel et al., 2016). Um die fehlende elterliche Unterstützung zu kompensieren, sollten SuS ihren Lernprozess selbstständig kontrollieren und regulieren können. Daraus resultiert unter anderem die Forderung nach Trainings zur Förderung der Kompetenz des selbstregulierten Lernens (SRL) (z.B. Artelt, 2003; Benner, Boyle & Sadler, 2016). Hier stellt sich aber die Frage, ob diese Trainings für SuS mit niedrigem SES wirksam sind oder womöglich zu einer Verstärkung des Zusammenhangs im Sinne eines Matthäus-Effekts (Renkl, 1996) führen. Für diese Studie stellt sich also die Frage, ob SuS mit hohem SES stärker von einem Training zur Förderung des SRL profitieren als SuS mit niedrigem SES.

Um dies zu untersuchen, wurden im Rahmen des Schulentwicklungsprojekts „Ganz In“ 118 SuS (M=10.5 Jahre; 50% weiblich) der fünften Jahrgangsstufe über ein halbes Schuljahr von Gymnasiallehrerinnen und -lehrern, die im Vorhinein von wissenschaftlichem Personal fortgebildet wurden, trainiert. Das erfolgreich evaluierte Training hat einen naturwissenschaftlichen Schwerpunkt und trainiert das SRL durch Experimentieren und aus Sachtexten (Stebner et al., 2015). Vor und nach dem Training wurde das Strategiewissen der SuS mit Hilfe des Multistrategietests (MST; Stebner et al., 2015; Cronbachs Alpha = .70) erhoben. Um den SES zu bestimmen, wurde sich zweier Methoden aus PISA 2003 bedient: einerseits wurden die SuS mit einem einzelnen Item nach der Anzahl an Büchern im Haushalt befragt. Zusätzlich wurden die SuS mit einer Skala bestehend aus zehn Items nach dem Besitz bestimmter Wohlstands- und Kulturgüter befragt (Cronbachs Alpha = .64). In beiden Fällen wurde ein Mediansplit durchgeführt, um die SuS zu gruppieren (hoher und niedriger SES). Zusätzlich zum SES wurden auch das Geschlecht, die Schulnoten und die kognitiven Fähigkeiten (KFT; Heller et al., 1985) in die Berechnungen einbezogen. Es wurden Varianzanalysen mit wiederholtem Messzeitpunkt und der MST-Leistung als abhängige und dem SES, dem Geschlecht, den Schulnoten und den kognitiven Fähigkeiten als unabhängige Variable gerechnet.



Die empirischen Ergebnisse zeigen keinen verstärkenden Matthäus-Effekt, sind jedoch abhängig von der Operationalisierung des SES. Wenn SES mit dem Besitz an Büchern operationalisiert wird, zeigt sich ein Haupteffekt des Trainings ( $F(1,116)=56.77, p=.00, \eta^2=.33$ ), jedoch keine Interaktion mit SES ( $F(1,116)<1, \eta^2<.01$ ). Operationalisiert man SES mit der Skala aus Wohlstands- und Kulturgütern, zeigt sich sowohl ein Haupteffekt des Trainings ( $F(1,116)=61.55, p=.000, \eta^2=.35$ ) als auch eine signifikante Interaktion mit SES ( $F(1,116)=4.58, p=.03, \eta^2=.04$ ) im Sinne eines kompensierenden Effekts. Demnach profitieren die SuS mit niedrigem SES stärker von dem Training als die SuS mit hohem SES. Darüber hinaus zeigen die Ergebnisse einen Haupteffekt für das Geschlecht (zugunsten der Mädchen,  $F(1,115)=8.01, p=.01, \eta^2=.07$ ) und die Schulnoten (zugunsten guter Noten,  $F(1,107)=7.33, p=.01, \eta^2=.06$ ), aber keine Interaktionseffekte. Die kognitiven Fähigkeiten stehen in keinem Zusammenhang mit der Wirksamkeit des Trainings.

Es lässt sich festhalten, dass das Training im Durchschnitt über alle SuS hinweg positiv wirkt. Es wird kein Matthäus-Effekt, sondern – je nach Operationalisierung des SES – sogar ein kompensierender Effekt gefunden, der die Bedeutung des Trainings für SuS mit niedrigem SES stärkt. Kritisch zu betrachten ist dabei die Frage nach der Operationalisierung des SES. Die zwei von uns angewandten Verfahren führen zu verschiedenen Interpretationen. Kommende Studien sollten daher nicht nur größere Stichproben heranziehen, sondern auch der Operationalisierung des SES besondere Aufmerksamkeit schenken. Bis zur Präsentation der Ergebnisse in Basel werden auch noch weitere Indikatoren für SES (u.a. Beruf der Eltern) in die Analysen aufgenommen.

### **Vermittlung von selbstreguliertem Lernen im Studium: Beeinflussen individuelle Merkmale der Studierenden den Zuwachs von selbstreguliertem Lernen?**

**Sophie Butz<sup>1</sup>, Silke Hertel<sup>1</sup>, Yves Karlen<sup>2</sup>, Henrik Bellhäuser<sup>3</sup>, Birgit Spinath<sup>1</sup>, Katharina Maag Merki<sup>4</sup>, Bernhard Schmitz<sup>5</sup>**

<sup>1</sup>Universität Heidelberg, <sup>2</sup>PH FHNW, <sup>3</sup>Goethe Universität Frankfurt, <sup>4</sup>Universität Zürich, <sup>5</sup>Technische Universität Darmstadt

#### Theoretischer Hintergrund

Die Fähigkeit zum selbstregulierten Lernen (SRL) ist ein wichtiger Prädiktor für akademischen Erfolg (Richardson, Abraham & Bond, 2012). Allerdings weisen viele Studierende deutliche Defizite im Bereich des SRL auf (Peverly, Brobst, Graham & Shaw, 2003; Stark & Mandl, 2005). Folglich ist die Förderung des SRL von Studierenden ein wichtiger Aspekt der universitären Lehre (Benz, 2010; Dignath & Büttner, 2008). SRL wird definiert als ein Prozess, der aus drei Phasen besteht (Schmitz & Wiese, 2006): Die präaktionale Phase ist durch die Lernvorbereitung bestimmt, die aktionale durch das eigentliche Lernen und die postaktionale durch die Reflexion des eigenen Lernprozesses.

Bislang unbeantwortet ist die Frage, ob es Studierendengruppen gibt, die besonders von einem SRL-Training profitieren. Es gibt Hinweise darauf, dass SRL u.a. abhängig ist von Persönlichkeitseigenschaften (Bidjerano & Dai, 2007; Papantoniou, Moraitou, Dinou & Katsadima, 2013). Außerdem konnte gezeigt werden, dass Studierende mit hoher Lern- und Leistungsmotivation höhere Selbstregulationskompetenzen aufweisen als Studierende mit niedriger Lern- und Leistungsmotivation (z.B. Baneshi, Samadieh & Ejei, 2015; Mouratidis, Vansteenkiste, Michou & Lens, 2013; Ohtani, Nakaya, Ito & Okada, 2012).

Hier knüpft unser Beitrag an und untersucht den Einfluss von individuellen Merkmalen der Studierenden (z.B. Abiturnote, Studiengang, Studiendauer), Vorwissen, Lerngewohnheiten sowie Persönlichkeits- und motivationalen Variablen auf den Zuwachs von SRL in zwei Lernumgebungen (Präsenzseminar und webbasierter Kurs).

#### Fragestellung

Der Beitrag befasst sich mit den folgenden Fragestellungen:

- (1) Lässt sich der Zuwachs an SRL durch die Abiturnote, das Geschlecht, das Studienfach und die Studiendauer vorhersagen?
- (2) Profitieren Studierende mit einem höheren deklarativen Wissen in SRL und günstigeren Lerngewohnheiten stärker von dem SRL-Training?
- (3) Beeinflussen Persönlichkeitsmerkmale und Motivation den Lernzuwachs?
- (4) Ergeben sich differentielle Effekte für die beiden Lernumgebungen?

#### Methode

Es wurde eine randomisierte Interventionsstudie mit Lehramts- ( $n = 164$ ) und Bachelorstudierenden ( $n = 54$ ) durchgeführt (Alter:  $M = 23.85$ ; 70% weiblich), die entweder am SRL-Seminar mit Präsenzterminen oder am webbasierten Kurs teilnahmen. Die Vermittlung des selbstregulierten Lernens erfolgte in 5 Lektionen.

Es wurden Fragebögen zu Selbsteinschätzung in der Anwendung der Strategien des selbstregulierten Lernens (SRL@U; Bellhäuser, Roth & Schmitz, 2015), der Lern- und Leistungsmotivation (SELLMO-ST; Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2012), der Lernstrategien im Studium (LIST; Wild & Schiefele, 1994) sowie eine Selbsteinschätzung der Persönlichkeitsfaktoren (BFI-K; Rammstedt & John, 2005) eingesetzt sowie Wissenstests zur Erfassung des deklarativen Wissens (u.a. Bellhäuser et al., 2016) und situationsspezifische Wissenstests zu Strategien beim wissenschaftlichen Schreiben (u.a. MCSWT; Karlen, 2017). Es wurden drei Messzeitpunkte (Prätest, Posttest, Stabilitätshebung) realisiert.

#### Ergebnisse

Die Auswertung der Daten erfolgte mit hierarchischen Regressionsanalysen. Abhängige Variablen waren ein Differenzwert (Post-Prä) über alle Phasen des SRL-Prozesses (Gesamtscore) sowie Differenzwerte für die drei Phasen. Für die erste Forschungsfrage zeigte sich für Studierende in Lehramtsstudiengängen sowie in höheren Fachsemestern ein höherer Zuwachs an SRL-Strategien der präaktionalen Phase. Für die Abiturnote und das Geschlecht ergaben sich keine signifikanten Effekte. Bezogen auf die zweite Forschungsfrage ergaben die Analysen, dass Studierende mit stärker ausgeprägten ressourcenorientierten Lernstrategien einen geringeren Zuwachs an SRL-Strategien der aktionalen Phase aufwiesen. Für das deklarative Vorwissen zeigten sich keine Zusammenhänge mit dem Zuwachs an SRL-Strategien. Hinsichtlich der Persönlichkeitsmerkmale und der Motivation (Forschungsfrage 3) ergab sich ein höherer Zuwachs an SRL-Strategien (Gesamtscore) sowie für SRL-Strategien der postaktionalen Phase für gewissenhaftere Studierende. Studierende mit einer ausgeprägteren Verträglichkeit sowie einer geringeren Offenheit für Erfahrungen wiesen einen stärkeren Zuwachs an SRL-Strategien der präaktionalen Phase auf. Für die Lern- und Leistungsmotivation ergaben sich keine Zusammenhänge mit dem Lernzuwachs. Für Forschungsfrage 4 zeigte sich, dass Studierende mit einer hohen Verträglichkeit stärker von Präsenzveranstaltungen profitierten, wohingegen Studierende mit hoher Gewissenhaftigkeit im webbasierten Kurs einen höheren Lernzuwachs aufwiesen. Die Bedeutung für die Vermittlung von SRL im Studium wird diskutiert.

## Prozessanalytische Betrachtung differenzieller Effekte einer Intervention zur Förderung des selbstregulierten Lernens bei Studierenden

Laura Dörrenbächer, Franziska Perels  
Universität des Saarlandes

Selbstreguliertes Lernen (SRL) wird definiert als „self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal goals“ (Zimmerman, 2000, S. 14) und beschreibt einen Prozess optimaler Zielerreichung im Lernkontext. Im Rahmen sozial-kognitiver Modelle wird SRL als zyklischer Prozess aus Planungs-, Handlungs- und Reflexionsphase beschrieben (Zimmerman, 2000). SRL weist positive Zusammenhänge mit akademischer Leistung in allen Bildungsabschnitten auf (z. B. Dignath, Büttner & Langfeldt, 2008; Kitsantas, Winsler & Huie, 2008) und wird daher als Bestandteil lebenslangen Lernens angesehen (Wirth & Leutner, 2008). Maßnahmen zur Förderung von SRL erscheinen vor allem bei der Zielgruppe der Studierenden relevant (z. B. Bembenuddy, 2011), weil diese Zielgruppe vor besondere Lernherausforderungen gestellt ist. In diesem Kontext zeigen Trainings zur direkten Strategievermittlung (z. B. Bail, Zhang & Tachiyama, 2008) sowie deren Kombination mit Lerntagebüchern (Dörrenbächer & Perels, 2016a) positive Effekte. Verschiedene Autoren fordern im Rahmen eines personenzentrierten Ansatzes (Niemivirta, 2002) die Berücksichtigung individueller Voraussetzungen bei Wirksamkeitsuntersuchungen von SRL-Fördermaßnahmen (Masui & DeCorte, 2005). In Bezug auf das Ausgangsniveau selbstregulativer Lernkompetenzen sprechen bisherige Befunde für einen Kompensationseffekt, also dafür, dass vor allem Studierende mit niedrigen SRL-Werten von einem Strategietraining profitieren (z.B. Gonzalez-Pienda, Fernandez, Bernardo, Nunez & Rosario, 2014).

Das Ziel der vorliegenden Untersuchung war daher die prozessanalytische Analyse differenzieller Wirkungen eines SRL-Präsenztrainings für Studierende in Abhängigkeit von deren SRL-Ausgangsniveau. Das Training umfasste neben sieben wöchentlichen Trainingseinheiten auch ein Lerntagebuch (paper-and-pencil) und konnte längsschnittlich als wirksam evaluiert werden (Dörrenbächer & Perels, 2016a). Der Einsatz von Lerntagebüchern ermöglicht zusätzlich eine prozessanalytische Betrachtung differenzieller Trainingsverläufe (Schmitz & Wiese, 2006).

Die Werte einer Stichprobe von  $N = 337$  Studierenden verschiedener Fachrichtungen ( $M_{\text{Alter}} = 23.48$ ,  $SD = 4.08$ , 71% weiblich) in einem Selbstregulationsfragebogen in Anlehnung an das Prozessmodell von Zimmerman (2000) wurden genutzt, um eine latente Profilanalyse (Vermunt & Magidson, 2002) durchzuführen. Die Ergebnisse indizierten vier SRL-Profile, die sich basierend auf den Werten in den SRL-Subskalen beschreiben lassen (niedriges SRL mit moderater Motivation, moderates SRL, widersprüchliches SRL mit hoher Motivation, hohes SRL, siehe auch Dörrenbächer & Perels, 2016b). Um die individuellen Effekte dieser Profile im Rahmen der Fördermaßnahme zu untersuchen, wurden die Daten von 27 Studierenden analysiert, die neben Trainingseinheiten zur Strategievermittlung auch täglich ein Lerntagebuch bearbeiteten. Die Studierenden verteilten sich gleichmäßig auf die vier SRL-Profile, was eine prozessanalytische Untersuchung der Tagebuchdaten differenziert nach Profilgruppe ermöglichte. Das Lerntagebuch erfasste über 48 geschlossene und vier offene Fragen verschiedene Aspekte des SRL. Fragen zur Planungsphase sollten vor dem Lernen, Items zur Handlungs- und Reflexionsphase nach dem Lernen ausgefüllt werden.

Die Ergebnisse prozessanalytischer Berechnungen zeigen für die Extremgruppen mit niedrigen und hohen SRL-Werten sowohl einen positiven linearen Trend ( $R^2 = .06$ ,  $p = .08$  bzw.  $R^2 = .08$ ,  $p = .06$ ) und einen quadratischen Trend ( $R^2 = .28$ ,  $p < .01$  bzw.  $R^2 = .17$ ,  $p < .05$ ) im selbstregulierten Lernen über den Interventionszeitraum hinweg. Während die Gruppe mit moderaten SRL-Werten einen positiven linearen Trend aufweist ( $R^2 = .08$ ,  $p = .05$ ), zeigt sich für die Gruppe mit widersprüchlichen Ausgangswerten im SRL keine Veränderung im Interventionsverlauf. Neben den Ergebnissen für die Gesamtskala SRL soll im Vortrag auf die differenziellen Interventionswirkungen hinsichtlich verschiedener SRL-Komponenten eingegangen werden (z.B. Zielsetzung). Die Ergebnisse können nicht im Sinne des Kompensationseffektes interpretiert werden, da entsprechend der prozessanalytischen Auswertungen auch Studierende mit hohen Ausgangswerten von der Intervention zu profitieren scheinen. Da jedoch nicht alle Studierenden gleichermaßen von der SRL-Intervention profitieren können, ist die Entwicklung adaptiver Trainingsprogramme zum optimalen Umgang mit bereits vorhandenen SRL-Kompetenzen wünschenswert.

### Wer nutzt ein freiwilliges Selbstregulationstraining? Prädiktoren und Outcomes der Nutzung eines webbasierten Trainings zur Steigerung des selbstregulierten Lernens

Henrik Bellhäuser<sup>1</sup>, Maria Theobald<sup>2</sup>, Margarete Imhof<sup>2</sup>  
<sup>1</sup>Goethe Universität Frankfurt, <sup>2</sup>Universität Mainz

Erfolg im Studium erfordert die Fähigkeit zum selbstregulierten Lernen (SRL) (Richardson, Abraham & Bond, 2012). Webbasierte Trainings (WBT) bieten die Möglichkeit, zeitlich und räumlich flexibel Teilnehmende zu fördern, indem sie über das Internet distribuiert werden (Sitzmann, Kraiger, Stewart & Wisher, 2006). Auf Grundlage des SRL-Prozessmodells (Schmitz & Wiese, 2006) wurde ein WBT entwickelt, welches im Rahmen eines Mathematikvorkurses für Studierende mathematisch orientierter Studiengänge bereits erfolgreich evaluiert wurde (Bellhäuser, Lösch, Winter & Schmitz, 2016).

Im vorliegenden Beitrag wird der Frage nachgegangen, welche Bedingungsfaktoren dazu beitragen, dass Studierende ein solches, freiwilliges Angebot wahrnehmen. Zudem soll analysiert werden, welche positiven Effekte die Nutzung des WBT auf Lernverhalten und Studienleistungen zeigt. Es wird erwartet, dass Eigenschaften der Studierenden, wie Vorwissen, Computerkenntnisse, Alter, Selbstwirksamkeitserwartung, Gewissenhaftigkeit, Offenheit für Erfahrungen und motivationale Faktoren wie Änderungsbereitschaft zur Nutzung des WBTs beitragen (Bidjerano & Dai, 2007; Dörrenbächer & Perels, 2016; Zimmerman, 2000).

Die Versuchspersonen wurden im Wintersemester 2016/17 in drei Vorlesungen rekrutiert: eine Psychologievorlesung im Bachelor Bildungswissenschaften, eine Vorlesung im Bachelor Biologie und eine Vorlesung im Bachelor Chemie. Die Studienteilnahme war freiwillig und wurde durch Versuchspersonengelder entlohnt. Es gab zwei Versuchsbedingungen, denen die Versuchspersonen randomisiert zugeordnet werden: Experimentalgruppe mit Zugang zum WBT und Kontrollgruppe ohne Zugang. Das WBT bestand aus sechs 45-minütigen Lektionen, in denen kognitive und metakognitive Strategien zur Planung, Umsetzung und Überwachung der Lernziele vermittelt wurden. Außerdem umfasste das WBT volitionale und motivationale Strategien zur Regulierung der Aufmerksamkeitsressourcen während des Lernens und der Bewertung der Lernleistungen.

Zum Semesterbeginn (T1) wurden demographische Angaben, Vorwissen, Motivation, Persönlichkeitseigenschaften sowie Selbstwirksamkeit, SRL-Kompetenzen und explizites Wissen über SRL erhoben. Zum Semesterende (T2) wurden neben Selbstwirksamkeit, SRL-Kompetenzen und explizitem Wissen über SRL auch die Noten in der Abschlussklausur erhoben. Prozessdaten wurden in Form von wöchentlichen Logfiles aus einer Online-Lernplattform erhoben.

Es konnten N=358 Versuchspersonen rekrutiert werden (n=216 weiblich, Alter: M=20.4 Jahre (SD=1.89); Fachsemester: M=1.7 (SD=.79). In der Experimentalgruppe erhielten n=192 Personen Zugang zum WBT. Davon bearbeiteten nur n=26 (14%) mindestens vier der sechs Lektionen, weitere 43 Personen (36%) bearbeiteten ein bis drei Lektionen.

In einem multiplen Regressionsmodell konnte gezeigt werden, dass die Anzahl bearbeiteter Lektionen vorhergesagt werden konnte aus der SRL-Kompetenz zu T1 (je niedriger die Ausprägung, desto mehr Beteiligung) und aus den positiven Erwartungen bezüglich des WBTs (je größere Erwartungen, desto mehr Beteiligung). Die weiteren zu T1 erhobenen Variablen erwiesen sich nicht als prädiktiv.

Um die Wirksamkeit des Trainings bei einer so geringen Teilnehmerzahl überprüfen zu können, wurde ein Matching-Ansatz gewählt. Dabei wurden für die n=26 TeilnehmerInnen, die das Training überwiegend bearbeitet hatten, jeweils individuelle Matchingpartner in der Kontrollgruppe ausgewählt, die hinsichtlich Abiturnote, Gewissenhaftigkeit und SRL-Kompetenz zu T1 vergleichbare Werte aufwiesen. Im Vergleich dieser beiden Gruppen zeigte sich, dass die Experimentalgruppe sich signifikant gegenüber der Kontrollgruppe in ihrer selbstberichteten SRL-Kompetenz steigerte und ebenfalls signifikant bessere Noten in der Abschlussklausur erzielte.

Die Ergebnisse zeigen, dass das WBT positive Effekte auf Verhalten und Leistung aufweisen kann. Dies setzt voraus, dass sich die TeilnehmerInnen hinreichend mit dem Training auseinandersetzen. Vergleichbar mit Erfahrungen aus der MOOC-Forschung (Onah, Sinclair & Boyatt, 2014) zeigte sich auch hier, dass nur eine geringe Prozentzahl der angesprochenen Personen wirklich das Training nutzt. Neben den eigenen SRL-Kompetenzen sind dabei die mit einem solchen Training verbundenen Erwartungen prädiktiv dafür, ob es zur Trainingsnutzung kommt oder nicht. Im Vortrag wird die Möglichkeit von individuellem Feedback zur Steigerung der Erwartungen diskutiert.

## D18

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H102

### **Bewertung professionellen Handelns von Lehrenden durch Videographie**

Chair(s): **Jennifer Iglar** (Technische Universität Dortmund, Deutschland), **Annika Ohle** (Technische Universität Dortmund, Deutschland), **Nele McElvany** (Technische Universität Dortmund, Deutschland)

DiskutantIn(nen): **Thamar Voss** (Albert-Ludwigs-Universität Freiburg)

In Bezug auf erfolgreiche Lernprozesse von Lernenden spielen Lehrkräfte als Vermittler eine wichtige Rolle. Für die Vorhersage qualitätvollen Unterrichts sowie der Leistung von Schülerinnen und Schülern rücken besonders individuelle Lehrkräfte Merkmale wie das professionelle Handeln in den Vordergrund. Die Bedeutung des professionellen Handelns von Lehrkräften zeigt sich unter anderem in der Prädiktivität für den Lernerfolg von Schülerinnen und Schülern (Kunter, Klusmann & Baumert, 2009).

Sowohl national als auch international hat sich die videobasierte Lehr-Lernforschung bereits etabliert (Seidel & Thiel, 2017). In der Praxis wird die Videographie häufig als Instrument der Professionalisierung von Lehrkräften zur Unterrichtsentwicklung oder zur Beurteilung professionellen Handelns verwendet (Clausen, Weingarten & Wegner, 2013). Videoaufzeichnungen ermöglichen qualitative sowie quantitative Analysen und lassen wiederholte Beobachtungen und Bewertungen zu (Helmke, 2007; Roth, 2009). Um Aussagen über die Qualität videographierten Unterrichts oder des aufgezeichneten Lehrkraftverhaltens treffen zu können, wird häufig auf hoch inferente Beurteilungen von geschulten Ratern zurückgegriffen. Der emotionale Abstand zum Unterrichtsgeschehen sowie ein höheres Ausmaß an Vergleichsmöglichkeiten sind Vorteile der Perspektive von externen Beobachterinnen und Beobachtern (Rakoczy, 2008). Beobachterratings gelten als am angemessensten für die Erfassung der Komplexität des Unterrichts und des professionellen Handelns. Nachteile sind mögliche Verzerrungen der Urteile durch Rater-Bias sowie fehlende Informationen über Schülerinnen und Schülern und den Klassenkontext.

Im Rahmen des Symposiums wird der Aspekt des professionellen Handelns von Lehrenden als Voraussetzung für qualitätvollen Unterricht und erfolgreiche Lernprozesse beleuchtet. Hierbei nehmen alle Beiträge die Videographie als Methode zur Beurteilung des professionellen Handelns sowie der Unterrichtsqualität in den Fokus. Der erste Beitrag betrachtet Zusammenhänge zwischen der Perspektive von angehenden Lehrkräften und externen Beobachterinnen und Beobachtern, indem Zusammenhänge von Selbsteinschätzungen Lehramtsstudierender und videobasierten Beurteilungen von zwei trainierten Beobachterinnen hinsichtlich der Handlungskompetenz im Bereich Feedback berichtet werden. Erste Resultate zeigten geringe Korrelationen zwischen den Selbsteinschätzungen und den videobasierten Beurteilungen. In dem zweiten Beitrag wird der Frage nachgegangen, inwiefern die Unterrichtsqualität in externen Beurteilungen von Ratern aufgrund einer unterschiedlich administrierten Reihenfolge der Unterrichtsvideos positiver oder negativer eingeschätzt wird. In Hinblick auf die Resultate des Experiments wurden signifikante Unterschiede in der Beurteilung der Unterrichtsqualität aufgrund der sequenziellen Ordnung sowie ein Einfluss von individuellen Merkmalen der externen Rater auf diese Abweichungen berichtet. Der dritte Beitrag bezieht sich auf die Bewertung professionellen Handelns und der Unterrichtsqualität in Schulen mit personalisierten Lernkonzepten. Mittels einer Kombination aus niedrig und hoch inferenten Ratings konnten interindividuelle Unterschiede zwischen Grundschullehrkräften in der Abstimmung des Lernangebotes auf die Bedürfnisse einzelner Schülerinnen und Schüler identifiziert werden. Im vierten Beitrag wurde in Anlehnung an bestehende Verfahren aus der Bildungsforschung ein hoch inferentes Ratingverfahren entwickelt, um die Lehrkompetenz von Hochschullehrenden videoanalytisch abzubilden. Ergebnisse einer längsschnittlichen Untersuchung mit insgesamt 112 Stunden Videomaterial geben aufgrund guter bis sehr guter Raterübereinstimmungen Hinweise auf die Qualität des entwickelten Instruments.

Aus den vier Beiträgen können neue Erkenntnisse zur Beurteilung des professionellen Handelns von Lehrenden mittels der Methode Videographie gewonnen werden. Abschließend an die Beiträge folgt eine Diskussion und Einordnung der Befunde aus methodischer und inhaltlicher Perspektive.

*Beiträge des Symposiums*

### **Wie geben Grundschullehramtsstudierende Schülerinnen und Schülern Feedback? Vergleich zwischen Selbsteinschätzungen und Videoanalysen**

**Miriam Hess<sup>1</sup>, Katharina Werker<sup>2</sup>**

<sup>1</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg, <sup>2</sup>Universität Kassel

Theoretischer Hintergrund: Obwohl Feedback mit einer Effektstärke von  $d=0.73$  eines der lernwirksamsten Unterrichtsmerkmale darstellt (Hattie & Timperley, 2007), zeigen Studien, dass Lehrpersonen den Lernenden nur selten informative Rückmeldungen geben (z.B. Lotz, 2015). Nach Hattie und Timperley (2007) sollte gutes Feedback Informationen zum Lernziel, zum Fortschritt und zu den weiteren Schritten enthalten. Zudem sollte es nicht personbezogen sein, sondern sich auf die Aufgabe, den Bearbeitungsprozess oder die Selbstregulation beziehen. Eine mögliche Ursache für die mangelnde Umsetzung guten Feedbacks in der Unterrichtspraxis könnte darin liegen, dass zentrale Aspekte der Lehrer-Schüler-Interaktion bereits im Studium zu wenig Berücksichtigung finden oder die Auseinandersetzung damit zu theoretisch erfolgt (z.B. Blomberg et al., 2013). Daher sollte dem Thema möglichst bereits in der universitären Lehrerinnen- und Lehrerbildung besondere Beachtung geschenkt werden. Um diese näher an den Erfordernissen der späteren Praxis auszurichten, wird insbesondere dem Einsatz von Videos ein großes Potenzial zugeschrieben (z.B. Goeze, Hefleisch & Schrader, 2013). Videoaufnahmen von eigenen Lehrversuchen bieten zudem die Möglichkeit, die Handlungskompetenz von Studierenden aktiv zu erproben und systematisch zu evaluieren. Häufig wird zur Evaluation von Maßnahmen in der Lehrerinnen- und Lehrerbildung lediglich auf Selbsteinschätzungen der Studierenden zurückgegriffen (zsf. Cramer, 2010; König, Kaiser & Felbrich, 2012). Aufgrund der Subjektivität und der damit vermuteten geringen Verlässlichkeit von Einschätzungen eigener Kompetenzen werden aber im Rahmen einer evidenzbasierten Lehrerinnen- und Lehrerbildung Evaluationsmaßnahmen gefordert, welche die professionellen Kompetenzen systematisch erfassen (König et al., 2012). Bisher haben sich nur wenige Studien mit der Frage beschäftigt, inwieweit Selbsteinschätzungen und mit Testverfahren erhobene Leistungen von (angehenden) Lehrpersonen zusammenhängen (zsf. König et al., 2012). Generell zeigen sich in diesen Studien eher geringe Zusammenhänge, was die Notwendigkeit einer testbasierten Erfassung unterstreicht (z.B. König & Tachtsoglou, 2012).

Fragestellungen: 1. Wie geben Studierende des Lehramts an Grundschulen Schülerinnen und Schülern Feedback? 2. Wie hoch sind die Zusammenhänge zwischen der Selbsteinschätzung durch die Studierenden und der videobasierten Beurteilung der Handlungskompetenz im Bereich Feedback?

Methode

Die Daten stammen aus vier Blockseminaren zum Thema Feedback, an denen im Wintersemester 2016/17 und im Sommersemester 2017 insgesamt 116 Studierende des Grundschullehramts der Universität Nürnberg teilgenommen haben (87% weiblich; durchschnittlich 4.-5. Fachsemester). Zur systematischen Evaluation der Seminare wurden Tests sowie Befragungen der Studierenden durchgeführt, wobei im Rahmen des Vortrags näher auf die Erfassung der Handlungskompetenz eingegangen wird: Dazu wurde nach dem Seminar jede/r Studierende in einer etwa 30-minütigen Interaktionssituation mit einem einzelnen Kind videografiert. Die Aufgabe bestand darin, an einer vorgegebenen komplexen mathematischen Aufgabe zu arbeiten und die/den Lernende/n dabei durch möglichst gutes Feedback zu unterstützen. Diese Videos werden von zwei trainierten Beobachterinnen unabhängig voneinander mit einem hoch inferenten Ratingsystem ausgewertet, wobei jedes Item auf einer vierstufigen Skala eingeschätzt wird. Das Beobachtungssystem dient zur Beurteilung der Qualität der erteilten Rückmeldungen und umfasst insgesamt 21 selbst entwickelte Items. Die Beobachterübereinstimmungen wurden über den gesamten Ratingprozess fortlaufend überprüft (relativer G-Koeffizient  $> .70$ ; vgl. Lotz, Berner & Gabriel, 2013). Zusätzlich wurden die Studierenden zu ihrer Selbsteinschätzung befragt: Neben einer schriftlichen Reflexion ihrer eigenen Videoaufnahme sollten sie ihrem eigenen Feedback-Verhalten im Video eine Schulnote geben. Zudem schätzten die Studierenden anhand einiger Items ihre Handlungskompetenz im Bereich Feedback ein (z.B. Ich kann Schülern gutes Feedback geben.).

#### Ergebnisse

Die bisherigen Ergebnisse aus dem Wintersemester 2016/17 zeigen, dass die Qualität des Feedbacks mit einem Mittelwert von 4.08 (SD=0.91) auf einer sechsstufigen Skala als eher hoch eingeschätzt wird. Die Studierenden selbst bewerten ihre Videos durchschnittlich mit der Note 2.82 (SD=0.77) und ihre Selbsteinschätzung zum Item Ich kann Schülern gutes Feedback geben liegt bei  $M=2.77$  (SD=0.78) auf einer vierstufigen Skala. Die Korrelationen zwischen der videobasierten Einschätzung und den Selbsteinschätzungen sind sehr gering und nicht signifikant. Ob sich diese Tendenz auch in der Gesamtstichprobe zeigt, wird im Vortrag berichtet.

### Reihenfolgeeffekte bei Videoratings von Unterrichtsqualität

**Jennifer Igler, Annika Ohle, Nele McElvany**  
Technische Universität Dortmund, Deutschland

#### Theoretischer Hintergrund

Für die Erfassung von Unterrichtsqualität wird neben der Lehrkraft- und Schülerperspektive häufig die Perspektive von externen Beobachterinnen und Beobachtern herangezogen. Untersuchungen zeigten, dass die Ratings über Unterrichtsqualität der verschiedenen Perspektiven nur gering übereinstimmen (vgl. Clausen, 2002; Fauth, Decristan, Rieser, Klieme & Büttner, 2014; Kunter & Baumert, 2006).

Als Merkmale qualitätsvollen Unterrichts haben sich die Kognitive Aktivierung, Klassenführung und Motivierung erwiesen (vgl. Klieme & Rakoczy, 2008). Hinsichtlich der Klassenführung ist die Bewertung von externen Ratern der direkteste Weg (Pianta & Hamre, 2009). Aufgrund der hohen Validität und Prädiktion für die Schülerleistungen (Fauth, Decristan, Rieser, Klieme & Büttner, 2014), wird die Befragung von externen Beobachterinnen und Beobachtern als Königsweg zur Erfassung der Unterrichtsqualität betitelt (Helmke, 2010). Allerdings ist die fehlende Stabilität von externen Ratings ein großer Nachteil. Untersuchungen zu Beobachterratings zeigten, dass teilweise hohe Varianzanteile auf Rater-Bias zurückzuführen sind (Pietsch & Tosana, 2008; Praetorius, 2014).

Ein Rater-Bias, der bei den Beurteilungen von externen Beobachterinnen und Beobachtern auftreten kann, ist der Reihenfolgeeffekt. Dieser Effekt ist ein Phänomen, der bei Gedächtnisprozessen auftritt und auf die Arbeitsweise des Kurzzeitgedächtnisses zurückzuführen ist (schon Ebbinghaus, 1885). Hierbei hat die sequenzielle Position, in der Informationen präsentiert werden, einen Einfluss auf die Beurteilung (Cushman & Mele, 2008). Darüber hinaus beeinflussen individuelle Faktoren wie Geschlecht, Alter und Studiengang (Praetorius, 2014), die Konzentrationsfähigkeit (Hogarth & Einhorn, 1992), aktuelle Stimmung (Murphy & De Shon, 2000) sowie Erfahrungen im schulischen Bereich (Strong, Gargani & Hacifazlioglu, 2011) Verzerrungen von Ratings. Jedoch untersuchten bisher nur wenige Studien mögliche Rater-Effekte in der Unterrichtsforschung (vgl. Pietsch & Tosana, 2008; Praetorius, 2014).

#### Fragestellungen

F1 Lassen sich systematische Unterschiede in den Ratings der Unterrichtsqualität aufgrund der sequenziellen Darbietungsordnung von Unterrichtsvideos mit unterschiedlichen Qualitäten finden?

H1 In Abhängigkeit von der sequenziellen Ordnung werden Unterrichtsvideos mit einer mittleren Qualität (a) nach einer vorigen Bewertung eines Unterrichtsvideos mit einer niedrigeren Qualität positiver oder (b) nach einer vorigen Bewertung eines Unterrichtsvideos mit einer höheren Qualität negativer bewertet.

F2 Welche Faktoren haben einen Einfluss auf die Unterschiede in den Ratings des gleichen Videos?

H2 Die Faktoren Geschlecht, Alter, Lehramtsstudium, Konzentrationsfähigkeit, aktuelle Stimmung und Erfahrungen im schulischen Bereich beeinflussen die Unterschiede in den Ratings.

#### Methode

Insgesamt nahmen  $N = 69$  Studierende ( $n = 34$  im Lehramt und  $n = 35$  in anderen Studiengängen) an der Studie teil ( $M_{\text{Alter}} = 24.92$ ,  $SD = 4.40$ ;  $M_{\text{Semester}} = 7.83$ ;  $SD = 4.66$ ). Um die Vergleichbarkeit der Beurteilungen sicher zu stellen, wurden alle Studierenden im Bereich der Unterrichtsqualität geschult. Die Interraterreliabilitäten erreichten nach einem zweiten Trainingsvideo Werte von ICCkognitive Aktivierung = .82., ICCKlassenführung = .86 und ICCMotivierung = .63. Für die Erfassung der Ratings von Unterrichtsqualität gab es drei Versuchsgruppen, die jeweils zwei zehnminütige Videos mit einer niedrigen, mittleren oder hohen Unterrichtsqualität in unterschiedlicher Reihenfolge präsentiert bekamen und direkt im Anschluss an jedes Video die Unterrichtsqualität anhand eines Ratingmanuals eingeschätzten.

#### Ergebnisse

Die Analysen wiesen signifikante Unterschiede in den Ratings der Kognitiven Aktivierung ( $F[2, 66] = 6.96$ ;  $p < .05$ ,  $\eta^2 = .17$ ), Klassenführung ( $F[2, 66] = 4.10$ ;  $p < .05$ ,  $\eta^2 = .11$ ) und Motivierung ( $F[2, 66] = 6.07$ ;  $p < .05$ ,  $\eta^2 = .16$ ) in Abhängigkeit von der sequenziellen Ordnung auf. Die Qualität des Unterrichtsvideos mit einer mittleren Unterrichtsqualität wurde positiver nach einer vorangestellten Bewertung eines Unterrichtsvideos mit einer niedrigeren Qualität und negativer nach einer vorigen Beurteilung eines Unterrichtsvideos mit einer höheren Qualität eingeschätzt. Einen Einfluss auf diese Unterschiede zeigten die Kovariaten Alter, Lehramtsstudium, Konzentrationsfähigkeit, aktuelle Stimmung und Erfahrungen im schulischen Bereich für einzelne Unterrichtsqualitätsmerkmalen.

Demzufolge ist die Einschätzung der Unterrichtsqualität von externen Ratern von der sequenziellen Ordnung der einzuschätzenden Videos abhängig. Die Bedeutung dieses Rater-Effekts für Forschung und Praxis wird diskutiert.

## **Bewertung professionellen Handelns und Unterrichtsqualität in Schulen mit personalisierten Lernkonzepten**

**Patricia Ackermann<sup>1</sup>, Christine Pauli<sup>2</sup>**  
<sup>1</sup>Universität Zürich, <sup>2</sup>Universität Freiburg

### Theoretischer Hintergrund

Schulen stehen aktuell vor vielfältigen Herausforderungen. Sie sollen u.a. produktiv mit heterogenen Schülervoraussetzungen umgehen, alle Lernenden an die curricularen Grundanforderungen heranführen, Kompetenzen des lebenslangen Lernens durch vermehrte Autonomiestpielräume und Gelegenheiten zum selbstregulierten Lernen fördern (Stebler, Pauli & Reusser, 2017). Um diesen Anforderungen gerecht zu werden, setzen Schulen zunehmend auf innovative, individualisierende Schul- und Unterrichtskonzepte wie beispielsweise „personalisiertes Lernen“. Unter diesen Sammelbegriff lässt sich eine wachsende Gruppe von Schulen in der deutschsprachigen Schweiz einordnen, die von traditionellen Organisationsprinzipien wie Jahrgangsklassen, dem Stundentakt oder fest zugeordneten Klassenzimmern teilweise stark abweichen und neue Formen der Schul- und Unterrichtsorganisation etabliert haben. Typischerweise beschäftigen sich die Lernenden an solchen Schulen über längere Zeitabschnitte in flexiblen Gruppen und räumlichen Anordnungen (z.B. „Lernlandschaften“ oder „Ateliers“) mit individuellen Lernangeboten. Dies bringt veränderte Tätigkeiten und Rollen der Lehrpersonen mit sich. Neben traditionellen Lehrtätigkeiten (z.B. Vorbereitung und Durchführung durchstrukturierter Unterrichtsstunden, erklären, beurteilen) erhalten insbesondere Formen der individuellen adaptiven Lernunterstützung, der Diagnostik individueller Lernstände und -ergebnisse und des Coachings selbstständig lernender Schülerinnen und Schüler mehr Gewicht. Angesichts der veränderten Lehr- und Lernaktivitäten ist davon auszugehen, dass die bisher vorliegenden Instrumente videobasierter Qualitätsbeurteilung nur bedingt geeignet sind, solchen Unterricht zu beurteilen. Die Instrumente sind in der Regel auf traditionelle Organisationsformen, d.h. fachlich und zeitlich klar abgrenzbare Unterrichtseinheiten im Klassenverband - geführt von der Lehrperson - ausgerichtet. Eine Qualitätsbeurteilung ist wichtig, da die bisher vorliegende empirische Evidenz zur Wirksamkeit individualisierender Unterrichtsformen in Bezug auf die Fachleistungen ein gemischtes Bild zeigt. Es ist davon auszugehen, dass die Unterrichtseffektivität von der tiefenstrukturellen Qualität abhängt (z.B. Bohl, Budde, Rieger-Ladich, 2017; Lipowsky & Lotz, 2015).

Im Rahmen des Forschungsprojekts „perLen“ (personalisierte Lernkonzepte in heterogenen Lerngruppen) wurde deshalb eine Videostudie (Teilprojekt perLa) in zwei Schulen durchgeführt. Diese hat zum Ziel, das professionelle Handeln der Lehrperson in personalisierten Lernlandschafts-Arbeitsphasen zu untersuchen. Am Symposium werden Instrumente, Durchführung und erste Ergebnisse präsentiert und zur Diskussion gestellt.

### Fragestellungen

Mit welchen methodischen Verfahren und Auswertungen kann das professionelle Handeln der Lehrperson im personalisierenden Unterricht erfasst und beurteilt werden? Wie werden die Schülerlernaktivitäten beim personalisierten Lernen adaptiv unterstützt?

### Methode

In zwei Schulen mit personalisierten Lernkonzepten wurde die Fachlehrperson Mathematik, je drei Zielschüler (9.Sj) und das gesamte Lernarrangement mit mehreren Kameras gefilmt. Dies hat dem Zweck, sowohl Lehrperson und Lernende als auch deren Interaktionen und die Lernumgebung zu erfassen. Die Filmaufnahmen aus drei Perspektiven (Stellkamera Lernraum, Verfolgerkamera Lehrperson, Kopfkamera Schüler) wurden zu je einem Strang pro Zielschüler synchronisiert.

Die Analyse der sechs Videostränge erfolgte in zwei Schritten mittels (1) eines Codesystems (time-sampling) und (2) eines Rasters zur Beschreibung von Episoden (event-sampling). Mit diesem Vorgehen wurden niedrig und hoch inferente Verfahren kombiniert. Aktuelle Arbeiten (z.B. Lotz, Gabriel, Lipowsky, 2013) zeigen, dass dieses kombinierte Vorgehen Zusammenhänge und Veränderungen aus multiplen Perspektiven dokumentieren lässt und dass reliable und valide Daten generiert werden können.

(1) Beim niedrig-inferenten Rating wurden mittels eines induktiv-deduktiv entwickelten Kategoriensystems 9 Kategorien im time-sampling (10-Sekunden-Intervalle) kodiert (z.B. Unterrichtsorganisation, Unterrichtsmethode/Handlungsmuster, Schülerrolle, Lehrpersonen-Kommunikation, Aufmerksamkeitsverhalten Zielschüler).

(2) Beim hoch-inferenten Rating wurden Ereignisse ausgewählt und im event-sampling kodiert und/oder eingeschätzt: z.B. Kontext-Kode (z.B. Raum, Regeln, Routinen, Rituale, Hilfesysteme, Kooperationsformen, besondere Vorkommnisse, Warten auf Unterstützung), Interaktion zwischen Lehrperson und Zielschülern (z.B. indiv. Lernunterstützung), Schüleraktivität (z.B. Exploration, Durcharbeiten, Vertiefen), Aktivität Zielschüler vor und nach der Unterstützung.

### Ergebnisse

Die Auswertungen konzentrieren sich auf die Abstimmung des Lernangebots auf die Bedürfnisse der einzelnen Schüler/innen und auf die durch diese wahrgenommene Lernunterstützung. Von Interesse ist, wie das professionelle Handeln der Lehrperson (z.B. Aspekte der Lernunterstützung) mit der Nutzung seitens der Lernenden zusammenhängen. Die ersten Analysen deuten darauf hin, dass beträchtliche interindividuelle Unterschiede bei der Art der Nutzung der Lernzeit durch die Zielschüler sowie bei der Art und der Qualität der Lernunterstützung bestehen.

## **Hoch inferentes Rating zur Bewertung von Lehrkompetenzentwicklung in der Hochschullehre**

**Carmen Heckmann**  
Goethe-Universität Frankfurt

Werden Lehrende als wissenschaftlich vorgebildete, reflektierte Praktiker (z. B. Tenorth 2006) betrachtet, die stets weiter lernen, kommt dieser Einschätzung im Hinblick auf Hochschullehrende eine besondere Bedeutung zu, da diese meist ohne vorherige systematische Ausbildung in die eigenständige Lehre einsteigen und teilweise über wenig systematisches Wissen hinsichtlich der Wirkfaktoren von Lehre verfügen. Die Qualität von Lehr-Lernprozessen im Unterrichtsverlauf wurde an Schulen bereits mit verschiedenen Schwerpunkten analysiert, um Faktoren zu finden, die Unterrichtsqualität zur Erreichung einer optimalen Lernleistung bestmöglich vorhersagen können. Die lernwirksamen Unterrichtskomponenten universitärer Lehre wurden dagegen bisher vergleichsweise wenig erforscht. Insbesondere der Forschungsansatz direkter Beobachtung mittels Videoanalysen zur Ermittlung von Lehrkompetenz einer universitären Lehrperson wurde bisher wenig beachtet. Um

Lehrkompetenz in der Hochschullehre möglichst objektiv erfassen zu können, scheint die direkte Beobachtung zur Beschreibung real gezeigter Verhaltensweisen jedoch ein interessanter und vielversprechender Forschungsansatz zu sein, da auf diese Weise Kompetenzen (Weinert, 2001) in Kompetenzprofilen (Oser, Curcio & Duggeli, 2007) gebündelt betrachtet und analysiert werden können.

Ausgehend von allgemeinen Mustern erfolgreichen Unterrichtens, (u.a. Helmke, 2009; Klieme, 2006; Baumert & Kunter, 2006), kann angenommen werden, dass viele Theorien, Modelle und wissenschaftliche Befunde grundsätzlich von der Lehr-Lern-Forschung an Schulen auf den Hochschulkontext übertragen oder für diesen adaptiert werden können (Marzano, 1998), wenn dies unter Berücksichtigung der unterschiedlichen Kontexte und Voraussetzungen der Lernenden erfolgt. Erfolgreiches Lernen an Hochschulen, so legen aktuelle empirische Studien nahe (z. B. Hattie, 2009; Johannes et al., 2011), wird zu einem großen Teil von den gleichen Einflussgrößen der Unterrichtsqualität bestimmt wie im schulischen Unterricht. Ausgehend von dieser Annahme wurden für die Entwicklung eines hoch inferenten Beobachtungsinstrumentes zur Erfassung der Lehrkompetenz in der Hochschullehre jene Lehrverhaltensweisen aufgegriffen, die vielfach Berücksichtigung in der Bildungsforschung fanden (z.B. Hiebert, et al., 2003; Hugener, Pauli & Reusser, 2006; Kunter, 2005; Seidel, Prenzel, Duit, & Lehrke, 2003).

In diesem Beitrag wird insbesondere der Frage nachgegangen, wie die Lehrkompetenz von Hochschullehrenden bestmöglich videoanalytisch abgebildet werden kann. Dabei wurde für die Entwicklung des hoch inferenten Ratingverfahrens auf das Vorgehen zur Entwicklung von Kategoriensystemen nach Bos und Tarnai (1999) zurückgegriffen und zunächst theoriegeleitet acht beobachtbare Indikatoren zur Erfassung von Merkmalen kompetenten Hochschullehrverhaltens aus etablierten Beobachtungssystem aus der Bildungsforschung adaptiert und ergänzende Indikatoren neu generiert (Lernzielorientierung, Kursorganisation, Organisation des Lernstoffs, Aktivierung der Studierenden, Interaktion Lehrende – Studierende, Lernatmosphäre, Präsentationskompetenz, Anforderungen). Jedes Merkmal wurde anhand von fünf bis neun Indikatoren (Prädiktoren) präzisiert und in einer Auswertungsanleitung beschrieben. Zur Bestimmung der Gütekriterien und der Änderungssensitivität des Instrumentes wurde eine längsschnittliche Untersuchung durchgeführt, an der 15 Lehrende der Goethe-Universität teilnahmen. Alle Lehrenden wurden einmal pro Semester während ihrer Seminarveranstaltung auf Video aufgezeichnet. Das entwickelte Ratingverfahren wurde anhand der so entstandenen 112 Stunden Videomaterial mit drei geschulten Ratern überprüft. Für die Berechnung der Interraterreliabilität der Rater ergab der ICC (3), 2-way-mixed average measure für die Beobachtungsmerkmale eine gute (Interaktion Lehrende-Studierende:  $F = 7.32$ ,  $p < .001$ ,  $ICC(3,3) = .81$ ) bis exzellente (Aktivierung der Studierenden:  $F = 22.8$ ,  $p < .001$ ,  $ICC(3,3) = .95$ ) Beobachterübereinstimmung. Um den auf die Beobachtungsbestandteile Unterricht, Rater und die unsystematische Fehlervarianz entfallenden Varianzen noch weiter zu differenzieren, wurde zusätzlich der Ansatz der Generalisierbarkeitstheorie von Cronbach et al. (1972) verfolgt und eine 2-Facetten-G-Studie durchgeführt. Die Ergebnisse der G-Analyse über 31 Videos bestätigte, dass bei der Beurteilung durch drei geschulte Rater der auf die Rater zurückzuführende Anteil der erklärten Varianz einen minimalen Anteil von 1-2% ausmacht und der durch unsystematische Fehler erklärbarer Varianzanteil mit  $\leq 25\%$  akzeptabel ausfällt. Der größte erklärende Varianzanteil entfällt somit wie intendiert mit  $\geq 73\%$  auf die Unterschiede zwischen den videografierten Unterrichtseinheiten. Abschließend wird in diesem Vortrag diskutiert, welchen Beitrag das Ratingverfahren zur Erweiterung des Repertoires an Messinstrumenten für den Forschungskontext sowie die Professionalisierung von Lehrkräften in der Hochschule leisten kann.

## D19

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H001

### Academic choices, prior knowledge and knowledge acquisition: Keeping students in the STEM fields

Chair(s): **Nele Kampa** (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN, Deutschland))

DiskutantIn(nen): **Richard Shavelson** (Stanford University, Graduate School of Education)

The three contributions of the symposium bring together two research strands in order to investigate predictors for successful science careers in higher education. The first strand targets the relevance of science-related academic choices in upper secondary education. These choices already lead to varying prior knowledge in the science subjects by the end of the school career and channel later choices of STEM studies (Pinxten, de Fraine, van den Noortgate, van Damme, & Anumendem, 2012). The second strand targets knowledge development during the beginning of STEM studies. Once students have chosen STEM-studies prior knowledge (Buschhüter, Spoden & Borowski, 2017) as well as knowledge development during the beginning of the studies is a good predictor for study success in STEM-studies. The contributions of the symposium focus on these two aspects in order to answer the overarching research question on what prepares students for and keeps them in STEM-studies. More specifically, the contributions shed light on the transition from school science to STEM-studies at university and predictors for STEM-study orientation and success from various disciplines of educational research. Therefore, this critical time period needs special attention from educational research.

The symposium unites presentations from empirical educational research as well as from science education research. The overarching research questions which choices STEM-students have made during their school education and how their knowledge develops during the introductory phase will be answered. The first contribution Who chooses physics studies focusses on academic choices in upper secondary education of physics students who became teachers and 1-field bachelor students. While existing studies mostly focus on earlier choices channeling later broad STEM choices, the authors take the next step by looking at their impact on the choices of two study options (teacher vs. 1-field) within the same field. The contribution gives profound evidence on how schools can better contribute to STEM-orientation of students.

The following two contributions focus on the next step on the pathway towards STEM-professions: the introductory phase of STEM-studies at university level. The introductory phase is a critical point because the majority of drop-outs and switching between studies takes place during this phase (Heublein, Ebert, Hutzsch, Isleib, König, Richter, Woisch, 2017). Both contribution investigate the development of domain-specific knowledge during that time and how this development is related to study success. The second contribution Development of different knowledge types focusses on four knowledge types in biology and physics students and find interesting differences between the two student groups. The third contribution Physical and mathematical modelling competence focusses on the domain-specific predictors content knowledge and mathematical ability as well as physical and mathematical modeling competence for passing introductory physics exams of physics students. The authors find differing results for high and low achievers and an incremental effect of modeling competence for study success. These two contributions give profound evidence on how universities can shape their courses in order to prevent drop-out and promote study success of STEM-students.

The renowned discussant, both in science education and in higher education, will discuss the three contributions and synthesize the results.

#### Beiträge des Symposiums

### Who chooses physics studies: Gender differences of educational pathways of teacher candidates and physics 1-field bachelor students academic choices at school

**Kampa Nele<sup>1</sup>, David Buschhüter<sup>2</sup>, Andreas Borowski<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), <sup>2</sup>Lehrstuhl Didaktik der Physik, Institut für Physik und Astronomie, Universität Potsdam

#### Introduction

Too few students, and even less women, decide for STEM studies (for Europe: <http://ec.europa.eu/eurostat/statistics-explained/index.php/>

[Tertiary\\_education\\_statistics](#)). Various studies showed that students' academic choices during their school career are a precursor for later choices for university studies (e.g., Pinxten, de Fraine, van den Noortgate, van Damme, & Anumendem, 2012). Students not choosing science tracks in school may be prepared for math- and science-related occupations (e.g., automotive mechatronics) but not for science studies (Author, 2016; 2017). Therefore, these early decisions need to be investigated. Also, in-depth studies with large sample sizes on students' choices within one STEM field (e.g., teacher candidates vs. 1-field bachelors) are missing.

In this light, we investigate differences between two physics student groups: female and male physics teacher candidates and 1-field bachelors. Teacher candidates who need to decide for at least two study fields might show a greater variety in interests leading to less choices of science courses. Women deciding for physics studies might need to act against the female stereotype when making that decision (Hannover & Kessels, 2004) which could lead to science career decisions in later years – at university level. If students do not feel sufficiently equipped for physics studies, they can take a preparatory course at university.

#### Methods

In 2013, 2322 freshmen physics students from 24 German universities took part in the data collection during the first three weeks of their first semester. Students were on average 20 years old (SD = 3.30) and 20 % were female (both 1 % missings). While 15 % of the students are teacher candidates, 83 % studied physics as a 1-field bachelor (1 % missings). The students filled in a questionnaire with school-related questions (e.g., chosen physics or mathematics advanced course, number of lessons in the three sciences and mathematics during upper secondary education). We report descriptive findings and correlational analyses on these academic choices during upper secondary education.

#### Results and Discussion

First we investigated the distribution of students across the two study choices (Table 1).

Table 1.



Distribution of gender within teacher candidates and physics 1-field bachelors in percentages (missings: 1.21 %)

Studies

Gender 1-field bachelor Teacher candidates

Male 68.69 9.73

Female 14.73 5.64

Less women opt for a 1-field bachelor but also to become a teacher candidate ( $t(2292)=8.29$ ,  $p < .05$ ,  $r = .17$ ). Also male 1-field bachelor students and male teacher candidates have significantly more often chosen an advanced physics course at school ( $t(1686) = 3.15$ ,  $p < .05$ ;  $r1\text{-field} = .08$ , missings = 0.29 %;  $t(320) = 1.64$ ,  $p < .05$ ;  $r\text{teacher} = .13$ ). We did not find differences for advanced mathematics courses in school ( $t(1687) = 1.64$ ,  $p > .05$ ;  $r1\text{-field} = .04$ ;  $t(327) = 1.56$ ,  $p > .05$ ;  $r\text{teacher} = .09$ ). However, female teacher candidates chose preparatory courses more often ( $t(350) = -3.15$ ,  $p < .05$ ;  $r = -.17$ ), but not female 1-field bachelor students ( $t(1922) = -1.65$ ,  $p > .05$ ;  $r = -.04$ ). To explore the effect of these choices on the amount of exposure to physics and mathematics, ergo the amount of university preparation, our analyses will be extended to the actually taken number of lesson hours.

The effect of gender on participation in preparatory courses shows that especially female teacher candidates take preparatory courses. Implications will be discussed for school and university studies. Either women might choose preparatory courses more often to compensate their choices during upper secondary education. Or male teacher candidates might be a risk group and preparatory courses should be designed to motivate their participation. Also, advanced as well as general physics and mathematics lessons in upper secondary education should prepare students for university studies.

### Development of different knowledge types during the first two semesters of Biology and Physics studies

**Torsten Binder<sup>1</sup>, Theyßen Heike<sup>2</sup>, Philipp Schmiemann<sup>1</sup>**

<sup>1</sup>Didaktik der Biologie, Universität Duisburg-Essen, <sup>2</sup>Didaktik der Physik, Universität Duisburg-Essen

Dropout rates in STEM subjects at university level are high, but varying (e. g. Biology 27% Physics about 40%). This calls for detailed analyses of the first semesters at university. Since, there is a strong relationship between prior knowledge and study success (or dropout respectively), the development of prior knowledge and the conditions for its development seem to be a reasonable focus. This study analyses the development of four different types of knowledge over the first two semesters in Biology and Physics taking into account affective entry conditions of the students.

Theoretical Background

Prior knowledge is one major predictor for university success (Burton & Ramist, 2001). Students prior knowledge is the base for future learning (Dochy, 1996; Thompson & Zamboanga, 2003). Thereby knowledge development is interrelated with students' prior knowledge and other non-cognitive variables (Krapp, 1999). In the context of knowledge assessment, Dochy, Segers, and Buehl (1999) emphasize to target different types of knowledge. This is reflected by a normative model of Hailikari, Nevgi and Lindblom-Ylänne (2007) describing four different types of knowledge: Knowledge of facts, Knowledge of meaning, Integration of knowledge and Application of knowledge. Hailikari (2009) also suggests using different assessment methods to assess these different types of knowledge.

Research Questions

Regarding the background, it seems reasonable to have a deeper look on students' knowledge development differentiating in Biology and Physics:

To what extent do the four types of knowledge develop within the first semester?

In addition we assess how non-cognitive variables are correlated with type-specific knowledge development.

Methods

We assessed students' knowledge at two German universities at the beginning and the end of the first semester (initial sample:  $N = 162$  Biology B.Sc. students and  $N = 100$  Physics B.Sc. students). For every type of knowledge, we applied a specific test in each domain:

- Knowledge of facts: Multiple Choice-Items e.g. naming of a plant organ
- Knowledge of meaning: Constructed response, e.g. giving a definition of the concept "Harmonic Oscillator"
- Integration of knowledge: Concept Maps (fill-in-the-map task)
- Application of knowledge: Sorting of subject-specific problem solving tasks (adapted from Friege, 2001)

To control for non-cognitive variables we also gathered data for learning strategies interest and motivation. For further analysis, we also assessed the types of knowledge after the second semester.

Results and Discussion

In a first step, the prior knowledge was tested for dimensionality using IRT models. Analysis reveals that a four dimensional model fits best - in both domains.

Regarding knowledge development, we found significant changes in different knowledge types: In Biology, the students gain knowledge in all four types within the first semester. The highest change is in knowledge of facts ( $\Delta M = .50$ ,  $SD = .69$ ,  $T = -7.4$ ,  $p < .001$ ).

In Physics, only integration and application of knowledge changed significantly during the first semester with higher change in application of knowledge ( $\Delta M = .48$ ,  $SD = 1.4$ ,  $T = -3.3$ ,  $p < .01$ ).

These findings seem reasonable, taking into account, that the first semester offers new fields for Biology students (e.g. zoology, botany, cellular biology) with high amount of factual knowledge while Physics students have to apply their school knowledge – at least partly – on novel problems, especially in mechanics and electrodynamics. Additional findings for the relationship between the development of the types of knowledge and non-cognitive variables and for the second semester will be presented in the talk.

### Physical and Mathematical Modelling Competence as Predictor for Success in Physics Degree Courses

## Theoretical Background

In order to account for the high number of physics students who drop out (40%) (Heublein et al. 2017; Nienhaus, 2010) at German universities, it is aimed at describing the success in physics degree courses by means of subject-related, psychological and demographic variables. Our study, funded by the German Research Community (DFG), is part of the ALSTER project which investigates academic learning and students' success in degree courses in science (and civil engineering) in the first year of university. In physics, for the majority of students who drop out, the subject-related requirements are decisive to stay at university (Albrecht, 2011). Accordingly, recent studies show that students' mathematical knowledge at the beginning of their first year predicts success in physics degree courses complemented by physical knowledge (Buschhüter, Spoden & Borowski, 2017; Buschhüter, Spoden & Borowski, 2016; Sorge, Neumann, & Petersen, 2016). Additionally, a successful physics student must be able or should learn to transform physical models into mathematical models and vice versa while solving typical tasks in modules of the first semesters. This transformation ability is introduced as physical and mathematical modelling competence in such a way that understanding physical description of a phenomenon involves learning to simultaneously apply and translate between various representations (Angell, Kind, Henriksen & Guttersrud, 2008; Opfermann, Schmeck & Fischer, 2017, Uhden, Karam, Pietrocola & Pospiech, 2011). We assume this relationship, the physical and mathematical modelling competence, to be a predictor for success in physics degree courses (Müller, Fischer, Borowski & Lorke, 2017). A pilot study was successfully completed indicating a good reliability (EAP Rel. from .78 to .89) of all instruments that measure the subject-related variables. Eventually, the goal of this study is to investigate the relationship between the students' modelling competence and their success in physics degree courses leading to two research questions.

1. Is there an interrelation between modelling competence and success in physics degree courses?
2. Does modelling competence account for a surplus to predict success in physics studies?

## Method

We operationalize the variables mentioned above by employing subject-related, psychological and demographic tests in a longitudinal design at three different points of measurement, namely before the start of the physics degree course (10/2016), half a year into the course (02/2017) and at the end of the first year (07/2017). This required the design of three tests; a test on content knowledge, a test on physical and mathematical modelling competence (both in a multiple-choice format) and a test on mathematical ability (open format). Further, we obtain additional data from the ALSTER project to what extent the ability to model is related to extracurricular variables, e.g. gender, socioeconomic status, school grades, and psychological variables, e.g. knowledge on metacognition, knowledge about learning strategies, the academic self-concept, general motivation to learn and study, expectations regarding the course, stress and subject-related interest.

## Results

In the first semester, there is an overall significant gain in knowledge in both content knowledge and in mathematical ability. However, a gain in modelling competence can only be measured for low achievers. Moreover, modelling competence discriminates well between students who pass their introductory physics exam and students who fail it. In a multiple linear regression the mathematical ability is the best predictor for exam grades at the end of the first semester, and modelling competence accounts for a surplus in predicting success in physics degree courses. Teaching modelling competence explicitly should be analyzed in an intervention and may prompt the development of a tuition concept for German, and ultimately European, physics departments.

## D20

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Kollegienhaus H120

### **Kommt es auf die Mitschüler/innen an? Effekte der Schul- und Klassenzusammensetzung auf verschiedene Kriteriumsvariablen**

Chair(s): **Arvid Nagel** (Institut Professionsforschung und Kompetenzentwicklung; Pädagogische Hochschule St.Gallen, Schweiz), **Horst Biedermann** (Pädagogische Hochschule St.Gallen)

DiskutantIn(nen): **Eckhard Klieme** (DIPF - Deutsches Institut für Internationale Pädagogische Forschung)

Wenn die Schul- und Klassenzusammensetzung der Schüler/innen in den bildungswissenschaftlichen Blick genommen wird, so verbinden sich damit (noch immer) viele unbeantwortete Fragen nach möglichen Auswirkungen für die einzelnen Schüler/innen wie die folgenden: Werden Fachleistungen bzw. Lernfortschritte der Schüler/innen durch die Zusammensetzung von Lerngruppen bzw. Klassen und Schulen beeinflusst? Wenn ja, welche Zusammensetzungen wirken sich wie aus? Sind neben den Fachleistungen auch Auswirkungen auf Aspekte wie das Selbstkonzept, die Motivation, Verhaltensprobleme oder das Wohlbefinden festzustellen? Hätte eine Schülerin/ein Schüler bei anderer Klassenzusammensetzung deutliche Lernzuwächse erreichen können? Die Beiträge in diesem Symposium nehmen sich derartiger Fragen an, indem Wirkungen der Zusammensetzung von Schüler/innen aus unterschiedlichen Perspektiven hinsichtlich unterschiedlicher Outcomes analysiert werden.

Beitrag 1 richtet den Blick, basierend auf für die USA national repräsentativen Early Childhood Longitudinal Study 2010-11, auf den Effekt der sozialen und ethnischen Komposition auf die Leistungsentwicklung im Lesen und Mathematik während der Vorschule, während der Sommerferien zwischen Vorschule und 1. Jahrgangsstufe, während des ersten Schuljahres, während des Sommers zwischen 1. und 2. Jahrgangsstufe sowie während des zweiten Schuljahres. Effekte der Schul- und Klassenzusammensetzung in der Primar- und Sekundarstufe I auf die Mathematikleistung von Schüler/innen in Österreich stehen im Mittelpunkt des zweiten Beitrags. Sowohl für die Primarstufe als auch die Sekundarstufe I zeigen sich substantielle, jedoch schwache Effekte der sozialen Zusammensetzung der Klassen auf die Mathematikleistungen der Schüler/innen. Es zeigt sich im Ergebnis, dass die auf den ersten Blick sehr starken sozialen Klassenkompositionseffekte in der Sekundarstufe I Unterschiede zwischen den bestehenden Schultypen abbilden, bzw. sie für die Hauptschule durch die (sozialeselektive) Zusammensetzung der Schüler/innen nach ihrem Leistungspotential in Lerngruppen zu erklären sind. Im dritten Beitrag betrachten die Autoren, basierend auf einer österreichischen Datengrundlage schulisches Problemverhalten als Mediator herkunfts- und leistungsbezogener Klassenkompositionseffekte auf die schulische Performanz. Der Beitrag zeigt auf Basis von Längsschnittdaten, dass bei simultaner Betrachtung von drei Klassenkompositionsmerkmalen (Leistung, Sozialstatus, Migrationshintergrund) nur die leistungsbezogene Komposition Effekte auf die Schüler/innenleistungen in Deutsch, Mathematik und Englisch aufweist. Ferner zeigen die Analysen, dass der Einfluss der Leistungskomposition größtenteils durch das Ausmaß des unterrichtsstörenden Verhaltens erklärt werden kann. Der vierte Beitrag fokussiert auf die Frage, inwieweit sich die Entwicklung individueller Verhaltensprobleme von Jugendlichen durch das Niveau an Verhaltensproblemen unter den Mitschüler/innen erklären lässt. Der Beitrag präsentiert eine Gesamtsicht auf die Ergebnisse der Freiburger Längsschnittstudie zum Peereinfluss in Schulen. Die Befunde lassen erkennen, dass die individuelle Entwicklung aggressiven, delinquenten und unterrichtsstörenden Verhaltens durch die Ausprägung dieses Verhaltens unter den Klassenkameradinnen und Klassenkameraden beeinflusst wird.

Im Rahmen des Symposiums werden wichtige Fragen bezüglich des Einflusses von Kompositionsmerkmalen im schulischen Kontext auf das Leistungsniveau von Schüler/innen sowie weiterer Outcomes beantwortet. Die Ergebnisse aller Beiträge werden abschließend in Hinblick auf mögliche schulpraktische Implikationen reflektiert. Die einzelnen Beiträge werden von Eckhard Klieme als ausgewiesenen Experten auf dem Gebiet der Schul- und Kompositionseffekte diskutiert.

#### *Beiträge des Symposiums*

### **Schulkompositionseffekte auf die Leistungsentwicklung in US-amerikanischen Grundschulen: Befunde aus der Early Childhood Longitudinal Study 2010-11**

**Hanna Dumont<sup>1</sup>, Douglas Ready<sup>2</sup>**

<sup>1</sup>DIPF Berlin, <sup>2</sup>Columbia University

#### Theoretischer Hintergrund

Die Untersuchung der Zusammensetzung der Schülerschaft von Schulen und die Frage, wie dies die Leistungsentwicklung von Schülerinnen und Schülern beeinflusst, hat in der Bildungsforschung eine lange Tradition (Coleman, 1966). Inzwischen liegen zahlreiche nationale und internationale Studien vor, die zeigen konnten, dass eine sozial privilegierte Schülerschaft vorteilhaft für die Leistungsentwicklung von Schülerinnen und Schülern ist (im Überblick siehe Dumont, Neumann, Maaz & Trautwein, 2013; Van Ewijk & Sleegers, 2010a). Für die ethnische Zusammensetzung ist die Befundlage hingegen weniger eindeutig. In deutschen Studien konnte bislang kein Hinweis darauf gefunden werden, dass der Anteil von Schülerinnen und Schülern mit Migrationshintergrund an Schulen über die soziale Zusammensetzung hinaus einen Einfluss auf die Leistungsentwicklung hat (Bellin, 2009; Gröhlich, Guill, Scharenberg, & Bos; Stanat, Schwippert, & Gröhlich, 2010). In zwei internationalen Metaanalysen konnte jedoch ein spezifischer Effekt der ethnischen Zusammensetzung gefunden werden (Mickelson, Bottia & Lambert, 2013; Van Ewijk & Sleegers, 2010b). Insbesondere in den USA wurden negative Effekte der ethnischen Zusammensetzung für bestimmte ethnische Minderheiten, nämlich des Anteils von afro- und hispanoamerikanischen Schülerinnen und Schülern pro Schule, berichtet (Benson & Borman, 2010; Ready & Silander, 2011).

#### Fragestellungen

Die Schulkomposition ist in den USA in hohem Maße mit der Komposition bzw. hohen Segregation in Schuleinzugsgebieten und damit sogenannten Nachbarschaftseffekten konfundiert (Sampson, Morenoff & Gannon-Rowley, 2002). Neben der simultanen Schätzung von Schulkompositions- und Nachbarschaftseffekten ist der Vergleich von Schulkompositionseffekten während des Schuljahres und während der Sommerferien eine weitere Möglichkeit, den Effekt der Schulkomposition zu isolieren (Ready & Silander, 2011). Vor diesem Hintergrund analysieren wir in der vorliegenden Studie den Effekt der sozialen und ethnischen Komposition auf die Leistungsentwicklung im Lesen und Mathematik während der Vorschule, während der Sommerferien zwischen Vorschule und 1. Jahrgangsstufe, während des ersten Schuljahres, während des Sommers zwischen 1. und 2. Jahrgangsstufe sowie während des zweiten Schuljahres.

#### Methode

Als Datengrundlage dienen Daten der ersten sechs Erhebungszeitpunkte der (für die USA national repräsentativen) Early Childhood Longitudinal Study 2010-11, die jeweils im Herbst und im Frühjahr der Vorschule, der ersten und der zweiten Jahrgangsstufe stattfanden. Insgesamt gingen 16182 Schülerinnen und Schüler aus 888 Schulen in die Analysen ein. Fehlende Werte wurden mit dem R Paket MICE (van Buuren & Groothuis-Oudshoorn, 2011) imputiert. Zur Analyse der Fragestellungen wurden Mehrebenenanalysen in HLM für die fünf betrachteten Zeiträume spezifiziert. Als abhängige Variablen fungierten die Lese- und Mathekompetenz der Schülerinnen und Schüler am Ende des jeweiligen Zeitraums, wobei immer für die Kompetenz zu Beginn des Zeitraums kontrolliert wurde. Die soziale Komposition wurde über den mittleren sozioökonomischen Status der Schule, die ethnische Komposition über den Anteil von afro- und hispanoamerikanischen Schülerinnen und Schülern pro Schule operationalisiert. In allen Modellen wurden für weitere Personenmerkmale kontrolliert.

#### Ergebnisse & Diskussion

Es zeigte sich sowohl für die Lese- als auch die Mathematikkompetenz, dass der mittlere sozioökonomische Status der Schule in keinem Zusammenhang mit der Leistungsentwicklung von Schülerin-nen und Schülern stand. Für den Anteil von afro- und hispanoamerikanischen Schülerinnen und Schüler pro Schule waren die Befunde heterogener. Eine höhere Konzentration dieser ethnischen Minderheit hatte negative Effekte auf die Leistungsentwicklung im Lesen während der drei betrachteten Schuljahre, die bis auf den Effekt in der zweiten Jahrgangsstufe auch unter Kontrolle der sozialen Komposition signifikant blieben. In Mathematik zeigte sich der negative Effekt der ethnischen Komposition lediglich im ersten Schuljahr, blieb jedoch auch hier unter Kontrolle der sozialen Komposition signifikant. Gleichzeitig konnte für die Lesekompetenz ein positiver Effekt der ethnischen Komposition während des Sommers zwischen Vorschule und erstem Schuljahr und für die Mathematikkompetenz nur ein geringer negativer Effekt für den Sommer vom ersten zum zweiten Schuljahr festgestellt werden, was Hinweise auf die Bedeutsamkeit der Schulkomposition im Gegensatz zur Komposition des Schuleinzugsgebiets liefert.

### **Effekte der Schul- und Klassenzusammensetzung in der Primar- und Sekundarstufe I auf die Mathematikleistung von Schüler/innen**

**Horst Biedermann<sup>1</sup>, Arvid Nagel<sup>2</sup>, Christoph Weber<sup>3</sup>**

<sup>1</sup>Pädagogische Hochschule St.Gallen, <sup>2</sup>Institut Professionsforschung und Kompetenzentwicklung; Pädagogische Hochschule St.Gallen, <sup>3</sup>Pädagogische Hochschule OÖ

Auf die Mitschüler/innen kommt es an? Diese Frage hat insbesondere seit der Fokussierung auf internationale und nationale Vergleichsstudien (Large-Scale-Assessments) wiederum unverkennbar an Interesse im Bereich der empirischen Bildungsforschung gewonnen. Dabei interessiert vor allem, inwieweit sich die Zusammensetzung der Schüler/innen (was mit Komposition zum Ausdruck gebracht wird) auf deren schulischen Leistungen auswirkt und wie dies erklärt werden kann. Während die Effekte unterschiedlicher Kompositionsmerkmale (so z.B. der sozio-ökonomische Hintergrund, der kulturelle Hintergrund oder das schulische Fähigkeitsniveau) der Schul- und Klassenebene auf Schüler/innenleistungen bei getrennter Betrachtung (d.h. entweder Schul- oder Klassenebene) empirisch weitgehend gut dokumentiert sind (vgl. u.a. Agirdag, van Avermaet & van Houtte, 2013; Baumert, Stanat & Watermann, 2006; Bellin, 2008; Dumont, Neumann, Maaz & Trautwein, 2013; Harker & Tymms, 2004; Rjosk, Richter, Hochweber, Lüdtke & Klieme, 2014), ist die empirische Befundlage in Hinblick auf Kompositionseffekte bei simultaner Berücksichtigung von Schul- und Klassenebene (Drei-Ebenen-Modelle) gegenwärtig noch dünn.

In diesem Beitrag wird aufbauend hierauf der Fragestellung nachgegangen: Lassen sich Effekte der Schul- und Klassenkompositionen auf die Mathematikleistung von Schüler/innen identifizieren? Konkret interessiert, welche Schul- und Klassenkompositionsmerkmale (Anteil Schüler/innen mit nichtdeutscher Alltagssprache, durchschnittlicher Sozialstatus) bei simultaner Betrachtung einen Effekt auf die Leistung in Mathematik aufweisen. Zur Beantwortung der Frage werden die österreichischen Daten der Bildungsstandarderhebung in Mathematik, 4. Schulstufe (BIST-M4; N = 73.655 Schüler/innen aus 4.920 Klassen und 3.050 Schulen; vgl. Schreiner & Breit, 2014) und Mathematik, 8. Schulstufe (BIST-M8; N = 79.678 Schüler/innen aus 4.074 Klassen und 1.416 Schulen; vgl. Schreiner & Breit, 2012) herangezogen. Die statistischen Analysen (Mehrebenenmodelle) wurden mit der Software Mplus (vgl. Muthén & Muthén, 2012) durchgeführt.

Sowohl für die Primarstufe als auch die Sekundarstufe I zeigen sich substanzielle, jedoch schwache Effekte der sozialen Zusammensetzung der Klassen auf die Mathematikleistungen der Schüler/innen. Die auf den ersten Blick sehr starken sozialen Klassenkompositionseffekte in der Sekundarstufe I bilden Unterschiede zwischen den bestehenden Schultypen (Allgemeinbildende Höhere Schulen vs. Hauptschulen/Neuen Mittelschulen) ab, bzw. sind für die Hauptschule durch die (sozialselektive) Zusammensetzung der Schüler/innen nach ihrem Leistungspotential in Lerngruppen zu erklären. Auf Schulebene besteht sowohl in der Volksschule, der Hauptschule und der Allgemeinbildenden Höheren Schule ein substanzieller, aber schwacher negativer Kompositionseffekt des Anteils der Schüler/innen mit nichtdeutscher Familiensprache auf die Mathematikleistung.

Im Rahmen des Beitrags werden die theoretische Grundlagen und empirischen Analysen dargestellt und die Ergebnisse bezüglich möglicher bildungspolitischer Konsequenzen reflektiert und diskutiert.

### **Schulisches Problemverhalten als Mediator herkunfts- und leistungsbezogener Klassenkompositionseffekte auf die schulische Performanz**

**Christoph Weber<sup>1</sup>, Heinz Leitgöb<sup>2</sup>, Johann Bacher<sup>3</sup>**

<sup>1</sup>Pädagogische Hochschule OÖ, <sup>2</sup>Universität Eichstätt-Ingolstadt, <sup>3</sup>Johannes Kepler Universität Linz

In der Literatur herrscht weitgehender Konsens, dass neben Individualmerkmalen wie z.B. dem kognitiven Fähigkeitsniveau, elterlichen Sozialstatus und Migrationshintergrund auch die Schul- bzw. Klassenzusammensetzung nach diesen Merkmalen die Leistungen von Schüler/innen beeinflusst (im Überblick Dumont et al., 2013). Demgegenüber ist jedoch die empirische Befundlage zur Erklärung solcher Kompositionseffekte noch dünn. Als mögliche Mediatoren für Kompositionseffekte werden u.a. Unterrichtsmerkmale (z.B. weniger anspruchsvoller Unterricht bei geringem Leistungsniveau), die normative Kultur der Elternschaft (z.B. geringes Engagement der Eltern bei ungünstigen Zusammensetzungen) oder auch Aspekte der normativen Kultur der Schülerschaft bzw. Interaktions- und Vergleichsprozesse innerhalb der Schülergruppen (zusammenfassend als Peer-Effekte bezeichnet) genannt (Ditton, 2013; Harker & Tymms, 2004). Bei Peer-Effekten wird in der Regel der Fokus auf „Ansteckungsprozesse“ (hohe Bildungsaspirationen in Klassen/Schulen schafft leistungsförderliches Klima) und soziale Lernprozesse gelegt.

Der Beitrag schließt an die Annahme von Peer-Mediatoreffekten an. Es wird argumentiert, dass die Merkmale elterlichen Sozialstatus und Fähigkeitsniveau mit schulischem Problemverhalten (z.B. zu spät kommen, den Unterricht stören) korrelieren (Hinshaw, 1992; Silver et al., 2005; Bradley & Corwyn, 2002) und somit Unterschiede in der sozialen und leistungsbezogenen

Zusammensetzung von Klassen auch mit Unterschieden im Ausmaß des schulischen Problemverhaltens einhergehen (Demanet & van Houtte, 2014). Ein vermehrtes Auftreten von schulischem Problemverhalten stört Lernprozesse und ist somit unmittelbar relevant für die kollektive Leistungsentwicklung. Darüber hinaus stellt Problemverhalten auch auf der Individualebene einen Leistungsprädiktor dar (Zimmermann et al., 2013).

Vor diesem Hintergrund geht der Beitrag (1) der Frage nach, welche Klassenkompositionsmerkmale (durchschnittliches Fähigkeitsniveau, durchschnittlicher Sozialstatus und Anteil Schüler/innen mit nichtdeutscher Alltagssprache) bei simultaner Betrachtung einen Effekt auf die Schülerleistungen haben. Anschließend (2) wird untersucht, ob potentielle Kompositionseffekte (nach Kontrolle des individuellen Problemverhaltens) durch das Ausmaß an Problemverhalten in der Klasse erklärt werden können.

Zur Beantwortung der Fragestellungen werden Daten der Evaluation der Österreichischen Neuen Mittelschule (NMS; Eder et al., 2015) herangezogen. Es liegen für zwei NMS-Kohorten (Kohorte 1: 3362 Schüler/innen aus 170 Klassen aus 67 Schulen; Kohorte 2: 6022 Schüler/innen aus 301 Klassen aus 105 Schulen) neben soziodemografischen Variablen u.a. auch Messungen des kognitiven Fähigkeitsniveaus (KFT; Heller & Perleth, 2000; 5. und 8. Schulstufe), der Leistungen (8. Schulstufe) in Deutsch und Englisch (jeweils Lesen) und Mathematik und des schulischen Problemverhaltens vor. Für die Analysen werden Zweiebenenmodelle und Zweiebenenmediationsmodelle unter Verwendung des doubly latent-Ansatzes (Lüdtke et al., 2011; Nagengast & Marsh, 2012) herangezogen.

Für beide Kohorten ergeben sich signifikante Effekte des durchschnittlichen Fähigkeitsniveaus (KFT) auf die Leistungen (Effektstärken zwischen 0,17 und 0,35). Für Englisch zeigt sich in beiden Kohorten ebenfalls ein schwacher signifikanter Effekt des durchschnittlichen Sozialstatus (Effektstärke jeweils 0,21). Effekte des Anteils der Schüler/innen mit nichtdeutscher Alltagssprache zeigen sich nicht. Die Mehrebenenmediationsmodelle bestätigen weitgehend die Annahme, dass in Klassen mit einem höheren Fähigkeitsniveau weniger Problemverhalten auftritt und ein geringeres Problemverhalten mit besseren Leistungen einhergeht. In beiden Kohorten fallen nach Berücksichtigung des Problemverhaltens die direkten KFT-Kompositionseffekte nicht mehr signifikant aus. Zusätzlich lassen sich signifikante bzw. marginal signifikante indirekte Kompositionseffekte des KFT über das Problemverhalten auf die Leistungen zeigen. Die direkten Effekte des durchschnittlichen Sozialstatus auf die Englischleistungen bleiben im Mediationsmodell jedoch bestehen.

### **Klassenkompositionseffekte auf Verhaltensprobleme – Eine Gesamtsicht auf die Ergebnisse der Freiburger Studie zum Peereinfluss in Schulen (FRI-PEERS)**

**Christoph Michael Müller**  
Universität Fribourg

Klassenkompositionseffekte wurden in den Bildungswissenschaften bisher vorwiegend mit Bezug zu der Entwicklung von Schulleistung und akademischem Selbstkonzept untersucht. Die Peereinflussforschung deutet jedoch darauf hin, dass Kinder und Jugendliche auch im Bereich ihrer sozialen Entwicklung von den Gleichaltrigen beeinflusst werden. Die Klassenkameradinnen und Klassenkameraden bilden dabei eine „unfreiwillige“, durch die Schule von außen zusammengestellte Peergruppe. Auf der Basis verschiedener Peereinflusstheorien (z.B. Akers, 2009; Bandura, 2011) kann erwartet werden, dass Schülerinnen und Schüler in ihrer eigenen Verhaltensentwicklung durch das Niveau an Verhaltensproblemen in ihrer Schulklasse beeinflusst werden. Zu klären gilt es dabei auch, durch welche Mechanismen sich solcher Peereinfluss vollzieht und von welchen Faktoren er abhängt. Mit einem Fokus auf aggressive und delinquente Verhaltensweisen unter Jugendlichen wird diesen Fragen in der Freiburger Studie zum Peereinfluss in Schulen (FRI-PEERS) seit 2011 nachgegangen. Basierend auf den bisher publizierten Originalbeiträgen ist nun eine Gesamtsicht auf die Ergebnisse möglich, die Perspektiven für weitere Forschungsarbeiten eröffnet.

FRI-PEERS ist eine Längsschnittstudie, bei der über sechs Messzeitpunkte (7. Klasse: T1-T4; 8. Klasse: T5; 9. Klasse: T6) die Entwicklung von 864 Jugendlichen auf der Sekundarstufe I verfolgt wurde (Alter zu T1: M=13.12 Jahre; SD=0.48). An der Studie nahmen Jugendliche aller 56 Schulklassen aus allen acht Sekundarschulhäusern einer Schweizer Region fast vollständig teil. Die Jugendlichen besuchten Klassen des Progymnasiums, der Sekundarschule, der Realschule oder sonderpädagogische Förderklassen, die in gemeinsamen Schulhäusern verortet waren. Die Erhebungen erfolgten größtenteils durch anonyme Selbst- und Peerauskünfte der Jugendlichen. Alle verwendeten Skalen waren standardisiert und genügten den psychometrischen Gütekriterien. Die Verhaltenscharakteristika der Schulklassen wurden i.d.R. als Klassenmittelwerte der individuellen Verhaltensberichte der Schülerinnen und Schüler operationalisiert, welche als Indikatoren für die vorherrschende deskriptive Gruppennorm in einer Klasse gelten (Henry et al., 2000). Zur Bestimmung von Klassenkompositionseffekten wurden etablierte Methoden der Klassenkompositions- und Peereinflussforschung eingesetzt (Übersicht s. Müller & Zurbriggen, 2016) und mehrebenenanalytisch vorgegangen. Die Ergebnisse zeigen eine moderate Zunahme von Verhaltensproblemen auf der Sekundarstufe I, wobei im Bildungsgang mit Grundansprüchen (Realschule) in der 7. Klasse ein überproportionaler Anstieg zu beobachten war (Müller & Hofmann, 2016). In Bezug auf Klassenkompositionseffekte wurde deutlich, dass die zukünftige Ausprägung aggressiven, delinquenten und unterrichtsstörenden Individualverhaltens durch das mittlere Niveau des jeweiligen Verhaltens unter den Klassenkameradinnen und Klassenkameraden zu einem früheren Zeitpunkt beeinflusst wird (unter statistischer Berücksichtigung des früheren Individualverhaltens; Müller et al., 2015; 2016a; early online; submitted). Dies bedeutet, dass ein hohes mittleres Niveau an Verhaltensproblemen in der Klasse zu einer Erhöhung der Auftretenswahrscheinlichkeit zukünftiger, individueller Verhaltensprobleme beiträgt. Der Einfluss der Verhaltenscharakteristika der Peers scheint zwar weniger bedeutsam zu sein als proximale, individuelle Faktoren (z.B. Impulsivität), erwies sich aber auch unter Kontrolle anderer relevanter Faktoren (z.B. Geschlecht, Impulsivität, sozioökonomischer Status, Herkunft, Bildungsgang, Schulnoten, Elternunterstützung, Verhaltensproblemen der Freunde von außerhalb der Klasse; Müller et al., 2016a; early online) als signifikant. Neben dem Klassenmittelwert hatte auch das wahrgenommene Verhalten unter den persönlichen Freundinnen und Freunden in der Klasse sowie unter den sozial dominanten Jugendlichen signifikanten Einfluss auf die individuelle Verhaltensentwicklung (2016b). Mediationsanalysen zeigten, dass die Effekte des Klassenverhaltens über die individuelle Wahrnehmung des Peerverhaltens sowie eigene Einstellungsveränderungen gegenüber Verhaltensproblemen vermittelt wurden (Müller et al., 2015; Hofmann & Müller, submitted). Jugendliche mit folgenden Charakteristika erwiesen sich im Bereich Delinquenz als besonders beeinflussbar durch ihre Peers: Männliches Geschlecht, hohe Impulsivität, hohe Risikobereitschaft, hohe soziale Dominanz und Popularität, hoher Anteil unstrukturierter Freizeitaktivitäten und wenig Elternaufsicht (Müller et al., early online). Als interessant und unterstützend wahrgenommener Unterricht der Klassenlehrperson stellte einen protektiven Faktor gegenüber negativem Peereinfluss dar (Müller et al., submitted). Die Gesamtergebnisse der Studie werden vor dem Hintergrund der Bedeutung des Schulkontexts für die Verhaltenssozialisation von Jugendlichen diskutiert. Implikationen für die zukünftige Erforschung von Verhaltensproblemen in der Schule werden aufgezeigt.



## D21

Zeit: Freitag, 16.02.2018: 11:15 - 13:00 · Ort: Bernoullianum 30, H148

### Professioneller Umgang mit Heterogenität und Inklusion: Herausforderungen für die Bildungsforschung

Chair(s): **Johannes König** (Universität zu Köln, Deutschland), **Cornelia Gräsel** (Bergische Universität Wuppertal)

DiskutantIn(nen): **Gabriele Kaiser** (Universität Hamburg/ Australian Catholic University, Brisbane (Australia))

Der Umgang mit Heterogenität und Inklusion in Schule und Unterricht gehört zu den Kernanforderungen professionellen Handelns. Zugleich gilt die Erforschung des Umgangs mit Heterogenität und Inklusion zu einer der aktuellen Herausforderungen der empirischen Bildungsforschung. Dabei stellen sich die zu bearbeitenden Forschungsfragen auf unterschiedlichen Ebenen und Funktionsfeldern von Schule und Unterricht (Gräsel, Decristan & König, 2017): Auf Ebene des Unterrichts werden mit verschiedenen Ansätzen beispielsweise die Heterogenität von Lerngruppen bzw. Maßnahmen des Umgangs mit heterogenen Lerngruppen im Unterricht hinsichtlich möglicher Einflüsse auf Lernfortschritte der Lernenden analysiert; auf Ebene der Einzelschule stehen stärker Maßnahmen der Organisation und Kooperation zur Bewältigung professioneller Anforderungen im Bereich Heterogenität und Inklusion im Mittelpunkt; auf Ebene der Lehrpersonen fragen zunehmend mehr Arbeiten nach den personalen Voraussetzungen, die benötigt werden, um professionelle Anforderungen im Bereich schulischer Heterogenität und Inklusion erfolgreich zu bewältigen; nicht zuletzt setzen sich verschiedene Arbeiten mit der Beschreibung und Analyse von Heterogenität und Inklusion bei Lernenden auseinander.

Das Symposium „Professioneller Umgang mit Heterogenität und Inklusion: Herausforderungen für die Bildungsforschung“ zielt darauf, diese Breite an Ansätzen und Forschungszugängen mithilfe von vier Vorträgen exemplarisch abzubilden. Dies geschieht explizit aus interdisziplinärer Perspektive: Es beteiligen sich Fachvertreterinnen und Fachvertreter verschiedener Disziplinen, u.a. der Erziehungswissenschaft, der Fachdidaktik, der Psychologie und der Sonderpädagogik. Damit sollen auch Berührungspunkte der Arbeiten sichtbar werden und mögliche Anschlussstellen für die zukünftige interdisziplinäre Bildungsforschung zur Diskussion gestellt werden.

Der erste Beitrag von Jasmin Decristan et al. verortet sich auf der Unterrichtsebene und analysiert die individuelle Gesprächsbeteiligung von Schülerinnen und Schülern in heterogenen Lerngruppen. Der zweite Beitrag von Lisa Hoffmann et al. analysiert die Sicht von Schülerinnen und Schülern auf die interprofessionelle Lehrerkooperation in inklusiven Schulsettings mit Blick auf mögliche Auswirkungen des Wohlbefindens und der sozialen Integration der Lernenden. Im dritten Beitrag von Kristina Gerhard et al. wird nach den kognitiven Voraussetzungen von Lehrkräften gefragt, die diese für einen erfolgreichen Umgang mit heterogener Schülerschaft benötigen, und es wird ein neu entwickelter Test zum pädagogischen Lehrerwissen für inklusiven Unterricht präsentiert. Der vierte Beitrag von Michael Grosche et al. setzt sich mit schulischer Heterogenität und Inklusion auf einer begrifflichen Ebene auseinander und stellt Ergebnisse von Experteninterviews mit Inklusionsforscher\*innen zur Suche nach einem gemeinsamen Verständnis vor. Die vier Beiträge werden abschließend von Gabriele Kaiser diskutiert, unter besonderer Berücksichtigung der fachdidaktischen Perspektive in der empirischen Bildungsforschung.

#### Beiträge des Symposiums

### Individuelle Beteiligung am Klassengespräch in heterogenen Lerngruppen: Wer beteiligt sich, wen beteiligt die Lehrkraft?

**Jasmin Decristan<sup>1</sup>, Benjamin Fauth<sup>2</sup>, Eva Heide<sup>3</sup>, Csaba Kurucz<sup>4</sup>, Franziska Locher<sup>5</sup>, Bianka Troll<sup>6</sup>, Mareike Kunter<sup>7</sup>, Gerhart Büttner<sup>7</sup>, Eckhard Klieme<sup>8</sup>, Ilonca Hardy<sup>7</sup>**

<sup>1</sup>Bergische Universität Wuppertal; <sup>2</sup>IDeA-Forschungszentrum, <sup>3</sup>Universität Tübingen; <sup>4</sup>IDeA-Forschungszentrum, <sup>5</sup>Goethe-Universität Frankfurt am Main, <sup>6</sup>Freie Universität Berlin, <sup>7</sup>Universität Bamberg, <sup>8</sup>Leuphana Universität Lüneburg, <sup>7</sup>Goethe-Universität Frankfurt am Main; <sup>8</sup>IDeA-Forschungszentrum, <sup>8</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt am Main; <sup>8</sup>IDeA-Forschungszentrum

#### Hintergrund

Der Umgang mit Heterogenität ist wesentlicher Bestandteil eines professionellen Handelns von Lehrkräften (KMK-Standards für die Lehrerbildung, 2014). Im Fokus steht die Frage, wie Lehrkräfte in ihrem Unterricht auf die unterschiedlichen Voraussetzungen der Lernenden eingehen können. Entsprechend dem Angebot-Nutzungs-Modell (Helmke, 2010) lässt sich annehmen, dass hierfür sowohl das Unterrichtsangebot der Lehrkräfte als auch die individuelle Nutzung des Angebots durch die Lernenden entscheidend sind. Als ein Nutzungsindikator lässt sich die Beteiligung am Unterrichtsgespräch heranziehen. Eine aktive Beteiligung an ko-konstruktiven Lernprozessen stellt eine wesentliche Grundlage eines sozial-konstruktivistischen Lehr-Lernverständnisses dar (z.B. Vygotsky, 1978).

Empirische Studien liefern erste Befunde dazu, dass eine Beteiligung am Unterrichtsgespräch mit höherer Motivation und größerem Kompetenzerleben und teils auch besseren Leistungen einhergeht (Carstens et al., 2016; Flieller, Jarlégan & Tazouti, 2016; Ing et al., 2016; Pauli & Lipowsky, 2007). Zugleich zeigt sich, dass vor allem leistungsstärkere Schülerinnen und Schüler am Unterrichtsgespräch beteiligt sind (z.B. Black, 2004; Carstens et al., 2016; Jurik, Gröschner & Seidel, 2013).

Diesem Befund können zwei unterschiedliche Mechanismen zugrunde liegen: (1) vor allem leistungsstärkere beteiligen sich und/oder (2) Lehrkräfte nehmen vor allem leistungsstärkere Schülerinnen und Schüler dran. Wenngleich diese unterschiedlichen Selektionsmechanismen mit unterschiedlichen Implikationen zum Umgang mit Heterogenität einhergehen, gibt es bislang kaum empirische Forschung dazu.

#### Fragestellungen

1. Sind leistungsstärkere Kinder eher am Unterrichtsgespräch beteiligt?
2. Beteiligen sich vor allem Leistungsstärkere am Unterrichtsgespräch oder nimmt die Lehrkraft vor allem Leistungsstärkere dran?

#### Methode

Die Daten stammen aus einer Längsschnittstudie im naturwissenschaftlichen Sachunterricht an Grundschulen, die zwei Unterrichtseinheiten (à 4.5 Doppelstunden) mit konstantem Stundenaufbau (Einstieg, Experimentierphase, Unterrichtsgespräch, Arbeitsblattphase, gemeinsame Reflexion) umfasste. Die vorgenommenen Videokodierungen beziehen sich auf eine max. zehnminütige Phase (M = 8.48 min) des Unterrichtsgesprächs nach dem Experiment.

Die 35 Lehrkräfte (89% weiblich) hatten eine mittlere Berufserfahrung von 16 Jahren, die 628 Kinder (48% weiblich) waren im Mittel 9 Jahre alt.

Die zu Beginn der Studie erhobenen Lernvoraussetzungen umfassten kognitive Grundfähigkeiten ( $\alpha = .72$ ), Leseverständnis ( $\alpha > .91$ ), Sprachverständnis ( $\alpha = .72$ ), naturwissenschaftliche Kompetenz (EAP/PV Rel. = .70) und Vorwissen im Fachinhalt (EAP/PV Rel. = .52) sowie Migrationshintergrund (39% gaben an, dass mindestens ein Elternteil nicht in Deutschland geboren wurde) und Geschlecht.

#### Ergebnisse Forschungsfrage 1

Immer, wenn eine Schülerin/ein Schüler einen mündlichen Beitrag am inhaltsbezogenen Unterrichtsgespräch leistete wurde eine Beteiligung (ein Fall) kodiert (N = 896; Beurteiler-Übereinstimmung: 80%). Diese Fälle verteilten sich auf 338 (54%) der 628 Kinder. Die Beteiligung am Unterrichtsgespräch korrelierte signifikant mit Vorwissen ( $r = .13$ ), Lese- ( $r = .08$ ) und Sprachverständnis ( $r = .09$ ) sowie Migrationshintergrund ( $r = .15$ ), jedoch nicht mit kognitiven Grundfähigkeiten, naturwissenschaftlicher Kompetenz oder Geschlecht. In einer multiplen Regressionsanalyse mit Standardfehlerkorrektur und MLR-Schätzer in MPlus (Muthén & Muthén, 2012–2017) erwiesen sich Vorwissen ( $b = .10$ ) und Migrationshintergrund ( $b = .09$ ) als eigenständige signifikante Prädiktoren.

#### Ergebnisse Forschungsfrage 2

Um schüler- und lehrergesteuerte Selektionsmechanismen voneinander zu trennen, wurden zusätzlich alle Meldungen der Kinder während der 10-minütigen Sequenz kodiert (N = 1.941; Übereinstimmung: 78%). Hieraus resultierten drei Vergleichsgruppen: Lernende, die beteiligt waren (G1, n=338, 54%), Lernende, die sich ein- oder mehrfach meldeten, aber nicht drangenommen wurden (G2, n=109, 17%) und Kinder ganz ohne Beteiligung (G3, n=182, 29%). Die Ergebnisse multipler logistischer Regressionsanalysen deuten auf einen zweifachen Selektionseffekt hin. Vor allem Lernende mit höherem Vorwissen ( $b = .12$ ) und ohne Migrationshintergrund ( $b = .09$ ) beteiligten sich (schülergesteuerte Selektion; Vergleich G3 vs. G1+G2) und von denen, die sich meldeten, nahm die Lehrkraft Lernende mit höherem Vorwissen ( $b = .13$ ) und Sprachverständnis ( $b = .10$ ) dran (lehrergesteuerte Selektion, Vergleich G1 vs. G2).

Vertiefende Analysen müssen nun der Frage nachgehen, inwieweit diese differenziellen Interaktionen einen adaptiven Umgang mit Heterogenität darstellen und somit auch, inwiefern hieraus unterschiedliche kognitive und motivationale Lernergebnisse resultieren.

### Die Sicht von Schülerinnen und Schülern auf die interprofessionelle Lehrkooperation in inklusiven Schulsettings

Lisa Hoffmann<sup>1</sup>, Mike Trauntschnig<sup>2</sup>, Susanne Schwab<sup>3</sup>

<sup>1</sup>Bergisch Universität Wuppertal, <sup>2</sup>Karl-Franzens-Universität Graz (Österreich), <sup>3</sup>Bergische Universität Wuppertal/ North-West University, Vanderbijlpark (Südafrika)

Die Zusammenarbeit zwischen allgemeinbildenden und sonderpädagogischen Lehrkräften hat sich als wesentliche Bedingung für das Gelingen von schulischer Inklusion herausgestellt (z. B. Urban & Lütje-Klose, 2014), um den heterogenen Lern-, Entwicklungs- und Förderbedürfnissen von Kindern und Jugendlichen mit und ohne sonderpädagogischen Förderbedarf (SPF) angemessen begegnen zu können. Bisherige Studien mit dem Fokus auf der Erforschung von interprofessioneller Lehrkooperation beruhen dabei fast ausschließlich auf der Sichtweise von Lehrkräften (z.B. Hellmich, Hoya, Görel & Schwab, 2017). Die vorliegende Studie nimmt jedoch die Sichtweise von Schülerinnen und Schülern in den Blick, um eine Verzerrung durch sozial erwünschte Antworten von Lehrkräften zu vermeiden. Es wird unter anderem der Fragestellung nachgegangen, inwiefern Schülerinnen und Schüler das gemeinsame Unterrichten als positiv wahrnehmen und inwieweit sie sich unterstützt fühlen, wenn zwei Lehrkräfte gemeinsam lehren. Zusätzlich wird fokussiert, ob Lehrkooperation und die Zusammenarbeit von Schülerinnen und Schülern miteinander zusammenhängen. Weiterhin wurde untersucht, ob es einen Zusammenhang zwischen Lehrkraftkooperation und den Qualitätsindikatoren (soziale Inklusion, emotionale Inklusion und akademisches Selbstkonzept) von Inklusion gibt. Die Datenerhebung für die vorliegende Studie erfolgte im Rahmen der Längsschnittstudie ATIS-STEP (Attitudes Towards Inclusive Schooling – Students, TEachers and Parents) am ersten Messzeitpunkt (Herbst 2016; Beginn des vierten Schuljahres). Insgesamt haben 721 Schülerinnen und Schüler (357 Schülerinnen und 364 Schüler) der vierten Schulstufe aus 48 inklusiven Klassen aus der Steiermark (Österreich) und 46 Lehrkräfte (42 Klassenlehrerinnen und vier Klassenlehrer) teilgenommen. 92 der befragten Schülerinnen und Schüler wiesen nach Auskunft der Lehrkraft einen SPF auf. Nach Angaben der Lehrkräfte variiert die Anzahl der Stunden im Team-Teaching zwischen 0 und 23 Stunden pro Woche. Im Rahmen der Papier-Bleistiftenerhebung sollten alle Schülerinnen und Schüler auf einer vierstufigen Likert-Skala („stimmt gar nicht“ bis „stimmt genau“) angeben, ob in der eigenen Klasse zwei Lehrkräfte im Team unterrichten. Ergänzend wurde gefragt, inwieweit Schülerinnen und Schüler es mögen, wenn die Lehrkräfte im Team unterrichten („Ich mag es, wenn zwei Lehrerinnen/Lehrer gemeinsam unterrichten.“) und ob sie sich gut unterstützt fühlen, wenn zwei Lehrkräfte gemeinsam unterrichten. Für die Erfassung der Kooperationsformen wurden die Lehrkräfte gebeten auf einer fünfstufigen Ratingskala (von „nie“ bis „immer“) anzugeben, wie oft folgende Formen von Team-Teaching eingesetzt wurden: Lehrkraft und Beobachtungsperson, Niveaudifferenzierter Unterricht, Assistenzunterricht, Gemeinsamer Unterricht und Individuelle Begleitung. Vorab konnte die vorliegende Studie in Bezug auf die Kooperationsformen verschiedene andere Arbeiten (z.B. Pancsofar & Petroff, 2016) bestätigen und zeigt, dass Teamarbeit vergleichsweise selten Co-Teaching ist, sondern häufiger in den Formen Assistenzunterricht oder niveaudifferenzierter Unterricht stattfindet und somit weiterhin die traditionelle Rollenverteilung, in der die Leitung des Unterrichts durch die allgemeinbildende Lehrkraft und die Unterstützung der Schülerinnen und Schüler mit SPF durch sonderpädagogische Lehrkraft erfolgt, beibehalten wird. Die Analysen machen deutlich, dass die Mehrheit der Schülerinnen und Schüler in der Fragebogenerhebung das gemeinsame Lehren im Team positiv wahrnimmt und sich durch die beiden Lehrkräfte ebenfalls gut unterstützt fühlt. Die Ergebnisse von Multilevel-Regressionsanalysen zeigten, dass Schülerinnen im Gegensatz zu Schülern die Teamarbeit der Lehrkräfte positiver bewerten und auch als unterstützender wahrnehmen. Die positiven Befunde in Bezug auf die Bewertung der Teamarbeit sowie die wahrgenommene Unterstützung lassen sich auf die zwei Kooperationsformen Co-Teaching und individuelle Begleitung zurückführen. Des Weiteren zeigt sich je positiver die Schülerinnen und Schüler die Teamkooperation der Lehrkräfte bewerten und als unterstützender wahrnehmen, desto stärker schätzen die Schülerinnen und Schüler auch Gruppen- und Partnerarbeit. Darüber hinaus konnte ein signifikant positiver Zusammenhang zwischen der Teamarbeit und der wahrgenommenen Unterstützung durch die Lehrkooperation mit dem schulischen Wohlbefinden und dem sozialen Integriertsein der Schülerinnen und Schüler ausgemacht werden.

### Testung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften



**Kristina Gerhard<sup>1</sup>, Kai Kaspar<sup>2</sup>, Johannes König<sup>1</sup>, Conny Melzer<sup>3</sup>**

<sup>1</sup>Universität zu Köln, Empirische Schulforschung-Quantitative Methoden, <sup>2</sup>Universität zu Köln, Department Psychologie, <sup>3</sup>Universität zu Köln, Department Heilpädagogik

Theoretischer Hintergrund:

Professionelles Lehrerwissen wird inhaltlich in die Trias von fachlichem, fachdidaktischem und pädagogischem Wissen (Shulman, 1987) gegliedert. Lehrpersonen benötigen in diesen Bereichen Wissen, um qualitätsvolle unterrichtliche Lerngelegenheiten anzubieten und damit die Leistungs- und Persönlichkeitsentwicklung ihrer Schülerinnen und Schüler zu fördern (Baumert & Kunter, 2011; Helmke, 2012). Fraglich ist jedoch, ob die bislang konzipierten und empirisch untersuchten Wissenskategorien die gewandelten Anforderungen inklusiver Bildung in Schule und Unterricht bereits hinreichend abdecken oder ob das geforderte Spektrum professionellen Wissens von Lehrerinnen und Lehrern mit Blick auf ihren inklusiven Unterricht erweitert werden muss. Wir konkretisieren die Erweiterung einer Kategorie des Lehrerwissens um einen Bereich des Wissens für inklusiven Unterricht. Dies geschieht am Beispiel des pädagogischen Wissens. Die Erweiterung wird als pädagogisches Wissen für inklusiven Unterricht bezeichnet (general pedagogical knowledge for inclusive teaching, GPK-IT). Der Anspruch ist es, Wissen zu definieren, das für Lehrkräfte aller Lehrämter, Schulstufen und Schulformen von Relevanz ist. Basierend auf früheren Forschungen zum Lehrerwissen definieren wir GPK-IT als zentralen Bestandteil professioneller Kompetenz (König et al., 2017). Wir bauen auf einen bekannten Test aus dem TEDS-M-Projekt auf (König et al., 2011). Der Test GPK-IT geht allerdings über die bisherigen Konzepte und Operationalisierungen von GPK hinaus, wie in TEDS-M definiert. Basierend auf Literaturrecherchen von Lehrerkompetenzkatalogen zur Inklusion fokussiert der Test auf spezifische Kenntnisse zu den Anforderungen der Diagnose und Intervention.

Fragestellung:

Bearbeitet werden zwei Fragestellungen, die sich auf die Konstruktvalidität des neuentwickelten Tests beziehen: (1) Ist der Test sensitiv gegenüber der Ausbildung (Curriculumvalidität)? (2) Zeigen sich theoretisch erwartbare Zusammenhänge mit der kognitiven Leistungsfähigkeit (indiziert über die Abiturnote) und berufsbezogenen Überzeugungen zur inklusiven Bildung (Diskriminanz- bzw. Konvergenzvalidität)? Wir untersuchen folgende Hypothesen:

Masterstudierende unabhängig ihres Lehramts schneiden im Test statistisch signifikant besser ab als Bachelor-Studierende (H1a). Lehramtsstudierende für Sonderpädagogik zeigen höhere Testleistungen als Lehramtsstudierende für Regelschulen, wobei die Unterschiede nur von geringer praktischer Bedeutsamkeit sind (H1b). Die Testergebnisse zeigen geringe, jedoch statistisch signifikante Korrelationen mit der Abiturnote (H2a). Das Testergebnis korreliert signifikant mit Überzeugungen, (im Sinne eines engen und weiten Inklusionsbegriffs) (H2b). Das Testergebnis korreliert positiv mit jenen Überzeugungen, die inhaltlich eine Befürwortung inklusiver Bildung bzw. adaptiver Förderung enthalten, und negativ mit solchen Überzeugungen, die einer adaptiven Förderung von Lernenden entgegenstehen (H2c).

Methode:

Verwendet wurde der Teildatensatz aus dem Hochschulweiten Bildungsmonitoring der Universität zu Köln, der im Rahmen des Projekts der „Qualitätsoffensive Lehrerbildung“ durchgeführten Kompetenzmessungen aus dem Sommersemester 2017. Die Stichprobe besteht aus 409 Bachelor- und Master-Lehramtsstudierenden. Eingesetzt wurde das neuentwickelte Testinstrument GPK-IT. Die Testitems des Tests GPK-IT wurden mit Unterstützung von Expertinnen und Experten aus dem universitären und schulischen Bereich sowie teilweise in Anlehnung an vorhandenen Testfragen des Educational Testing Service, Princeton (2016) entwickelt. Zudem wurde das neuentwickelte Instrument zu Einstellungen und Überzeugungen zu Inklusion (vgl. Strauß & König, 2017), das auf 7 Skalen basiert, die insgesamt 27 Items beinhalten, eingesetzt.

Ergebnisse:

Die Item-Kennwerte für die Skalierung des Tests an der Gesamtstichprobe können als gut bezeichnet werden, da die gewichteten Abweichungsquadrate innerhalb des empfohlenen Bereichs liegen (.80 und 1.20, vgl. Bond & Fox, 2007). In einer eindimensionalen IRT-Skalierung unter Einschluss von 44 Items erweist sich der GPK-IT-Test als reliabel, Reliabilität von .60 ist noch akzeptabel; die durchschnittliche Item-Diskrimination liegt über .30. Die Ergebnisse geben erste Hinweise zur Konstruktvalidität: Der Test ist sensitiv insbesondere gegenüber dem Studienabschnitt, der spezifische Lerngelegenheiten im Bereich der Inklusion vorsieht, aber auch geringfügig sensitiv gegenüber erweiterten Lerngelegenheiten wie sie Studierende für das sonderpädagogische Lehramt erhalten. Ferner lassen sich theoretisch zu erwartende Zusammenhänge zwischen getestetem Wissen und Abiturnote sowie Überzeugungen zur Inklusion belegen. Die Skalen, die eine Zustimmung zur Inklusion enthalten, korrelieren positiv, die Skalen, die eine Ablehnung adaptiver Unterrichtsgestaltung enthalten, negativ mit den Testergebnissen.

### **Eine empirische Explikation von vier Definitionen schulischer Inklusion und ihr konsensueller Kern – Ergebnisse von Experteninterviews mit Inklusionsforscher\*innen zur Suche nach einem gemeinsamen Verständnis**

**Michael Grosche<sup>1</sup>, Anne Piezunka<sup>2</sup>, Tina Schaffus<sup>3</sup>**

<sup>1</sup>Bergische Universität Wuppertal, <sup>2</sup>Wissenschaftszentrum Berlin für Sozialforschung, <sup>3</sup>Universität Potsdam

Theoretischer Hintergrund: Nach wie vor wird die Erforschung von schulischer Inklusion durch theoretische und konzeptuelle Unklarheiten des Begriffs schulischer Inklusion erschwert (Grosche, 2015; Moser, 2012; Werning, 2010). So ist in den letzten Jahren der Eindruck entstanden, dass es zwei Pole der deutschsprachigen Inklusionsforschung gebe: Auf der einen Seite stehe die Inklusionsforschung, die den Fokus auf Schulleistungstests und die Optimierung von Diagnoseverfahren legt und somit eher der quantitativen Sonderpädagogik und Bildungsforschung nahesteht. Auf der anderen Seite gebe es einen Forschungsstrang, der die bestehenden Strukturen und Praktiken im Bildungssystem auf normativer Ebene kritisch betrachtet, seine Forschung durch den Verweis auf Menschenrechte legitimiert und sich aus vielen Disziplinen und Zugängen speist (z.B. Disability Studies, kritische Theorie, verstehende Sonderpädagogik). Beide Pole nehmen sich in der Regel nicht zur Kenntnis oder werten sich gegenseitig ab, wobei das kokonstruktive Potential einer interdisziplinären Bildungsforschung ungenutzt bleibt.

Fragestellungen: Um einen gemeinsamen Bezugspunkt für einen zielführenden und verständigungsorientierten interdisziplinären Diskurs zu liefern, möchten wir erstmalig empirisch statt theoretisch explizieren, welche Definitionen von schulischer Inklusion den wissenschaftlichen Diskurs im deutschsprachigen Raum prägen, worin sie sich unterscheiden und ob sie einen konsensuellen Definitionskern besitzen, der sich durch alle Definitionen zieht und somit für eine interdisziplinäre Zusammenarbeit genutzt werden könnte.

Methode: Wir interviewten elf Inklusionsforscher\*innen aus unterschiedlichen wissenschaftlichen Disziplinen (u. a. Disability Studies, Erziehungswissenschaft, Grundschulpädagogik, Psychologie, Rechtswissenschaften, Sonderpädagogik, Soziologie).

Für die Datenauswertung verknüpften wir das methodische Vorgehen der qualitativen Inhaltsanalyse für Experteninterviews nach Gläser und Laudel (2009) mit dem phänomenographischen Ansatz nach Marton und Booth (2014). Zwei Personen führten die Extraktion von Dimensionen schulischer Inklusion unabhängig voneinander durch. Zur Analyse diente der phänomenographische Ansatz als Perspektive auf die vorhandenen Daten. Im ersten Schritt arbeiteten wir heraus, welche Dimensionen von Inklusion relevant sind, um unterschiedliche Verständnisse zu analysieren. Im nächsten Schritt betrachteten wir die Ausprägungen der einzelnen Dimensionen und arbeiteten mehrere Inklusionsverständnisse heraus, die sich in den Dimensionen unterscheiden. In jedem Verständnis arbeiteten wir einen Kern heraus, durch den die Verständnisse anhand ihrer Komplexität hierarchisch geordnet werden können.

Ergebnisse: Wir identifizierten die folgenden vier Definitionen von Inklusion. (1) Definition anhand der UN-Behindertenrechtskonvention: Schulische Inklusion ist der Rechtsanspruch von Schüler\*innen mit Behinderungen auf Platzierung in allgemeinen Schulen. (2) Pragmatische Definition im Sinne von Leistungsentwicklung: Schulische Inklusion ist die bestmögliche akademische Förderung von Schüler\*innen durch Berücksichtigung ihrer lernbezogenen Bedürfnisse, um individuell bestmögliche Lernergebnisse zu erreichen. (3) Definition von Inklusion als Teilhabe / Anerkennung / Wohlfühlen: Schulische Inklusion ist die bestmögliche Leistungsförderung und die unbedingte Anerkennung aller Schüler\*innen unabhängig von jeglichen Voraussetzungen als vollständiger Teil der Gemeinschaft, um ihnen vollständige soziale Teilhabe an sämtlichen schulischen Handlungen zu ermöglichen. (4) Definition von Inklusion als Utopie: Schulische Inklusion ist die Vision bzw. Utopie eines Denkens und Handelns ohne Kategorien. Beim Vergleich dieser vier Definitionen wird deutlich, dass alle vier Verständnisse auf die Überwindung von Diskriminierung aufgrund von sozial konstruierter Gruppenzugehörigkeit abzielen, um Teilhabe in Schule und Gesellschaft zu ermöglichen. Dieser konsensuelle Kern durchzieht alle vier Definitionen und ordnet sie anhand ihrer Komplexität hierarchisch.

Diskussion: Wir leiten aus unseren Ergebnissen ab, dass in jedem der erforschten Konzepte dieser konsensuelle Kern als Konsens vorliegt und somit expliziert werden könnte. Durch unsere Ergebnisse werden Forscher\*innen mit unterschiedlichen Inklusionsverständnissen in die Lage versetzt, zu konkretisieren, wie sie über die Überwindung von Diskriminierung in ihrer Forschung nachdenken, wie sie die Überwindung von Diskriminierung theoretisch in ihre Forschung eingearbeitet haben und wie sie die Überwindung von Diskriminierung operationalisieren. Somit bietet sich der konsensuelle Kern an, um ein interdisziplinäres Verständnis von schulischer Inklusion zu ermöglichen.

## E01

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S103

### Urteile über Schüler/-innen mit Migrationshintergrund: Zur Bedeutung von Diagnose und Prognose

**Meike Bonefeld, Karina Karst, Oliver Dickhäuser**

Universität Mannheim, Deutschland; [meike.bonefeld@staff.uni-mannheim.de](mailto:meike.bonefeld@staff.uni-mannheim.de)

Lehrkräfte fällen im Schulalltag regelmäßig Urteile über Schüler/-innen. Die Akkuratheit dieser Urteile kann durch verschiedene Lehrer-, Schüler- sowie Urteilsmerkmale beeinflusst werden (Stückamp, Kaiser & Möller, 2012). Analysen zu kulturellen Stereotypen in Deutschland konnten zeigen, dass Menschen mit Migrationshintergrund weniger stark mit leistungs- und erfolgsbezogenen Attributen beschrieben werden als Menschen ohne Migrationshintergrund (Kahraman & Knoblich, 2000). Das Schülermerkmal „Migrationshintergrund“ könnte so die Urteilsakkuratheit beeinflussen, und zwar umso stärker, je unsicherer das zu fällende Urteil ist. Zwei Urteile, die sich in dem Ausmaß ihrer Unsicherheit unterscheiden sind diagnostische und prognostische Urteile. Zu Prognosen könnten Erwartungen über zukünftige Leistungen, wie sie beispielsweise auch bei Schullaufbahnpfehlungen eine Rolle spielen, zählen. Zu Diagnosen könnten unter anderem eine einfache Kategorisierung von Schüler/-innen zu einem Leistungsniveau zählen.

Beide Urteile können auf der gleichen Leistungsinformation beruhen (Ingenkamp & Lissmann, 2008), unterscheiden sich jedoch in ihrer Sicherheit (größere Unsicherheit bei langfristigen Prognosen). Diese Urteile orientieren sich nur teilweise an unmittelbar beobachtbarem und beschreibbarem Verhalten und erfordern ergänzende Schlussfolgerungen aus beschreibbaren Faktoren

Die vorliegende Arbeit untersucht daher vergleichend diagnostische sowie prognostische Urteile auf Grundlage der gleichen Leistungsinformation. 171 angehende Lehrkräfte gaben Leistungseinschätzungen (diagnostisches Urteil) sowie Leistungserwartungen (prognostisches Urteil) über Grundschüler/-innen mit und ohne Migrationshintergrund ab. Den Teilnehmer/-innen wurden Mathematikleistungen in einem Multiplikationstest von zwölf Grundschüler/-innen der zweiten Jahrgangsstufe präsentiert. Die Kinder unterschieden sich dabei in Bezug auf ihr Leistungsniveau, das Geschlecht und den Migrationshintergrund (türkisch vs. deutsch). Geschlecht sowie Migrationshintergrund wurden durch die Namen der Kinder operationalisiert, welche im Rahmen von Vortests pilotiert wurden. Dabei wurden vergleichend verschiedene Urteilsarten gegenübergestellt: Auf Grundlage der gleichen Leistungsinformation sollten sie je eine Leistungseinschätzung des vorliegenden Kindes abgeben sowie eine Erwartung über zukünftige Leistungen des Kindes. Einerseits sollten die Lehrkräfte das vorliegende Leistungsniveau der Schüler/-innen auf einer 5-stufigen Skala von „sehr gut“ bis „sehr schlecht“ bewerten (Leistungseinschätzung, diagnostisch) sowie andererseits angeben welche Leistungen der Schüler/-innen sie in einem ähnlichen Test sieben Tage später erwarten (Leistungserwartung, prognostisch).

Die Analysen wurden mit Hilfe von latenten linearen Mischmodellen (zur Berücksichtigung der genesteten Datenstruktur) im Softwarepaket lme4 (Bates, Maechler, Bolker & Walker, 2015) in R (R Development Core Team, 2013) durchgeführt.

Es zeigte sich, dass die Probanden wie erwartet besonders bei der zukünftigen Leistungserwartung Kindern mit vermeintlichem Migrationshintergrund bei gleichem Leistungsniveau schlechtere Leistungen prognostizierten als den Kindern ohne Migrationshintergrund. In Bezug auf die aktuelle Leistungseinschätzung fällten die Lehrkräfte keine statistisch signifikant unterschiedlichen Urteile über die Schüler/-innen mit und ohne Migrationshintergrund. Sie ordneten die Kinder unabhängig ihres Migrationshintergrundes vergleichsweise akkurat verschiedenen Leistungsniveaus zu.

Festgehalten werden kann, dass Disparitäten nach dem Migrationshintergrund in ihrem Auftreten nach der Urteilsart variieren können. Dass gerade die Erwartungen von Lehrkräften über die zukünftigen Leistungen von Schüler/-innen in Abhängigkeit des Migrationshintergrundes verschieden sind deuten wir als einen Hinweis auf einen Einfluss von Stereotypen bei prognostischen Urteilen. Die Ergebnisse stützen die Vermutung, dass im Rahmen prognostischer Urteile wie der Erwartungsbildung Urteilsfehler in Abhängigkeit des Migrationshintergrundes zum Tragen kommen während solche Einflüsse offenbar nicht gleichermaßen auf diagnostische Urteile (Wiedergabe des Leistungsniveaus) wirken. Dies eröffnet erste Ansatzpunkte für die Vermeidung dieser Fehler.

### Differenzielle Effekte des Unterrichtsklimas bei Schüler\*innen mit Migrationshintergrund

**Cansu Topalak<sup>1</sup>, Samuel Merk<sup>1</sup>, Benjamin Fauth<sup>1</sup>, Jasmin Decristan<sup>2</sup>, Gerhard Büttner<sup>3</sup>, Eckhard Klieme<sup>4</sup>**

<sup>1</sup>Universität Tübingen, Deutschland; <sup>2</sup>Institut für Bildungsforschung, Bergische Universität Wuppertal; <sup>3</sup>Institut für Psychologie, Goethe Universität Frankfurt am Main; <sup>4</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF); [cansu.topalak@uni-tuebingen.de](mailto:cansu.topalak@uni-tuebingen.de)

Theoretischer Hintergrund:

Das Unterrichtsklima ist ein konstitutives Merkmal der Unterrichtsqualität. Die vorliegende Arbeit fokussiert Interaktionen zwischen Schülerinnen und Schülern wie Hilfsbereitschaft, keine Ausgrenzungen Einzelner und gegenseitigen Respekt und Anerkennung als Merkmale des Unterrichtsklimas. Ausgangspunkt sind verschiedene Ansätze in der empirischen Unterrichtsforschung, die Evidenz dafür zeigen, dass Beziehungsaspekte als Qualitätskriterium von Unterricht gelten (u.a. Hamre & Pianta, 2005). Auf der Grundlage des Angebot-Nutzungs-Modells (Helmke, 2014) ist anzunehmen, dass es Wechselwirkungen zwischen individuellen Schülermerkmalen und Unterrichtsmerkmalen gibt. Beruhend auf den Ergebnissen der internationalen Schulleistungsstudien, sowohl in der Primar- wie auch in der Sekundarstufe, haben Schülerinnen und Schüler mit Migrationshintergrund schlechtere Schulleistungen als Kinder ohne Migrationshintergrund (z.B. Bos et al., 2007). Bereits mit Eintritt in die Grundschule starten Schülerinnen und Schüler mit Migrationshintergrund mit einem Risikofaktor ihre Schullaufbahn (u.a. Schultheis, 2012). Statistisch ausgedrückt bedeutet es, dass die Zugehörigkeit zu einer Migrationsgruppe die Wahrscheinlichkeit erhöht, geringere Bildungserfolge zu erzielen (Seiz et al., 2016). Entsprechend wird in der Unterrichtsforschung sowie in Studien zur Leistungsentwicklung für Lernende mit Migrationshintergrund der Begriff „Risikoschüler bzw. Risikoschülerin“ verwendet (u.a. Hasselhorn et al., 2014). Beruhend auf ersten Befunden wird von kompensatorischen Wirkungen von Klassenführung und Unterstützung für Schülerinnen und Schüler mit Migrationshintergrund ausgegangen (Seiz et al., 2016). Die Frage, inwieweit Schüler-Schüler-Beziehungen einen Einfluss auf die schulische Leistung von Kindern mit Migrationshintergrund haben, konnte in der empirischen Unterrichtsforschung bis heute noch nicht zufriedenstellend beantwortet werden.

Fragestellung:

Vor dem Hintergrund der skizzierten theoretischen Zugänge und empirischen Befunde im Kontext des Angebot-Nutzungs-Modells untersucht diese Studie differenzielle Wirkungen allgemeiner Unterrichtsmerkmale. Es wird geprüft, ob ein positives

Unterrichtsklima besonders lernförderlich für Schülerinnen und Schüler mit Migrationshintergrund ist und somit deren tendenziell ungünstigere Leistungsentwicklung (partiell) kompensieren kann. Die grundlegende Forschungsfrage ist die Frage nach der Bedeutsamkeit des Unterrichtsklimas und dessen Auswirkungen für Schülerinnen und Schüler mit Migrationshintergrund in Bezug auf ihre Leistungsentwicklung. Auf folgende konkrete Forschungsfrage soll am Ende eine Antwort gegeben werden: Zeigen sich differentielle Effekte eines positiven Unterrichtsklimas für Schülerinnen und Schüler mit Migrationshintergrund in Bezug auf ihre Leistungsentwicklung und sind diese durch Variablen konfundiert, welche dem Lernprozess theoretisch proximaler sind?

Methode:

Für die Untersuchung wurde anhand von videogestützten Analysen ein Instrument zum Unterrichtsklima entwickelt. Die Unterrichtsvideos sind im Rahmen des Projekts IGEL (Individuelle Förderung und adaptive Lern- Gelegenheiten in der Grundschule) entstanden (Hardy et al., 2011). Grundlage dieser Arbeit ist ein längsschnittlich angelegter Datensatz mit N = 721 Schülerinnen und Schülern aus 37 Grundschulklassen. Die Leistungen der Schülerinnen und Schüler wurden über standardisierte Leistungstests erfasst. Für die in dieser Studie vorgenommenen Analysen der Schüler-Schüler-Beziehungen wurde ein hoch inferentes Kodiersystem entwickelt. Hierfür wurde das Unterrichtsklima in den videographierten Unterrichtsstunden anhand von drei Items (Wertschätzender Umgang, geteilte Lernfreude und Lerngemeinschaft) erfasst. Zur Analyse der Daten wurden multiple lineare Regressionsmodelle berechnet. Da es sich um eine geschachtelte Struktur der Daten handelt, nämlich Schülerinnen und Schüler in Klassen, wurden Parameter von Mehrebenenmodellen in Mplus geschätzt (Muthen & Muthen, 1998- 2012).

Ergebnisse:

Der allgemeine Einfluss des Unterrichtsklimas auf die Post-Testleistung wird in den Analysen nicht bestätigt. Jedoch wird in den hierarchisch linearen Modellen mit Cross-Level-Interaktionen ein positiver Interaktionseffekt zwischen Unterrichtsklima und Migrationshintergrund auf die Post-Testleistung evident. Das bedeutet, dass ein positives Unterrichtsklima die schlechteren Leistungen von Schülerinnen und Schülern mit Migrationshintergrund partiell kompensieren kann. Darüber hinaus zeigt sich dieser Interaktionseffekt auch nach Kontrolle von Variablen, die diesen möglicherweise konfundieren (Selbstkonzept, soziale Eingebundenheit, Klassenklima und Sprache.), da sie dem Lernprozess proximaler sind.

### **Akkulturationsorientierungen von Schülerinnen und Schülern mit Migrationshintergrund und ihren Lehrpersonen – der Einfluss von Passung und Nicht-Passung auf die Schulzufriedenheit und das akademische Selbstkonzept**

**Andrea Haenni Hoti, Christine Wolfgramm, Marianne Müller, Alois Buholzer**

Pädagogische Hochschule Luzern, Schweiz; [andrea.haenni@phlu.ch](mailto:andrea.haenni@phlu.ch), [christne.wolfgramm@phlu.ch](mailto:christne.wolfgramm@phlu.ch)

Nach Berry, Poortinga, Breugelmans und Chasiotis, sind Schulen das wichtigste Umfeld für die Akkulturation von Kindern und Jugendlichen mit Migrationshintergrund. Schulen und Lehrpersonen repräsentieren die Mehrheitskultur und können diese den Schülerinnen und Schülern näher bringen (2011, p. 326). Eine gute Beziehung zur Lehrperson ist ausserdem für die soziale Integration und das schulische Wohlbefinden der Schülerinnen und Schülern (SuS) wichtig (Haenni Hoti, Heinzmann, Müller & Buholzer, 2017).

Die Befundlage zur Frage, welche Akkulturationsorientierung mit schulischem Wohlbefinden (und Bildungserfolg) zusammenhängt, ist inkonsistent. Studien haben wiederholt auf die Vorteile der kulturellen Integration hingewiesen (z.B. Nguyen & Benet-Martínez, 2013). Andere Befunde zeigen jedoch, dass kulturell assimilierte SuS in einem auf Assimilation ausgerichteten Schulsystem besser zurecht kommen können und eine Integrationsorientierung eines diversitätsfreundlichen Umfeldes in der Schule bedarf (z.B. Makarova & Birman, 2015). Es ist deshalb naheliegend, den Fokus der Akkulturationsforschung stärker auf das Zusammenspiel von Akkulturationsorientierungen der SuS und dem von ihnen erlebten Schulkontext, insbesondere die Wahrnehmung ihrer Klassenlehrperson zu richten.

Die Studie geht der folgenden Frage nach: Welche Bedeutung hat die Passung zwischen den Akkulturationsorientierungen von SuS mit Migrationshintergrund und denjenigen ihrer Lehrpersonen für die Schulzufriedenheit und das akademische Selbstkonzept der SuS?

Die Stichprobe (convenience sample) besteht aus 1'112 SuS in 60 Schulklassen der 5. Klasse in der deutschsprachigen Schweiz (53% Mädchen, 10-12 Jahre alt) und ihren Klassenlehrpersonen (N=60). Während 79% der befragten SuS einen Migrationshintergrund besitzen, weist die Mehrheit der Klassenlehrpersonen (83%) keinen Migrationshintergrund auf. Für die Erhebung der Akkulturationsorientierungen von Primarschulkindern wurden eigene Skalen entwickelt. Diese beinhalten neben den zwei Dimensionen „Erhalt der Minderheitenkultur“ und „Übernahme der Mehrheitskultur“ die Dimensionen „Multikulturelle Orientierung“ und „kulturelle Indifferenz“. Die selbst berichteten sowie die von den SuS wahrgenommenen Akkulturationsorientierungen der Lehrpersonen wurden teils mit neuen und teils mit erprobten Skalen erhoben (Fürstenau, Huxel & Diekmann, 2014). Die Schulzufriedenheit und das akademische Selbstkonzept wurden mittels Instrumenten von Venetz, Zurbriggen und Eckhart (2014) erfasst. Darüber hinaus wurden die Qualität der Lehrer-Schüler-Beziehung sowie zahlreiche schulkontextuelle und demographische Variablen kontrolliert. Die Daten wurden mithilfe von Multiplen Regressionen und Mehrebenenanalysen ausgewertet.

Die Ergebnisse zeigen, dass die Passung zwischen der multikulturellen Orientierung der SuS und ihren Lehrpersonen (z.B. die Schülerin/der Schüler findet es gut, wenn die Klasse die verschiedenen Sprachen der SuS kennt und die Lehrperson findet wichtig, dass die Schule Kinder dabei unterstützt, sich in der multikulturellen Gesellschaft zurechtzufinden, oder beide lehnen dies ab) mit einer höheren Schulzufriedenheit der SuS einhergeht. Auch eine Passung bezüglich der Mehrheitsorientierung (beide Seiten sind z.B. der Ansicht, dass sich SuS mit Migrationshintergrund kulturell assimilieren sollten, oder beiden Seiten unterstützen dies nicht) hängt mit einer höheren Schulzufriedenheit der SuS zusammen. Zudem sagten auch der Migrationshintergrund der SuS, das Geschlecht, das Alter der Lehrperson und die Qualität der Beziehung zwischen Lehrperson und SuS deren Schulzufriedenheit voraus ( $R^2 = .30$ ).

Bezüglich des akademischen Selbstkonzepts zeigte sich, dass eine Nicht-Passung zwischen der eigenen nationalen Identifikation und der von der Lehrperson wahrgenommenen Zuschreibung (z.B. das Kind sieht sich als Schweizer/-in, denkt aber, dass die Lehrperson es als Albaner/-in sieht) mit einem weniger guten akademischen Selbstkonzept einhergeht, allerdings nur bei älteren Lehrpersonen ab 50 Jahren. Eine Passung bezüglich der Minderheitenorientierung steht allerdings in einem negativen Zusammenhang mit dem akademischen Selbstkonzept der SuS. Dieser unterwartete Effekt hat möglicherweise eine methodische Ursache (eine Lehrperson ohne Migrationshintergrund besitzt selber keine Minderheitenorientierung). Des Weiteren wird das akademische Selbstkonzept durch die Qualität der Lehrer-Schüler-

Beziehung und demographische Faktoren erklärt: Mädchen, Kinder ohne Schweizer Staatsangehörigkeit und Kinder mit weniger Bildungsressourcen zu Hause, gemessen an der Anzahl Bücher haben ein schlechteres schulisches Selbstkonzept ( $R^2 = .22$ ). Dagegen beeinflusst eine gute Beziehung zwischen Schüler/Schülerin und Lehrperson das akademische Selbstkonzept der SuS positiv unabhängig von der Passung oder Nicht-Passung der Akkulturationsorientierungen.

### **Herkunftseffekte auf Studienaspirationen und Studieneintritt: Eine Herausforderung für die Beratungskompetenz von Lehrkräften?**

**Claudia Schuchart, Benamin Schimke**

BU Wuppertal, Deutschland; [schuchart@uni-wuppertal.de](mailto:schuchart@uni-wuppertal.de)

Bildungsaspirationen gehören zu den stärksten Prädiktoren von Bildungsentscheidungen (Eccles 2004). Die Beratung im Blick auf ihre Realisierung gehört zu den professionellen Aufgaben von Lehrkräften (Hertel 2009). Insofern muss Wissen um die Determinanten von Aspirationen und ihre Entwicklung als Teil professionellen Handlungswissens verstanden werden.

Ein prägender Faktor von Aspirationen ist die soziale Herkunft. Herkunftseffekte auf den Übertritt in das Studium haben sich in den letzten Jahren verschärft (Schindler & Reimer 2010). Sie sind neben schichtspezifischen Differenzen in den Leistungen vor allem auf schichtspezifische Entscheidungsdifferenzen zurückzuführen. Bisherige Studien blenden Fragen zur Entwicklung der der Entscheidung vorausgehenden Aspiration weitestgehend aus. Damit besteht auch wenig Wissen darüber, wo z.B. Beratungslehrkräfte mit Maßnahmen zu einer Minderung von Herkunftseffekten ansetzen können. Im Mittelpunkt des geplanten Vortrags steht die Frage, ob schichtspezifische Entscheidungen für oder gegen ein Studium am Ende der SEK II das Ergebnis eines gleichbleibenden, abnehmenden oder zunehmenden Einflusses der sozialen Herkunft im Verlauf der SEK II sind.

Bildungsentscheidungen werden als Folge einer Bildungsaspiration verstanden, die im Sinne der Theorie des „planned behavior“ (TPB, z.B. Ajzen 1991) von drei Hauptkonzepten bestimmt wird: Der Einstellung zur Bildungsalternative (wird diese z.B. als nützlich und zielführend wahrgenommen), der subjektiven Norm (Erwartungen der als relevant bewerteten Anderen) und der wahrgenommenen Kontrolle über die Realisierung der Absicht (sog. Verhaltenskontrolle in Bezug auf wahrgenommene Schwierigkeiten, die die Realisierung verhindern können). Im Vortrag soll die schichtspezifische Entwicklung der Hauptkonzepte und ihre Bedeutung für die Formung von Bildungsaspiration und Studienentscheidung untersucht werden.

Folgende Entwicklungslinien sind grob skizziert denkbar: 1) Bereits am Ende der SEK I haben SchülerInnen eine stabile Bildungsaspiration unter dem Einfluss schichtspezifischer Sozialisationsbedingungen geformt (z.B. Sewell et al. 1969). Es besteht ein Herkunftseffekt auf die Hauptkonzepte zu Beginn der Sek II, der im Verlauf der Sek II stabil bleibt. 2) Der Übertritt in einen Bildungsgang der SEK II, der zur Studienberechtigung führt, könnte für SchülerInnen aus unteren Schichten zunächst noch mit einer idealistischen Studienaspiration (Stocké 2013) verknüpft sein. Die Bewusstwerdung von Hindernissen wie Kosten und Leistungsvoraussetzungen könnte im Verlauf der Sekundarstufe in dieser Gruppe jedoch stärker werden. Der Einfluss der sozialen Herkunft auf die Aspiration würde steigen aufgrund einer Verringerung der Wahrnehmung der Verhaltenskontrolle bei Schülern unterer Schichten. 3) Schließlich ist der Übergang in die SEK II für viele SchülerInnen mit einem neuen Kontext verbunden, der für Nichtgymnasiasten, die zugleich eher aus unteren Herkunftsgruppen kommen (Schindler 2014), im Vergleich zur vorherigen Schulform (z.B. Hauptschule, Realschule) oft mit einer stärkeren Studienorientierung verbunden ist (eigene Auswertungen). Der Einfluss der sozialen Herkunft auf die Studienaspiration könnte sinken aufgrund einer positiven Veränderung der Einstellung zum Studium bei SchülerInnen unterer Schichten (bei dennoch insgesamt geringerem Ausgangs- und Endniveau der Studienaspiration).

Diese drei Annahmen werden mit einer für NRW repräsentativen Stichprobe von 5562 SchülerInnen in allgemeinbildenden und beruflichen Bildungsgängen, die zur Studienberechtigung führen, geprüft. Die SchülerInnen wurden zu Beginn ( $t_1$ ) und zum Ende ( $t_2$ ) der SEK II sowie ein Jahr nach Erwerb der Studienberechtigung befragt. Die abhängige Variable ist die Studienentscheidung, die auf die zu  $t_1$  und  $t_2$  gemessenen Hauptkonzepte und die zu  $t_2$  gemessene Studienaspiration zurückgeführt wird. Die letztgenannten Merkmale werden als Skalen erfasst, die insgesamt sehr hohe Reliabilitäten aufweisen ( $\alpha \geq .80$ ), kontrolliert wird für Noten und Bildungsgang.

Gerechnet werden Mehrebenenstrukturgleichungsmodelle (Mplus 8), fehlende Werte werden imputiert. Die Analysen zeigen, dass Herkunftseffekte auf den Studienübertritt bereits durch eine schichtspezifische Ausprägung der Hauptkonzepte zu  $t_1$  erklärt werden können. Eine schichtspezifische Entwicklung der Aspiration als Folge einer differentiellen Entwicklung der Hauptkonzepte liegt nicht vor. Das spricht für die oben formulierte Annahme 1. Schichtspezifische Studienaspirationen scheinen sich bereits in der SEK I zu stabilisieren, was für die besondere Bedeutung von Maßnahmen der Bildungsberatung bereits zu diesem Zeitpunkt spricht.

## E02

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S104

### **Ist traditionelle Männlichkeit nicht mit guten Schulleistungen vereinbar? Der Zusammenhang von Geschlechternormen und -unterschieden in Noten bei Jugendlichen**

**Zerrin Salikutluk<sup>1</sup>, Georg Lorenz<sup>2</sup>, Sarah Gentrup<sup>2</sup>, Kristin Schotte<sup>1,2</sup>**

<sup>1</sup>Berliner Institut für empirische Integrations- und Migrationsforschung, Humboldt-Universität zu Berlin, Deutschland; <sup>2</sup>Institut zur Qualitätsentwicklung im Bildungswesen, Humboldt-Universität zu Berlin; [Zerrin.Salikutluk@hu-berlin.de](mailto:Zerrin.Salikutluk@hu-berlin.de)

Mit der Bildungsexpansion ab den 1960ern wurden Prozesse in Gang gesetzt, die Veränderungen in den geschlechtsspezifischen Bildungschancen hervorgebracht haben. In den meisten westlichen Nationen lässt sich eine Umkehrung des Geschlechterverhältnisses bei akademischen Erfolgen in vielen Etappen des Bildungssystems verzeichnen (Breen et al. 2010; Buchmann et al., 2008). Dieser Beitrag setzt sich insbesondere mit dem Zusammenhang zwischen Geschlechternormen von Jugendlichen sowie dem damit verbundenen geschlechtsspezifischen Verhaltensmustern und ihren Leistungen in der Schule auseinander.

Vor allem in der frühen Adoleszenz nehmen Abgrenzungsprozesse anhand des Geschlechts eine vorrangige Bedeutung bei der Identitätsentwicklung ein (Hill und Lynch, 1983). So gehen im Jugendalter vorherrschende Männlichkeitsnormen (im Schulkontext) mit einer demonstrativen Ablehnung von schulischem Engagement, anti-sozialem Verhalten gegenüber Schulautoritäten und dem Abweichen von geltenden Regeln im Unterricht einher. Engagement für gute Leistungen und gute Noten hingegen werden als zentrale Eigenschaften der Weiblichkeitsvorstellung gesehen (vgl. Hadjar et al., 2012; Legewie und DiPrete, 2012). Im Schulalltag durchlaufen Jungen über Anerkennung und Abgrenzung durch ihre gleichgeschlechtliche Peergroup eine soziale Konditionierung dieser männlichen Verhaltensweisen und weisen häufiger deviantes Verhalten auf. Verstärkt werden die dem Schulerfolg entgegenstehenden Einstellungs- und Verhaltensmuster durch die Freizeitgestaltung von Jungen, die ebenfalls aus als maskulin geltenden Tätigkeiten besteht. Besonders prominent ist die Zurückführung devianter Verhaltensweisen auf den Konsum von Filmen und Videospiele, deren Inhalte und Handlungsstränge diese begünstigen sollen (Baier und Pfeiffer, 2011). Diese Aktivitäten können gemeinsam mit der Peergroup unternommen werden, wodurch auch über Klassengrenzen hinweg die Auslebung der Vorstellung von traditioneller Maskulinität gefördert wird. Im Gegensatz dazu widmen sich Mädchen in ihrer Freizeit eher leistungsfördernden Aktivitäten wie Zeichnen, Musizieren oder Lesen und verbringen mehr Zeit mit Hausaufgaben (Baier und Pfeiffer, 2011). Zusammengefasst gehen wir davon aus, dass Geschlechterunterschiede in den Noten (zumindest teilweise) über Geschlechternormen und den damit verbunden Verhaltensweisen erklärt werden können.

Für unsere Analysen verwenden wir die Daten des Projektes Children of Immigrants Longitudinal Survey in Four European Countries für Deutschland (N=3.577 in 126 Klassen; 51% weiblich; Durchschnittsalter=14 Jahre). Da die Daten im Rahmen von Klassenbefragungen erhoben wurden, berechnen wir genestete Mehrebenenmodelle. Im ersten Schritt analysieren wir den Zusammenhang zwischen den Geschlechternormen der Jugendlichen, den Normen und Verhaltensweisen ihrer Klassenkamerad\*innen und dem schulbezogenen Verhalten und Freizeitaktivitäten der Jugendlichen. Im zweiten Schritt prüfen wir, inwiefern diese Indikatoren Geschlechterunterschiede im Notendurchschnitt in den Fächern Deutsch, Mathe und Englisch am Ende der neunten Klasse erklären können.

Erwartungsgemäß zeigen unsere Ergebnisse, dass Jungen sich häufiger deviant in der Schule verhalten als Mädchen. Traditionelle Geschlechternormen sowie Klassenkamerad\*innen, die zu abweichendem Verhalten neigen, verstärken dies bei Jungen. Generell haben Mädchen, wie bereits in früheren Studien vielfach gezeigt wurde, unter Kontrolle der sozialen und ethnischen Herkunft bessere Noten im Durchschnitt als Jungen. Dabei wirken sich traditionelle Geschlechternormen bei Schülern negativ auf deren Noten aus, während bei Schülerinnen kein statistisch signifikanter Zusammenhang zu finden ist. Ein großer Teil des negativen Zusammenhangs von traditionellen Geschlechternormen bei Jungen scheint über das schulbezogene Verhalten und die Freizeitaktivitäten zu verlaufen. Werden beide Indikatoren – das Verhalten in und außerhalb der Schule – in die Analysen aufgenommen, verringert sich einerseits der negative Effekt traditioneller Geschlechternormen für Jungen und andererseits verbleiben keine statistisch signifikanten Geschlechterunterschiede zugunsten von Schülerinnen.

### **Determinanten musikalischer Freizeitaktivitäten bei Jungen und Mädchen**

**Marcus Penthin<sup>1</sup>, Eva Susanne Fritzsche<sup>2</sup>, Kröner Stephan<sup>1</sup>**

<sup>1</sup>Friedrich-Alexander Universität Erlangen-Nürnberg, Deutschland; <sup>2</sup>Technische Universität München, Deutschland; [marcus.penthin@fau.de](mailto:marcus.penthin@fau.de)

Die Grundlagen musikalischer Bildung werden häufig schon in der Kindheit gelegt (Gembris, 2013): im Grundschulalter – teilweise bereits nach dem Durchlaufen einer musikalischen Früherziehung – beginnen viele Kinder in Musikschulen (Verband deutscher Musikschulen, 2015) oder in sonstigen privat organisierten Settings mit musikalischen Aktivitäten (Corrigall & Schellenberg, 2015; Zerte, 2008). Dabei tun Mädchen dies häufiger als Jungen (Statistisches Bundesamt, 2015). Nicht alle Kinder werden jedoch musikalisch aktiv. Wodurch lässt sich erklären, in welchem Ausmaß Schülerinnen und Schüler Musizieren? Welche Determinanten tragen dazu bei und unterscheidet sich die Relevanz dieser geschlechtsspezifisch? Bisher zeigten sich im Grundschulbereich direkte Effekte des musikalischen Selbstkonzepts, welches bei Mädchen höher ausgeprägt ist (Nonte & Schwippert, 2012) und indirekte Effekte der wahrgenommenen musikalischen Praxis der Eltern (Kröner, Schwanzer & Dickhäuser, 2009) sowie der Bedeutung von Musik im Elternhaus (Krupp-Schleußner & Lehmann-Wermser, 2016) auf musikalische Aktivitäten. Weiterhin verweisen deskriptive Befunde auf die Rolle der Eltern sowie auf die Bedeutung des affektiven Einstellungsaspekts (Grgic & Züchner, 2013), der bei Mädchen höher ausgeprägt ist (Medienpädagogischer Forschungsverbund Südwest, 2017). Dem Rahmenmodell kultureller Partizipation (Kröner, 2013) zufolge sollten darüber hinaus der instrumentelle Einstellungsaspekt und die Kontrollierbarkeit ebenfalls relevante Determinanten für die Aufnahme musikalischer Aktivitäten darstellen. Zu den Determinanten musikalischer Aktivitäten besteht nach wie vor Forschungsbedarf, insbesondere im Hinblick auf das Grundschulalter sowie bzgl. der Entwicklung geeigneter Forschungsinstrumente (Corrigall & Schellenberg, 2015; Grgic & Züchner, 2013). Entsprechend den im theoretischen Rahmenmodell enthaltenen Determinanten wurden Skalen entwickelt und N = 685 Viertklässler aus 54 Klassen an 16 mittelfränkischen Grundschulen befragt (n = 343 Mädchen, n = 342 Jungen; M = 9.88 Jahren, SD = 0.52; vgl. Wießnet, Kröner, Penthin & Fritzsche in diesem Band). Erhoben wurden der affektive (sechs Items, kongenerische Reliabilität [Omega] = .83) und der instrumentelle Einstellungsaspekt (Kosten/Nutzen der Aktivität, drei Items, Omega = .80); je drei Items zur wahrgenommenen musikalischen Praxis naher Bezugspersonen (deskriptive subjektive Norm in Bezug auf die Mutter, Omega = .64, den Vater, Omega = .77 sowie die Freunde, Omega = .80) und zur wahrgenommenen Rückmeldung auf eigenen musikalischen Aktivitäten (injunktive subjektive

Norm in Bezug auf die Mutter, Omega = .88, den Vater, Omega = .89 sowie den Freunden, Omega = .81); das musikalische Selbstkonzept (fünf Items, Omega = .89) und die subjektive Einschätzung von Umwelteinflüssen, die die Ausführung musikalischer Aktivitäten erleichtern oder erschweren (Kontrollierbarkeit, sechs Items, Omega = .87). Als Kriterium dienten drei Items zur Häufigkeit und Dauer des freizeithlichen Musizierens und des außerschulischen Musikunterrichts (Omega = .81). Alle Determinanten korrelierten statistisch signifikant mit dem Kriterium ( $|.19| \leq r \leq |.83|$ ) sowie untereinander ( $|.31| \leq r \leq |.79|$ ). Das entwickelte Messmodell wies strikte geschlechtsspezifische Invarianz auf. Mädchen waren musikalisch aktiver als Jungen ( $r = -.21$  [.03],  $p < .001$ ). Im Strukturgleichungsmodell (MLR-Schätzer;  $\chi^2 = 1130$  [715]  $p < .001$ , SRMR = .042, RMSEA = .029 [90% = .026, .032], CFI = .966, TLI = .961) wurden diese Geschlechtseffekte komplett über die Determinanten vermittelt, wobei direkte Effekte der Kontrollierbarkeit (Beta = .47 [.12]), des musikalischen Selbstkonzepts (Beta = .37 [.08]), des instrumentellen Einstellungsaspekts (Beta = -.26 [.06]) und der wahrgenommenen musikalischen Aktivität der Mutter (Beta = .16 [.06]) auf das Kriterium auszumachen waren ( $R^2 = .76$ ). Die Relevanz der Determinanten unterscheidet sich nach Geschlecht: Zeigten sich für die Jungen direkte Effekte des musikalischen Selbstkonzepts, des instrumentellen Einstellungsaspekts und der wahrgenommenen musikalischen Aktivitäten der Mutter, so trug bei den Mädchen lediglich die Kontrollierbarkeit und das musikalische Selbstkonzept zur direkten Erklärung musikalischer Aktivitäten bei. Diese geschlechtsspezifische Variation der Determinanten sollte bei der Förderung musikalischer Aktivitäten und künftigen Feldexperimenten zur Wirkung musikpädagogischer Interventionen beachtet werden. Die in der vorliegenden Arbeit vorgestellten Skalen können als Ausgangspunkt für zukünftige Feldexperimente genutzt werden.

### **Geschlechtsspezifische Unterschiede in der Wahrnehmung positiven und negativen Feedbacks im Leseunterricht der Grundschule**

**Fabian Hoya, Frank Hellmich**

Universität Paderborn, Deutschland; [fabian.hoya@campus.uni-paderborn.de](mailto:fabian.hoya@campus.uni-paderborn.de)

Feedback von Lehrkräften stellt eine wichtige Determinante im Zusammenhang mit den Lernprozessen von Schülerinnen und Schülern im Unterricht dar (Hattie, 2009). Unter Feedback werden Informationen verstanden, die Lehrkräfte Kindern zur Unterstützung ihrer Lernprozesse im Unterricht zukommen lassen (Hattie & Timperley, 2007). Mit dem Blick auf den aktuellen Forschungsstand wird Feedback als Unterstützungsmaßnahme von Mädchen und Jungen im Unterricht häufig differenzial wahrgenommen. Im Detail konnte dabei herausgestellt werden, dass Jungen in der Regel häufiger negatives Lehrkraftfeedback als Mädchen erhalten. Im Vergleich hierzu nehmen Mädchen häufiger positives Feedback wahr als Jungen (Chen, Thompson, Kromrey & Chang, 2011). Noch weitgehend ungeklärt ist, ob und inwiefern diese Geschlechterunterschiede in der Wahrnehmung von positivem und negativem Feedback im Allgemeinen zu beobachten sind oder sie in Abhängigkeit der einzelnen Unterrichtsfächer, des jeweiligen auf das Unterrichtsfach bezogenen Leistungsverhaltens sowie der jeweiligen Selbstkonzepte und Interessen variieren. Obschon über Unterschiede in den Wahrnehmungen von positivem und negativem Feedback in den einzelnen Unterrichtsfächern bei Mädchen und Jungen wenig bekannt ist, geht aus verschiedenen Untersuchungen – wenn auch sicherlich nicht vollständig einheitlich – hervor, dass Mädchen über signifikant höhere Leseleistungen sowie über signifikant höhere lesebezogene Selbstkonzepte verfügen als Jungen (vgl. Bos, Bremerich-Vos, Tarelli & Valtin, 2012; Valtin, Wagner & Schwippert, 2005). Umgekehrt zeigen Jungen signifikant höher ausgeprägte Leistungen in Mathematik und in den Naturwissenschaften sowie höher ausgeprägte mathematische Selbstkonzepte (Brehl, Wendt & Bos, 2012).

Vor diesem Hintergrund ist noch weitgehend ungeklärt, ob und inwiefern geschlechtsspezifische Unterschiede in den Wahrnehmungen von positiven und negativen Rückmeldungen bei Kindern im Grundschulalter vorliegen. Aufgrund des engen Zusammenhangs zwischen Feedback und Lernleistungen (Hattie, 2009) nehmen wir an, dass Mädchen im Grundschulalter signifikant häufiger positives und signifikant weniger negatives Feedback im Leseunterricht der Grundschule wahrnehmen als die Jungen. Diese Unterschiede werden vermutlich auch unter Kontrolle ihrer Leseleistungen, ihrer lesebezogenen Selbstkonzepte ebenso wie ihrer intrinsischen Lesemotivation bestehen bleiben.

In unserer Studie wurden  $N=682$  Kinder der dritten und vierten Jahrgangsstufe anhand eines 'Paper and Pencil'-Fragebogens gebeten, Stellung zu der Häufigkeit ihres wahrgenommenen positiven und negativen Feedbacks im Leseunterricht der Grundschule zu nehmen. Die Kinder beantworteten Fragen zu ihrer intrinsischen Lesemotivation sowie zu ihren lesebezogenen Selbstkonzepten. Darüber hinaus bearbeiteten sie den Leseverständnistest ELFE (Lenhard & Schneider, 2006). Die Auswertung der Studie erfolgte im Rahmen von univariaten zwei-faktoriellen Varianzanalysen. Dabei bildeten das durch die Kinder perzipierte positive und negative Feedback jeweils die abhängigen Variablen. Die Kinder wurden einerseits anhand ihres Geschlechts und andererseits anhand ihrer Leseleistung (leistungsschwach/ leistungsstark) in Gruppen unterteilt. Als Kontrollvariablen wurden die lesebezogenen Selbstkonzepte sowie die intrinsische Lesemotivation der Kinder bei den Analysen berücksichtigt.

Die ermittelten Ergebnisse verdeutlichen signifikante geschlechtsspezifische Unterschiede in den Wahrnehmungen des positiven und negativen Feedbacks im Leseunterricht. Während die an der Studie beteiligten Mädchen signifikant häufiger positives Feedback im Leseunterricht wahrnehmen als die Jungen, perzipieren die Jungen signifikant mehr negatives Feedback als die Mädchen. Signifikante Haupteffekte zeigen sich allerdings jeweils für die Leseleistungen. Bei der Berücksichtigung der Leseleistungen in den Analysen werden keine geschlechtsspezifischen Unterschiede in den Wahrnehmungen des positiven oder negativen Feedbacks im Leseunterricht mehr deutlich. Bei den Analysen zur Wahrnehmung des positiven Feedbacks lässt sich ein signifikanter Interaktionseffekt zwischen der Leseleistung und dem Geschlecht nachweisen. Die leistungsstärkeren Mädchen erfahren demnach mehr positives Feedback als die leistungsstärkeren Jungen. Die leistungsschwächeren Jungen bekommen aus ihrer Sicht deutlich weniger positives Feedback als die leistungsschwächeren Mädchen. Dieser signifikante Interaktionseffekt bleibt auch unter Kontrolle der lesebezogenen Selbstkonzepte und der intrinsischen Lesemotivation der Kinder erhalten.

Die Befunde aus unserer Studie geben erwartungsgemäß Hinweise darauf, dass auch unter der Kontrolle der Leseleistung geschlechtsspezifische Unterschiede in den Wahrnehmungen von positivem Feedback im Leseunterricht der Grundschule bestehen.

## E03

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S105

### Unterschiede in der ökonomiebezogenen Studieneingangskompetenz zwischen den deutschen Hochschultypen und Hochschulrankingsgruppen

**Carla Kühling-Thees<sup>1</sup>, Olga Zlatkin-Troitschanskaia<sup>1</sup>, Roland Happ<sup>1</sup>, Judith Jitomirski<sup>2</sup>, Jasmin Schlax<sup>1</sup>, Manuel Förster<sup>1</sup>, Sebastian Brückner<sup>1</sup>, Hans Anand Pant<sup>2</sup>**

<sup>1</sup>Johannes Gutenberg-Universität Mainz, Deutschland; <sup>2</sup>Humbolt-Universität zu Berlin; [c.kuehling-thees@uni-mainz.de](mailto:c.kuehling-thees@uni-mainz.de)

#### Theoretischer Hintergrund

Die Domäne der Wirtschaftswissenschaften gilt in Deutschland als der meiststudierte Studienbereich als Haupt- oder Nebenfach (Statistisches Bundesamt, 2017), und auch in anderen Fachbereichen wie den Rechts- und Ingenieurwissenschaften werden ökonomische Inhalte gelehrt (Siegfried & Walstad, 2014; Walstad et al., 2007; Hoyt & McColdrick, 2012). Mit den steigenden Studierendenzahlen nimmt auch die Heterogenität in den Eingangsvoraussetzungen bei StudienanfängerInnen zu, was u.a. im (fachrelevanten) Vorwissen erkennbar wird. Der Kompetenzzuwachs im Verlauf des Studiums hängt stark von den Eingangsvoraussetzungen der Studierenden ab. Dabei zeigt sich studienfachbezogenes Vorwissen als einer der bedeutsamsten Prädiktoren für den Studienerfolg (authors, 2015a; authors, 2016a; Owen, 2012). Ein weiterer Prädiktor für den Studienerfolg sind die allgemeinen kognitiven Eingangsvoraussetzungen wie die intellektuelle Leistungsfähigkeit (z.B. Coates & Friedman, 2010). Nach Abschaffung des ZVS-Verfahrens versuchen die Hochschulen zunehmend, die „besten“ Studierenden zu gewinnen. Dabei werden oftmals solche Annahmen unterstellt (Kramer et al., 2011), dass die Leistungsfähigkeit der StudienanfängerInnen im Vergleich der beiden in Deutschland vertretenen Hochschultypen – Fachhochschulen und Universitäten sowie auch nach den Hochschulrankingsgruppen (wie CHE-Ranking) sich systematisch differenzieren. Demnach würden Studierende an Universitäten über eine höhere ökonomiebezogene Studieneingangskompetenz (ÖSEK) als Studierende von Fachhochschulen verfügen (authors, 2015b); gleichbedeutend würde es für die hochgerankten Hochschulen versus niedrigergerankten Hochschulen gelten.

#### Forschungsfragen

Um die Annahmen zu überprüfen, ob sich die ÖSEK der StudienanfängerInnen in der Domäne Wirtschaftswissenschaften systematisch unterscheidet, präsentieren wir die Ergebnisse aus einer deutschlandweiten Erhebung (authors, 2016b). Wir fokussieren uns auf die Ergebnisse der Eingangsdiagnostik zur ÖSEK sowie zur allgemeinen intellektuellen Leistungsfähigkeit bei StudienanfängerInnen und auf folgende Forschungsfragen:

1. Gibt es Unterschiede in der ÖSEK der StudienanfängerInnen unter Berücksichtigung der verschiedenen Hochschultypen?
2. Gibt es Unterschiede in der ÖSEK der StudienanfängerInnen unter Berücksichtigung der verschiedenen Hochschulrankingsgruppen?
3. Lässt sich ein Zusammenhang zwischen den anderen kontrollierten Merkmalen der ökonomischen Vorbildung (wie abgeschlossene Berufsausbildung, Berufstätigkeit) und der allgemeinen intellektuellen Leistungsfähigkeit der StudienanfängerInnen und dem Hochschultyp erkennen?

#### Methode

Die Eingangsdiagnostik erfolgte in einer deutschlandweiten kontrollierten Feldbefragung im Wintersemester 2016/2017, an welcher 33 Universitäten und 21 Fachhochschulen in 10 Bundesländern teilnahmen. Dabei wurde die ÖSEK mittels der deutschen Adaption des TEL-IV (Walstad et al. 2013) in Kombination mit der deutschen Adaption des TUCE-IV erfasst (Walstad et al. 2007). Zur Erfassung der allgemeinen intellektuellen Leistungsfähigkeit wurde der BEFKI in einer Kurzform angewendet (Schipolowski et al. 2016). Insgesamt nahmen N = 7.640 StudienanfängerInnen der Wirtschaftswissenschaften an der Befragung teil. Die Durchschnittsnote der Hochschulzugangsberechtigung ist 2,35; 14.1 % haben eine abgeschlossene kaufmännische Ausbildung und 30.5 % einen wirtschaftsbezogenen Leistungskurs absolviert. Das Alter der Stichprobe liegt im Mittel bei 20,5 Jahren (SD = 2.7) und 54 % sind männlich.

#### Ergebnisse

Die Ergebnisse zeigen systematische Unterschiede in der ÖSEK sowie in der allgemeinen intellektuellen Leistungsfähigkeit der StudienanfängerInnen, und dass die ÖSEK signifikant mit dem Hochschultyp zusammenhängt; durchschnittlich ist die ÖSEK bei Universitätsstudierenden (M=13.62, SD=0.06) höher ausgeprägt als bei Fachhochschulstudierenden (M=12.12, SD=0.09). Dasselbe spiegelt sich auch in der intellektuellen Leistungsfähigkeit wieder: Universitätsstudierende (M=8.57, SD=0.04) und Fachhochschulstudierende (M=7.61, SD=0.05). Insgesamt lassen die Befunde annehmen, dass der Stand des (fachrelevanten) Vorwissens, der allgemeinen intellektuellen Leistungsfähigkeit sowie der ökonomischen Vorbildung die Wahl des Hochschultyps beeinflussen. Aus den Befunden lassen sich weitreichenden Implikationen für die eingangsdiagnostische Praxis beim Hochschulzugang ableiten, die kritisch diskutiert werden.

### Lernen am Arbeitsplatz im dualen Studium

**Alexander Brodsky<sup>1</sup>, Jürgen Seifried<sup>1</sup>, Gerald Sailmann<sup>2</sup>**

<sup>1</sup>Universität Mannheim, Deutschland; <sup>2</sup>Hochschule der Bundesagentur für Arbeit, Deutschland; [brodsky@bwl.uni-mannheim.de](mailto:brodsky@bwl.uni-mannheim.de), [seifried@bwl.uni-mannheim.de](mailto:seifried@bwl.uni-mannheim.de), [Gerald.Sailmann@arbeitsagentur.de](mailto:Gerald.Sailmann@arbeitsagentur.de)

Duale Studiengänge kombinieren Praxisphasen in Unternehmen mit dem Studium an (dualen) Hochschulen (Baethge & Wolter, 2015; Graf, 2013) und sind derzeit auf Expansionskurs (Krone, 2015; Leichsenring, König & Göser, 2015). Doch gerade das Lernen in der betrieblichen Praxis, das ein wichtiger Bestandteil dieser Studienmodelle ist, wurde im Kontext des dualen Studiums bisher nicht hinreichend untersucht. Deshalb soll in diesem Forschungsprojekt der Frage nachgegangen werden, wie sich Praxisphasen auf die professionelle Expertise von Studierenden auswirken. Dabei werden Bachelorstudierende der Hochschule der Bundesagentur für Arbeit über die gesamte Studiendauer hinweg begleitet.

Das Analysemodell der Untersuchung basiert auf einer Kombination von Input-Prozess-Output Modellen (z.B. das 3-P model of workplace learning; Tynjälä, 2013) und Prozessmodellen (z.B. das Job Characteristics Modells von Hackman & Oldham (1976)). Dabei ist zum einen der Prozess des Lernens am Arbeitsplatz selbst (u.a. in Abhängigkeit von Aufgabenmerkmalen) von Interesse. Zum anderen sind auch die Beziehungen zwischen Input-, Prozess- und Outputfaktoren relevant. Die Input-Komponente beinhaltet die individuellen Voraussetzungen des Lernenden (z.B. Studieninteresse). Die Prozessdimension setzt sich aus den Bestandteilen Lern- und Arbeitsbedingungen und der individuellen Verarbeitung des Lernenden zusammen. Es



wird angenommen, dass diese Bestandteile neben einem gegenseitigen Einfluss auch die Output-Komponente (Kompetenzentwicklung der Studierenden) beeinflussen.

In der Erhebung zu Beginn des ersten Praxistristesters wurden (1) Erwartungen an die Kompetenzentwicklung (in Anlehnung an Sailmann, 2014), (2) Lern- und Leistungsziele (Spinath et al., 2012), (3) das Studieninteresse (Schiefele et al., 1993), (4) die berufliche Selbstwirksamkeit (in Anlehnung an Schwarzer & Jerusalem, 1999) und (5) soziodemographische Daten erfragt. Während des Praxistristesters wurden dann fünf Lerntagebuch-Erhebungen durchgeführt. Das Lerntagebuch beinhaltete die Häufigkeit verschiedener Arbeitstätigkeiten sowie die Beschreibung und Einschätzung (Neuheit, Interessantheit, Lernmöglichkeit) der zwei zeitintensivsten Tätigkeiten in der zurückliegenden Arbeitswoche. In der Ausgangserhebung wurden (1) die Erfüllung der Erwartungen, (2) die Lern- und Arbeitsbedingungen (Morgeson & Humphrey, 2006; Rausch, 2012; Velten & Schnitzler, 2012), (3) die Wahrnehmung motivationsrelevanter und situationaler Bedingungsfaktoren (Prenzel et al., 1996), (4) die Erreichung der Lern- und Leistungsziele, (5) die Nutzung von Studieninhalten aus den Hochschulphasen sowie (6) die Gesamtzufriedenheit und der wahrgenommene Nutzen der Praxisphase erfasst. Sämtliche Erhebungen wurden auf der Lernplattform ILIAS durchgeführt. Bei insgesamt 483 eingeschriebenen Studierenden lag die Beteiligung an den freiwilligen Erhebungen im ersten Praxistristester zwischen 208 und 66 Studierenden, die um Durchschnitt 22,3 Jahre alt und zu 79 Prozent weiblich waren.

Die meisten Studierenden berichten zu Beginn der Praxisphase über ein großes Interesse an ihrem Studienfach und die berufliche Selbstwirksamkeit ist stark ausgeprägt. Auch die Lern- und Leistungsziele sowie die Erwartungen an die eigene Kompetenzentwicklung sind hoch.

Die Ergebnisse der Lerntagebücher zeigen dann, dass der Anteil selbst durchgeführter Tätigkeiten im Verlauf der Praxisphase ansteigt und die Studierenden sukzessive über mehr Eigenaktivitäten berichten. Dennoch waren über das gesamte Praxistristester hinweg Formen der Hospitation dominant. Zudem arbeiteten die Probanden an ihrem Praktikumsbericht (Prüfungsleistung der Hochschule).

Die Ausgangserhebung am Ende des Praxistristesters zeigt schließlich, dass die Erwartungen der Studierenden überwiegend erfüllt wurden. Die Studierenden schätzen die Lern- und Arbeitsbedingungen als überwiegend gut ein (z.B. KollegInnen), auch wenn einige Skalen nur Mittelwerte im mittleren Skalenbereich erreichen. Auch die motivationsrelevanten Bedingungen werden als gut (soziale Eingebundenheit, Autonomie) bis mittelmäßig (Kompetenzunterstützung) bewertet. Der Gesamtnutzen und die Zufriedenheit mit dem Praxistristester wurden von einem großen Teil der Studierenden als relativ hoch eingeschätzt.

Mit den Ergebnissen des Projekts sollen Ansatzpunkte zur Professionalisierung von Praxisphasen in dualen Studiengängen und zur Abstimmung zwischen den Ausbildungsphasen an der Hochschule sowie der Praxis gewonnen werden.

## **Lehrerinnen- und Lehrerfortbildung in Baden-Württemberg. Analyse des Gesamtangebots**

**Karen Johannmeyer, Colin Cramer, Martin Drahm**

Universität Tübingen, Deutschland; [karen.johannmeyer@uni-tuebingen.de](mailto:karen.johannmeyer@uni-tuebingen.de)

### **THEORETISCHER HINTERGRUND**

Die Forschung zur »Dritten Phase« kann im Vergleich zu Studien im Umfeld zur Ersten und Zweiten Phase der Lehrerbildung als marginalisiert gelten, wenngleich weitestgehend Einigkeit darüber besteht, dass den Fort- und Weiterbildungsangeboten ein zentraler Stellenwert für die Qualität des Unterrichts zukommt (Lipowsky, 2014). Konform zum Forschungsstand ist einerseits zu konstatieren, dass die Fort- und Weiterbildungsangebote bislang kaum systematisch und verbindlich gestaltet sind und darüber hinaus – im Gegensatz zur Ersten und Zweiten Phase – keine allgemeingültigen Standards vorliegen (Daschner, 2009). Dennoch liegen verschiedene Rahmenmodelle zur Evaluierung und Wirksamkeitsmessung vor, wie das erweiterte Angebots- und Nutzungsmodell zur Erklärung des beruflichen Lernens von Lehrpersonen im Rahmen von Fort- und Weiterbildungen (Lipowsky, 2014). Die empirische Forschung weist auf eine Vielzahl von Gelingensbedingungen für wirksame Angebote hin (Desimone, 2009). Einerseits haben „kurzzeitige One-Shot-Fortbildungen“ (Lipowsky, 2014, S. 518) kaum Wirkung auf den Lernerfolg der Lehrkräfte, andererseits kann aber u.a. die subjektiv wahrgenommene Relevanz der Inhalte oder das Lernen in professionellen Gemeinschaften positive Effekte auf die Teilnehmenden und ihr berufliches Handeln haben (Caena, 2011; Lipowsky, 2010), und es zeigen sich Zusammenhänge zwischen einer hohen Fortbildungsteilnahme und einer höheren Selbstwirksamkeit und Kooperation im Kollegium (Richter u.a., 2013).

### **FRAGESTELLUNG**

Ausgehend vom Mangel an einer vollständigen Analyse des Angebots der Lehrerinnen- und Lehrerfortbildung in einem klar abgegrenzten Lehrerbildungssystem im deutschsprachigen Raum, erscheint eine differenzierte Analyse der Angebote von Bedeutung. Nur unter Kenntnis des real existierenden Angebots können evidenzbasiert Aussagen über die Ausbildungsstruktur getroffen werden, die vor dem Hintergrund der gegenwärtigen Reformbestrebungen der Bildungspolitik von hoher Relevanz erscheinen. Die Studie geht daher der Frage nach, wie sich das Angebot der Lehrerinnen- und Lehrerfortbildung in Baden-Württemberg darstellt: Welche Zusammenhänge zeigen sich im Angebot etwa bezüglich der Inhalte, des Fachbezugs, der Schulart, der Zielgruppe und der Region? Zusätzlich wird explorativ untersucht, inwiefern sich Ungleichgewichte hinsichtlich dieser Merkmale in der Angebotsstruktur zeigen.

### **METHODEN**

Als Stichprobe dient das gesamte Fortbildungsangebot der Landesakademien, Regierungspräsidien und Schulämter, das in einer Online-Datenbank indiziert ist. Die Stichprobe umfasst N=10588 angebotene Fortbildungen, die über ein gesamtes Jahr hinweg angeboten wurden (01.07.2016 bis 30.06.2017). Die Angebote werden zunächst mittels Inhaltsanalyse (Krippendorf, 2013) kriteriengeleitet kodiert. Im Anschluss erfolgt eine differenzierte Analyse, in der deskriptive Analysen sowie Zusammenhangs- und Regressionsanalysen zur Untersuchung kategorialer Daten (z.B. Logistische Regression: Eid et al. 2015, S. 799-839) durchgeführt werden. So werden Zusammenhänge zwischen Variablen wie beispielsweise Region, Schulart oder Fortbildungsinstitution und dem fachlichen Bezug der Angebote aufgezeigt. Aufgrund der Vollerhebung kann so ein repräsentatives Bild von der Angebotsstruktur der Lehrerinnen- und Lehrerfortbildung gegeben werden.

### **ERGEBNISSE und DISKUSSION**

Die Ergebnisse der Analysen zeigen ein differenziertes Bild der Fortbildungsangebote in Baden-Württemberg. So zeigt sich ein starkes Ungleichgewicht in Bezug auf die Anzahl der Angebote in verschiedenen Regionen. Weitere Ungleichgewichte lassen sich für die Schulfächer feststellen, für die Fortbildungen durchgeführt werden, sowie für die Schulformen. Weitere Analysen tragen dazu bei, Zusammenhänge zwischen den Merkmalen erkennen zu können. So zeigt sich z.B. ein Mangel an Angeboten in bestimmten Fächern in Abhängigkeit von der Region oder von der Zielschulart oder die Abhängigkeit fachbezogener

Fortbildungen von Schularten und Fortbildungsinstitutionen. Die zu gewinnenden Erkenntnisse sind von Bedeutung für Überlegungen zu einer Optimierung des Fortbildungsangebots. Sie ergänzen mit weitreichenden Strukturinformationen die bislang vorliegenden Erkenntnisse zu Gelingensbedingungen einzelner Fortbildungsangebote. Erstmals wird mit dem Vorhaben eine Vollerhebung und systematische Analyse des Fort- und Weiterbildungsangebots in Baden-Württemberg durchgeführt. Da bislang keine vergleichbaren Strukturinformationen erhoben wurden, kommt dem Projekt eine Pilotfunktion zu. Das methodische Vorgehen und die basalen gewonnenen Erkenntnisse könnten ein Modell für vergleichbare Forschung in anderen (Bundes-)Ländern sein oder einen Ausgangspunkt für korrespondierende Forschung darstellen.

## **Anlässe und Barrieren von Randgruppen in beruflichen Weiterbildungen**

**Christin Siegfried<sup>1</sup>, Eveline Wuttke<sup>1</sup>, Susan Seeber<sup>2</sup>**

<sup>1</sup>Goethe-Universität Frankfurt, Deutschland; <sup>2</sup>Georg-August-Universität Göttingen, Deutschland; [Siegfried@em.uni-frankfurt.de](mailto:Siegfried@em.uni-frankfurt.de), [wuttke@em.uni-frankfurt.de](mailto:wuttke@em.uni-frankfurt.de)

### Theoretischer Hintergrund

Auf dem deutschen Arbeitsmarkt ist in den kommenden 20 Jahren aufgrund des technologischen und demografischen Wandels und des daraus resultierenden Fachkräftemangels mit Engpässen bei der Besetzung von Stellen zu rechnen (Calo, 2008). Insbesondere die Branchen Gastronomie, Gesundheitswesen, Handel und Technologie werden betroffen sein (Helmrich et al., 2012). Auf der anderen Seite gibt es nicht erschlossenes Potenzial von Arbeitnehmern, speziell in bestimmten Randgruppen, um die sich Unternehmen bemühen und sie mit gezielten Weiterbildungen zu Fachkräften ausbilden müssen, um damit den eigenen Erfolg auch in Zukunft sichern zu können. Der Begriff Randgruppe bezieht sich auf Personen mit Einschränkungen in der Teilnahme an verschiedenen Lebensbereichen, z.B. Ausbildung, Einkommen, soziale Positionierung und Beteiligung (Trostel, Walker & Woolley, 2002). In unserer Studie beziehen wir uns auf fünf Randgruppen (Seeber, Wuttke, Rosemann, 2017): (1) Niedrig oder nicht formal Qualifizierte (Schulte-Braucks, 2013; Solga, 2005; Krenn, 2010), (2) Arbeitnehmer mit Migrationshintergrund (Autorengruppe Bildungsberichterstattung, 2016), (3) Wiedereinsteiger in den Beruf (Allmendinger, 2010), (4) ältere Arbeitnehmer (Gillen, Elsholz & Mayer, 2010) und (5) Arbeitnehmer mit Migrationshintergrund mit gleichzeitiger niedrig Qualifizierung. Diese Gruppen sind in der als notwendig erachteten Weiterbildung stark unterrepräsentiert, betrachtet man jedoch den Stand der Forschung, wird deutlich, dass wenig bekannt ist, welche Gründe für bzw. Argumente gegen die Weiterbildungsteilnahme bestehen.

### Forschungsfragen

Ziel des vorliegenden Beitrags ist es deshalb, Anlässe und Barrieren für die Teilnahme an beruflichen Weiterbildungen von Randgruppen in den Branchen Gesundheitswesen, Sozialdienstleistungen, Hotel- und Gastronomie, Handel und Technologie zu ermitteln. Hierfür sind drei zentrale Fragestellungen zu beantworten:

- (1) Welche individuellen hinderlichen und förderlichen Faktoren für eine Weiterbildungsteilnahme lassen sich bei den betrachteten Zielgruppen identifizieren?
- (2) Inwiefern bestehen Unterschiede zwischen den Gruppen?
- (3) Welche individuellen Einflussfaktoren erklären Unterschiede in der gruppenspezifischen Nennung von Anlässen und Barrieren?

### Methode

Für die Beantwortung der Fragestellungen wurden zwischen September 2016 und Juni 2017 Leitfadenterviews sowohl mit Arbeitnehmern (N=75 davon (1)=7, (2)=11, (3)=17, (4)=24, (5)=17), als auch Arbeitgebern (N=53) durchgeführt. In diesem Beitrag konzentrieren wir uns auf die Arbeitnehmerperspektive. Neben Interviews wurden standardisierte Fragebögen zur Intelligenz (Baudson & Preckel, 2015), Selbstwirksamkeit (Rigotti, Schyns & Mohr, 2008) und Zielorientierung (Spinath, 2014) eingesetzt und analysieren, inwieweit diese und weitere individuelle Faktoren wie Geschlecht, Bildungshintergrund sowie Alter die von den Interviewteilnehmern genannten Gründe und Barrieren beeinflussen.

Alle Interviews wurden professionell transkribiert und anschließend mithilfe eines Kodierleitfadens ausgewertet. Um die Qualität der Kodierungen zu gewährleisten, wurden zunächst 10% der Interviews in einem Team von drei Kodierern, anschließend weitere 30% durch Kodiererpaare (alle Kodierer können miteinander verglichen werden) kodiert. Erst nachdem sich eine kontinuierlich gute Interraterreliabilität eingestellt hatte (Cohen's Kappa von .70 - .85), wurden Kodierungen durch einzelne Kodierer zugelassen.

### Ergebnisse:

Insgesamt nannten die Befragten 22 verschiedene Anlasstypen und 14 verschiedene Barrieretypen für ihre Weiterbildungsteilnahme. Dabei scheinen Anlässe wie „berufliche Veränderungen“, ein „generelles, inhaltliches Interesse oder Spaß am Thema“ und die „Aufrechterhaltung nicht unmittelbar berufsbezogenen Wissens“ für alle Randgruppen eine ähnlich hohe Bedeutung zu haben (50-95% aller Interviewpartner nennen diese Kategorien), während andere Anlässe nur von spezifischen Randgruppen Erwähnung finden. So wird das „berufliche Networking“ vor allen Dingen von der Randgruppe (3) und (5) genannt, während die „Teilnahmepflicht“ eher für die Gruppe (1) und (4) eine Rolle zu spielen scheint.

In Bezug auf die Barrieren für die Teilnahme an beruflichen Weiterbildungen wird gruppenübergreifend vor allen Dingen „konkurrierende Verpflichtungen“ und „angebotsbezogene Barrieren“ genannt (Nennung von 40 bis 75% aller Interviewpartner). Andere Barrieren wie „Barrieren durch den Arbeitgeber/Veranstalter“ oder „physische/psychische Erkrankungen“ werden nur von der Randgruppe (5) und (3) genannt.

Unterschiede in der Nennung der Anlässe und Barrieren lassen sich vor allen Dingen durch die kognitive Fähigkeit (Intelligenz) erklären (33% Varianzaufklärung), das Geschlecht, die Ausprägung der Selbstwirksamkeit oder Zielorientierung scheinen keine Rolle zu spielen.

## E04

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Rosshofgasse, S02

### New methodological approaches to studying acquisition of additional languages: An empirical study with third-age learners

**Simone Pfenninger**

Universität Salzburg, Österreich; [simone.pfenninger@sbg.ac.at](mailto:simone.pfenninger@sbg.ac.at)

This paper deals with cutting-edge conceptualisations and empirical research techniques for studying second language acquisition (SLA) in different educational settings. Recent decades have seen a “social turn” and an “individual turn” in SLA, which capture and capitalize on individuality in the language learning process that is constantly being shaped by and shaping the social environment (Block, 2003; MacIntyre, 2014). When nonlinear system dynamics was introduced into SLA research – under various rubrics such as chaos theory (Larsen-Freeman, 1997), emergentism (Ellis & Larsen-Freeman, 2006), dynamic systems theory (de Bot et al., 2007) and complexity theory (Larsen-Freeman & Cameron, 2008) – the new approach fundamentally questioned the feasibility of investigating (linear) cause-effect relations and macro-developments, the traditional bases of generalizable theories. This notion fundamentally challenges empirical conceptualisations of SLA processes: the question arises as to how we can reconcile large-sample methods with individual-level analyses. If we aim to maximize the generalizability of our findings, to what extent do we lose sight of variability?

In this paper I examine these challenges from the viewpoint of classroom research that has focused on the identification of age effects and their interaction with contextual, personal and affective variables. Specifically, I present the methodological design and theoretical rationale of a longitudinal project called “Dynamic analysis of language learning in the third age (DYNAGE3)”, which explores how older adults learn a new language in a school setting and how good – linguistically and neurocognitively speaking – additional language learning is for THEM. In the pilot studies of this project 26 German-speaking subjects, half of them bilinguals, between 65 and 90 years of age attended four-week intensive English courses for beginners, and were tested on a range of linguistic, cognitive, neurofunctional and socio-affective parameters. The relevant instruments included: vocabulary and listening tests, working memory tasks, EEG experiments, and interviews about (linguistic) self-confidence, L2 learning motivation, general communicative skills and overall well-being. These were administered before, during and after the training, with a view to identifying factors that facilitate successful learning in the third age and uncovering inter-individual differences that are age-independent. The results of quantitative and qualitative analyses suggest that (1) the learning of an additional language in the third age can be a promising way of contributing to healthy and active ageing, since not only was cognitive fitness shown to play a decisive role in the learning outcome in question, but, contrariwise, L2 learning had a positive effect on executive functions; (2) age-related social, psychological and contextual factors seemed to play as significant a role as strictly maturational factors.

Such research promises in due course to inform adult educators about the establishment of relevant third-age learner profiles as well as the development of appropriate teaching materials and the design of individualized third-age language training. Furthermore, I will outline the advantages of employing statistical analyses such as multilevel models and generalized additive models that (1) allow for the simultaneous generalization of the results to new items and new participants, (2) take account of non-linear L2 learner trajectories and (3) run on unaveraged data and thus work around the problem of discretizing continuous variables such as age, bilingualism, or L2 proficiency. Such individual-centered statistical approaches offer a viable alternative to the group-based methodology that examines the relationship between well-defined variables in relative isolation (as in ANOVA-type analyses), which has traditionally dominated classroom research in SLA.

### Einflüsse der schweizerdeutschen dialektalen Familiensprache auf Ergebnisse eines Deutschtests auf universitärem Niveau?

**Karin Gehrler<sup>1</sup>, Maren Oepke<sup>2</sup>, Franz Eberle<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für Bildungsverläufe (LifBi), Deutschland; <sup>2</sup>Universität Zürich Institut für Erziehungswissenschaft Abteilung Lehrerinnen- und Lehrerbildung Maturitätsschulen; [karin.gehrer@lifbi.de](mailto:karin.gehrer@lifbi.de)

#### Theoretischer Hintergrund

Seit der Evaluation der Maturitätsreform 1995 EVAMAR II wissen wir empirisch gesichert, dass die Schweizer Maturandinnen und Maturanden an der Schnittstelle Gymnasium – Universität eine auffallend breite Streuung der getesteten Leistungen und Fähigkeiten aufweisen, mit teilweise substanziellen Defiziten (Eberle et al., 2008). Die für den Matura-Abschlussjahrgang 2007 repräsentative Studie belegte für die Deutschschweiz und die Romandie, dass deren Maturandinnen und Maturanden eine ausgesprochen breite Varianz von Kompetenzen in den getesteten Bereichen Mathematik, Biologie und Deutsch bzw. Französisch aufwiesen. Aus den Befunden lässt sich folgern, dass eine beachtliche Minderheit der möglichen StudienanfängerInnen für ein künftiges Studium an den Universitäten nicht die notwendigen sprachlichen Eingangskompetenzen mitbringt (Eberle et al., 2008, 374, 383-384). Diese sprachlichen Lücken stellen eine Herausforderung für die Bildungslandschaft Schweiz, insbesondere deren handelnde Akteure in Gymnasium und Universität, dar.

Das Plurizentrik-Paradigma (Ammon 1995; Davies, Häcki Buhofer, Schmidlin, Wagner & Wyss, in press; zur Gegenposition Koller 1999) geht davon aus, dass für viele Sprachen und so auch für das Deutsche unterschiedliche Sprachzentren mit eigenen Sprachnormen bestehen, diese sich relativ unabhängig voneinander entfalten und für ihre regional beschränkten Sprachmitglieder prägend und verbindlich sind. Keine der so regional definierten Sprachnormen ist jedoch grundsätzlich der anderen überlegen, sondern nur im eigenen Sprachraum durch die Normsetzung führend bzw. bindend. In der Deutschschweizer Sprachregion wachsen die Kinder meist mit der dialektalen Familiensprache Schweizerdeutsch auf. Hauptsächlich der Institution Grundschule obliegt die Aufgabe, die dialektal geprägten Kinder an den Gebrauch der deutschen Standardsprache heranzuführen.

#### Fragestellung

Die sprachlichen Leistungsunterschiede der schweizerischen MaturandInnen werden in diesem Beitrag vor dem Hintergrund der sprachwissenschaftlichen Plurizentrik-Debatte daraufhin geprüft, ob sie gegebenenfalls von der Zugehörigkeit zu der einen oder anderen „Sprachgruppe“ abhängen oder beeinflusst sind. Die Fragestellung lautet, ob zwischen der Gruppe der MaturandInnen mit Schweizerdeutsch bzw. Schweizerdeutscher Standardsprache als Familiensprache (Gruppe 1, n=1527) und den Gruppen der MaturandInnen mit Bundesdeutscher Standardsprache eines Elternteils (Gruppe 2, n=188) respektive beider Eltern (Gruppe 3, n=44) beim Ablegen einer Kompetenztestung in Deutsch auf dem Niveau von universitären Anforderungen starke Leistungsunterschiede auftreten oder nicht. Die Vermutung besteht, dass mit dem Schweizerdeutschen Dialekt

aufgewachsene SchülerInnen benachteiligt sind, weil sie durch ihre mündliche dialektale Familiensprache in einer Form sozialisiert werden, die nicht der universitären Sprachnorm entspricht.

#### Methode

Die Hypothese wird anhand des bei der EVAMAR II-Studie durchgeführten rasch-skalierten Erstspracheleistungstests mit den Subskalen „Detail-Leseverständnis“, „Allgemeines Leseverständnis“, „Wortschatz“ und „Grammatik/Orthographie“ überprüft. Zur Ermittlung der drei Gruppen wurden mittels Hilfsvariablen die Angaben zur Familiensprache mit Auskünften zu den Geburtsländern der Familienmitglieder kombiniert, weil keine direkten Befragungsdaten zur Sprachsozialisation zur Verfügung stehen. Zur Prüfung der Hypothese wurden univariate Varianzanalysen mit dem Faktor „Gruppenzugehörigkeit“ und den abhängigen Variablen der Leistungstestergebnisse durchgeführt. Als zusätzlicher Faktor wurde die berufliche Stellung der Eltern in den Varianzanalysen kontrolliert (in der Gruppe 3 sind überzufällig häufig Eltern mit Hochschulabschluss anzutreffen).

#### Ergebnisse

Für alle sprachlichen Subskalen des EVAMAR II-Erstsprachetests zeigen sich ebenso wie beim Gesamttest ( $F(2, 1756) = .515$ ,  $p = .598$ , Effektstärke  $\eta^2 = .001$ ) keine signifikanten Effekte der Gruppenzugehörigkeit aufgrund der unterschiedlichen Familiensprachen.

Unter dem Vorbehalt, dass die Gruppen relativ klein sind, und unter der Annahme, dass die Gruppenzugehörigkeit die Sprachgewohnheiten Schweizerdeutsch (Dialekt) vs. Hochdeutsch/Standarddeutsch in den Familien adäquat widerspiegeln, spielt es unseren Ergebnissen zufolge für das Abschneiden der MaturandInnen beim Deutsch-Leistungstest keine Rolle, ob in den Familien überwiegend Hochdeutsch/Standarddeutsch oder Schweizerdeutsch gesprochen wird.

Offensichtlich gelingt es entweder den Gymnasien, die Norm der deutschen Standardsprache für alle gleichermaßen als Orientierungspunkt gut zu vermitteln, so dass die sprachliche familiäre (dialektale) Alltagsgewohnheit weder eine hinderliche noch eine förderliche Rolle zu spielen scheint, oder andererseits gelingt es den im dialektalen Schweizerdeutsch aufgewachsenen Jugendlichen insgesamt als Gruppe, ohne Leistungseinbußen gegenüber anderen Sprachgruppen, die Sprachanforderungen des universitären Standarddeutschen zu erfüllen.

### **Performanznahe Messung von Deutsch-als-Zweitsprache (DaZ)-Kompetenz bei (angehenden) Lehrkräften: Effekte von mündlichen versus schriftlichen Antwortformaten – eine experimentelle Studie**

**Svenja Lemmrich, Anika Bahls, Timo Ehmke**

Leuphana Universität Lüneburg, Deutschland; [svenja.lemmrich@stud.leuphana.de](mailto:svenja.lemmrich@stud.leuphana.de)

Ziele oder Fragestellungen im Kontext des theoretischen Rahmens und Forschungsstandes

Um Schülerinnen und Schüler mit nichtdeutschen Herkunftssprachen angemessen fördern zu können, müssen Lehrkräfte die nötigen Kompetenzen aufweisen, ihren Fachunterricht sprachsensibel zu gestalten (vgl. Köker, Rosenbrock-Agyei, Ohm, Carlson, Ehmke, Hammer, Koch-Priewe & Schulze 2015). Diese Kompetenzen wurden bereits in einem theoretischen Kompetenzmodell für DaZ-Kompetenz modelliert. Basierend darauf wurde ein Paper-Pencil-Test entwickelt, der die Messung der Facetten der DaZ-Kompetenz ermöglicht (vgl. Köker et al. 2015). Ziel ist nun eine performanznahe Messung der DaZ-Kompetenz, um auch Kompetenzen auf Expertenniveau abzubilden. Um neben kognitiven Facetten von Kompetenz, auch Handlungsfacetten messbar zu machen, sind methodisch innovative Testverfahren erforderlich, die videobasierte Testinstrumente nutzen, da diese authentischer sind und situationsspezifisches Wissen abfragen (vgl. Gold & Holodynski 2017). Das BMBF-Projekt DaZKom-Video setzt an dieser Stelle an. Im Rahmen einer computerbasierten Testumgebung wird (angehenden) Lehrkräften ein videobasierter Aufgabenstimulus vorgelegt, welcher dann eine spontane, mündliche Antwort fordert. Dabei ist die Annahme, dass durch das mündliche Antwortformat, eine spontanere und performanznahe Messung realisiert werden kann.

#### Fragestellung

In dieser experimentellen Studie soll folgende zwei Fragestellungen untersucht werden:

- Welche Effekte hat das Antwortformat, schriftlich versus mündlich, auf die Messung von DaZ-Kompetenz bei (angehenden) Lehrkräften in Bezug auf die Korrektheit und die Ausführlichkeit der Antworten?
- Inwieweit lassen sich dabei differentielle Effekte für unterschiedliche Personenmerkmale feststellen (u.a. Geschlecht)?

#### Methodik

Um die Fragestellungen zu beantworten, wurden in einer experimentellen Studie mit randomisierter Stichprobenzuweisung zwei Varianten von computerbasiert dargebotenen Tests gegeneinander getestet. Beide Testvarianten erfassen die DaZ-Kompetenz bei (angehenden) Lehrkräften und bestehen aus jeweils drei Videostimuli mit jeweils drei Items. Neben den gemeinsamen Itemstimuli und den gleichen Fragestellungen der Items diente das variierte Antwortformat (schriftlich versus mündlich) als between-subject-factor. Die Stichprobe ( $N = 90$ ) wurde aus Seminaren des Lehramtsstudiums generiert (B.A./M. Ed, alle Semester). Insgesamt liegen 48.9 % mündliche Antworten und 51.1 % schriftliche Antworten vor. Die Antworten werden mithilfe einer Kodieranweisung in Doppelkodierung kodiert und ausgewertet. Lerngelegenheiten und Personenmerkmale wurden in einem Fragebogen abgefragt.

#### Ergebnisse bzw. Schlussfolgerungen

Die Ergebnisse zeigen, dass das Antwortformat keinen Effekt auf die Korrektheit der Antworten hat. Jedoch gibt es Unterschiede in der Ausführlichkeit: Mündliche Antworten sind durchschnittlich ausführlicher als schriftliche Antworten ( $MW = 67$  Wörter schriftlich versus  $MW = 99$  Wörter mündlich,  $p < .05$ ). In Bezug auf die Gruppenunterschiede nach Geschlecht gibt es keine statistisch signifikanten Unterschiede in der Korrektheit und der Ausführlichkeit ( $MW$  Studentinnen = 80.41 Wörter /  $MW$  Studenten 82.24 Wörter, ( $p = n.s.$ )). Zusammenfassend lässt sich festhalten: Eine performanznahe Kompetenzmessung durch ein mündliches Antwortformat führt demnach zu einer ausführlicheren Antwort. Sie führt aber nicht dazu, dass ein nachweisbarer Mode-Effekt im Hinblick auf das zu messende Konstrukt (in dieser Studie DaZ-Kompetenz) entsteht.

### **Linguistic Analysis of Multilingual Learning Data of Students in Higher Education**

**Alexander Mehler<sup>3</sup>, Olga Zlatkin-Troitschanskaia<sup>1</sup>, Wahed Hemati<sup>3</sup>, Dimitri Molerov<sup>2</sup>, Andy Lücking<sup>3</sup>, Susanne Schmidt<sup>1</sup>**

<sup>1</sup>Johannes Gutenberg-Universität Mainz, Deutschland; <sup>2</sup>Humboldt-Universität zu Berlin; <sup>3</sup>Goethe-Universität Frankfurt am Main; [plato@uni-mainz.de](mailto:plato@uni-mainz.de), [plato@uni-mainz.de](mailto:plato@uni-mainz.de), [susanne.schmidt@uni-mainz.de](mailto:susanne.schmidt@uni-mainz.de)

## Theoretischer Hintergrund

The influence of language and phrasing of task description on students' performance in educational assessments has long been discussed (Solano-Flores et al. 2006). This is particularly important regarding tests that assess the same construct with different languages, e.g. as illustrated in the comparison of the TUCE (Walstad et al. 2007) between the US, Japan, and Germany (Authors 2014, 2015a). Our study focuses on results showing language effects that could not be explained from an educational research perspective alone (Authors 2015a,b; 2016b), calling for a more in-depth analysis of the language in the test.

Empirical examination of the effects of language features on educational assessment is still scarce (Bühner&Pargent 2017). Few recent international studies examining item surface features and their effects on performance in specific domains, such as economics, have shown that students from different countries perform differently on certain tasks, even when personal and institutional influence factors and DIFs are controlled (Authors 2015a; 2016a,b).

There are various explanatory approaches for these differences in performance, which are still underresearched. One explanation focuses on students' knowledge acquisition history and experiences with applying subject-specific concepts in pre-tertiary education, situated in a specific culture and language (Rüede 2015). A second explanation foregrounds differences in cognitive load (Sweller et al. 2011) arising from complexity of language structures in test items and students' familiarity with them.

## Fragestellung

Based on the theoretical considerations, we assume that one can predict the rates of correct student responses by analyzing quantitative features of the test tasks' verbal presentations, and we focus on two questions:

- (i) To what extent do certain linguistic features in the German and English TUCE versions correlate with the test results?
- (ii) Does the correlation allow for characterizing each language version by distinct patterns of linguistic features?

## Methode

TUCE data is structured in terms of questions and corresponding response options. As a necessary prerequisite for explaining any language effects in an empirically substantiated way, we carry out an in-depth analysis of linguistic features of the TUCE. Analyzing these questions requires a new quantitative model that includes various linguistic variables correlating with the rates of correct responses (our target variable). To this end, we developed and tested a new linguistic model of task descriptions for the German and English TUCE in order to generate quantitative representations of specific linguistic features (including content-related characteristics) in both languages and phrasing of test items that can be related to the target variable.

The German TUCE data included 1,629 students from 33 universities (Authors 2015b). The US data included 4,400 students from 51 colleges. In both countries, the test was administered to economics students at the beginning of their 1st, 2nd, and 3rd year of undergraduate studies.

The analyses included mostly automatically annotated features and few manually annotated features, such as domain-specific terms and quantitative vocabulary. Language patterns in test tasks were examined in detail using methods from computational linguistics, and relations between the specific language features and the test-takers' response distributions on the multiple-choice test items were analyzed by calculating distance correlations for each item.

## Ergebnisse

Our findings indicate a strong correlation between linguistic features of task descriptions and rates of correct responses on these test tasks. This correlation gets remarkably higher if feature subsets are considered in contrast to single features. This assessment is also supported by the Monte-Carlo test according to which we can speak of significant correlations. This indicates a certain statistical interdependence among the analyzed linguistic features making the target variable, more predictable than in cases where single features are considered. The practical meaning of correlations and implications will be critically discussed in our presentation.

## E05

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Rosshofgasse, S01

### „Sekundeninterventionen“ – LehrerInnenhandeln in Momenten der Unbestimmtheit

**Katharina Rosenberger**

Kirchliche Pädagogische Hochschule Wien/Krems, Österreich; [katharina.rosenberger@kphvie.ac.at](mailto:katharina.rosenberger@kphvie.ac.at)

Das unterrichtliche Handeln von Lehrpersonen kann zwar in seinen Grundzügen vorbereitet werden, sein konkreter Verlauf und seine vielfachen Wirkungen sind aber in wesentlichen Teilen weder vorhersehbar noch gänzlich kontrollierbar. Eine Lehrperson kann etwa den sachlogischen Aufbau einer Unterrichtsstunde planen, gezielte Aufgabenstellungen dazu wählen, sich überlegen, in welcher Art die SchülerInnen wie lange einen Lernstoff erarbeiten sollen – wie sie dies in der Unterrichtsstunde aber dann tatsächlich zur Ausführung bringt und wie die SchülerInnen agieren werden, bleibt letztlich immer eine kontingente Angelegenheit. Wir können dazu davon ausgehen, dass das unterrichtliche Handeln zum Teil widersprüchlich und nicht geradlinig verläuft und pädagogische Interaktionen in ihrer Einzigartigkeiten von Unbestimmtheiten gekennzeichnet sind. Dieser Herausforderung begegnen LehrerInnen auf Basis eines Konglomerats aus erworbenen Fertigkeiten, von Wissen, Erfahrungen und stabilisierenden Routinen – eingebettet in kontextgebundene Bedeutungen, in kulturell bedingte Ordnungsmuster in und sozial ausgehandelte Regeln. Das Wirken der Lehrenden ist also Teil eines komplexen Geschehens mit bestimmten materiellen Konstellationen und sozialen Positionierungen. Darüber hinaus enthält es, da es nicht nur ein routinehaftes Handeln ist, improvisierende und kreative Elemente. Was in welcher Weise zu deuten ist, was am besten zu tun wäre und wie dies zu tun ist ‚entscheidet‘ eine Lehrperson im Augenblick und im Handeln.

Für die empirische Forschung ist dieser häufig implizite, praktische und personengebundene Charakter sozialen (unterrichtlichen) Handelns sowie die situative Verankerung eine besondere Herausforderung und erfordert Methoden und konzeptionelle Zugänge, die den Blick auf die meist diskreten und von den AkteurInnen oftmals nur peripher wahrgenommenen Prozesse unterstützen. Im Zuge eines empirisch-qualitativ ausgerichteten Forschungsprojekts, das sich auf Basis praxistheoretischer Ansätze dem Zusammenspiel zwischen AkteurInnen und Artefakten in Unterrichtspraktiken widmet (vgl. Rosenberger 2015), wurden Unterrichtsvideografien auf mikrosoziologischer Ebene ausgewertet, um der Situativität des LehrerInnenhandelns gerecht werden zu können. In Anlehnung an die von Dinkelaker und Herrle (2009) dargelegte Segmentierungsanalyse wurde ein Videoanalyseverfahren entwickelt, um die unterrichtlichen Prozesse als Teil der schulischen Alltagskultur rekonstruieren zu können. Das Auswertungsverfahren zielte darauf ab, den Ablauf der Interaktion zwischen den Handelnden in verschiedene Sequenzen, z.B. Veränderungen in den Bezügen zwischen Raum und Körper, in Mustern des SprecherInnenwechsels oder im Wechsel der Themen herauszuarbeiten.

Im Tagungsbeitrag soll anhand der praxistheoretischen Interpretation eines kurzen Videobeispiels die methodische Herangehensweise diskutiert werden, durch die gezeigt werden kann, wie eine Lehrerin (im Zusammenwirken mit den SchülerInnen) das ‚Spiel Unterricht‘ vollzieht. Das Forschungsinteresse fokussiert dabei auf das praktische Wissen bzw. Können der Lehrerin im Sinne eines adaptiven Eingehens auf die pädagogischen Herausforderungen einer Situation. Im Vordergrund dabei steht die Frage, welche Aufgaben von der Lehrerin angesichts ihrer Kontingenzerfahrung bewältigt werden müssen und wie sie dies tut. Die Praxisanalyse richtet sich dabei gemäß ihrer theoretischen Verortung „weniger auf die hermeneutische Entschlüsselung von Sinn und Bedeutung, als auf die beobachtende Entdeckung der Praxis“ (Breidenstein 2008, S. 110). Die Feinanalyse zeigt die im Unterricht ablaufenden „situierten Interaktionen und Interdependenzgeflechte“ (Schmidt 2012, S. 44) und macht die Ökonomie des LehrerInnenhandelns deutlich, die der Lehrkraft hilft, sowohl den erzieherischen als auch den didaktisch-methodischen ‚roten Faden‘ der Unterrichtsstunde in der Hand zu behalten.

### Die Auseinandersetzung angehender Lehrkräfte mit pädagogischen Situationen: Eine qualitative Analyse von Argumentationsstrukturen und genutzten Wissensquellen.

**Kati Trempler, Ulrike Hartmann**

Bergische Universität Wuppertal, Deutschland; [trempler@uni-wuppertal.de](mailto:trempler@uni-wuppertal.de)

Im Fokus des Beitrages steht die Auseinandersetzung mit pädagogischen Situationen durch angehende Lehrkräfte anhand von wissenschaftlichem Wissen. Dies stellt eine Facette pädagogischer Kompetenz dar, die normativ gefordert und deren Bedeutsamkeit für den schulischen Alltag in der wissenschaftlichen Diskussion häufig akzentuiert wird (Bauer et al., 2017; Hartmann et al., 2016). Gegenstand des vorliegenden Beitrages sind insbesondere kognitive Faktoren, die für die Auseinandersetzung mit wissenschaftlichem Wissen von Bedeutung sind.

Die Nutzung von wissenschaftlichem Wissen setzt Fähigkeiten wissenschaftlichen Denkens voraus (Bauer et al., 2017), das verschiedene Teilprozesse umfasst – von der Identifikation von Problemen über die Generierung von Fragestellungen und die Bewertung von Theorien und Befunden. Die aus diesem Prozess gezogenen Schlussfolgerungen können unter anderem durch Argumentationstexte repräsentiert werden (Fischer et al., 2014). In Anlehnung daran gehen wir davon aus, dass (wissenschaftliches) Denken sich in der Struktur argumentativer Texte wieder spiegeln kann (Iordanou et al., 2016).

Eine vollständige Argumentationsstruktur besteht allgemein aus Behauptung, Argument(en), Gegenargument(en) und Schlussfolgerung (Duncan & Chinn, 2016). Argumentationen bestehend aus Argumenten und Gegenargumenten werden als two-sided argumentations, Argumentationen mit einem Argument als one-sided argumentations und das fehlen jeglicher Argumente als non-argumentation eingestuft (Iordanou, 2016). Bisherigen Erkenntnissen folgend, werden two-sided argumentations mit einer evaluativen und kritischen Denkweise in Verbindung gebracht (Kuhn, 2001; Wolfe, 2012). Weil die Struktur allerdings nicht immer ein Indikator für die Argumentationsqualität ist, schlagen Duncan und Chinn (2016) vor, die Qualität einer wissenschaftlichen Auseinandersetzung zusätzlich anhand ihrer Wissensquellen zu beurteilen.

Da von Lehrkräften erwartet wird, für ihr professionelles Handeln auf wissenschaftliches Wissen zurückzugreifen, und die Fähigkeit des wissenschaftlichen Denkens auch an Lernende zu vermitteln (Muis, Bendixen, & Hearle, 2006), möchten wir untersuchen, inwieweit angehende Lehrkräfte bei der Auseinandersetzung mit pädagogischen Situationen in Form von Lerntagebucheinträgen evidenzbasiert und wissenschaftlich denkend vorgehen. Fragestellungen: (1) Welche Argumentationsstruktur weisen Lerntagebücher von Lehramtsstudierenden auf? (2) Welche Wissensquellen werden für die Erklärung pädagogischer Situationen verwendet?

Wir analysieren Lerntagebucheinträge von 50 Masterstudierenden im Praxissemesters (83% weiblich, Alter M= 26.23). Für die Erstellung der Einträge wurden die Studierenden gebeten, eine schulische Situation auszuwählen, diese zu beschreiben und anhand bildungswissenschaftlicher Literatur zu erklären, um Schlussfolgerungen für das eigene Handeln abzuleiten. Insgesamt

verfassten die Studierenden vier Einträge im Verlauf des Semesters. Für die vorliegenden Analysen nutzen wir den vierten Eintrag der Studierenden.

Die Lerntagebücher werden mittels Qualitativer Inhaltsanalyse untersucht. Die Kodierung erfolgt durch zwei unabhängige Beurteiler anhand folgender Aspekte:

Argumentationsstruktur. In Anlehnung an Arbeiten von Iordanou et al. (2016) und Wolfe (2012) analysieren wir die Erklärungstexte nach Vollständigkeit (Behauptung, Argument, Gegenargumenten, Schlussfolgerung) und betrachten, ob es sich dabei um two-sided-, one-sided- oder non-argumentations handelt.

Wissensquellen. Angelehnt an Arbeiten von Moshman und Tarricone (2016) unterscheiden wir nicht-epistemische (z.B. Erfahrung, Wahrnehmung) und epistemische (z.B. bildungswissenschaftliche Theorien und Befunde) Wissensquellen. Letztere dienen der Wissensgewinnung und dem Verstehen und bilden die Grundlage für wissenschaftliches Denken (Chinn & Rinehart, 2016).

Erste Analysen zeigen, dass drei von vier Studierenden bei der Auseinandersetzung mit schulischen Situationen Argumente und Gegenargumente anführen und Schlussfolgerungen für ihr Handeln ziehen (two-sided argumentations).

Drei von vier Studierenden nutzen nicht-epistemische Grundlagen für Ihre Argumentation (u.a. Erfahrung, Quellen ohne inhaltlichen Bezug), während ein Viertel bildungswissenschaftliche Theorien – jedoch keine empirischen Befunde – für die Erklärung nutzt. Schlussfolgerungen für die pädagogische Praxis scheinen vielmehr auf nicht epistemischen Wissensquellen zu beruhen. Auffallend ist, dass kaum Verbindungen zwischen den Argumenten und zwischen Argumenten und Schlussfolgerungen erkennbar sind. Diese Ergebnisse weisen darauf hin, dass Schlussfolgerungen für die pädagogische Praxis und wissenschaftliches Denken kaum miteinander in Zusammenhang stehen. Die Ergebnisse unterstreichen die Forderung einer systematischeren Untersuchung des wissenschaftlichen Denkens und Argumentation im Bildungsbereich (Fischer et al., 2014). Die geringe Nutzung von theoretischem und empirischem Wissen durch Studierende bestätigt bereits bestehende Forschungsergebnisse (Hetmanek et al., 2014).

## **Zur Wirksamkeit der Arbeit an video- und textbasierten Unterrichtsfällen: eine Interventionsstudie zur Entwicklung feedbackbezogener Performanzen bei angehenden Sportlehrkräften**

**Matthias Baumgartner**

Eidg. Hochschule für Sport, Schweiz; [matthias\\_baumgartner79@gmx.ch](mailto:matthias_baumgartner79@gmx.ch)

Theoretischer Hintergrund

Bezüglich der Diagnostik und der Förderung der Entwicklung des professionellen Handelns von (angehenden) Lehrkräften wird in jüngster Zeit innerhalb der empirischen Bildungsforschung eine performative Wende gefordert (z. B. Kuhn, Zlatkin-Troitschanskaia, Pant & Hannover, 2016). Es wird moniert, dass in der Kompetenzdiagnostik der Fokus nicht einzig auf die einzelnen Aspekte professioneller Kompetenz (z. B. professionelles Wissen) gelegt werden soll (z. B. Kuhn et al., 2016), sondern ein Bezug zum konkreten Handeln in der Komplexität und Ungewissheit von Unterricht hergestellt wird. Denn ‚Kompetenz‘ stellt mehr als Summe der Einzelteile dar, sondern weist einen unmittelbaren Bezug zum Handeln in einer realen Welt auf (Shavelson, 2013). Demgemäss wird aus ökologischen Validitätszwecken gefordert, die Kompetenzdiagnostik verhaltensnaher zu realisieren und die Wirksamkeit der ausbildungsbezogenen Interventionen stärker anhand des Performanzzuwachses der angehenden Lehrkräfte zu messen (vgl. Baumgartner, 2017a, 2017b; Shavelson, 2013). Bislang ist jedoch kaum erforscht, welche Einflüsse verschiedene hochschuldidaktische Inszenierungsformen auf den kompetenzbereichsbezogenen Performanzzuwachs (im Sinne anforderungsbezogene Könnensleistungen) bei angehenden Lehrkräften haben. Folgende Interventionsstudie soll einen Beitrag zur Bearbeitung dieser Forschungslücke leisten und zielt darauf ab, Wissen über die Effektivität von zwei verschiedenen hochschulmethodischen Inszenierungsformen (Arbeit an videobasierten Unterrichtsfällen; Arbeit an textbasierten Unterrichtsfällen) hinsichtlich des feedbackbezogenen Performanzfortschritts bei angehenden Sportlehrkräften zu generieren.

Fragestellung

Im vorliegenden Beitrag werden die Fragestellungen bearbeitet, ob angehende Sportlehrkräfte durch a) die Arbeit an video- sowie b) die Arbeit an textbasierten Unterrichtsfällen innerhalb eines gleichen hochschuldidaktischen Lehr-Lernarrangements ihre feedbackbezogenen Performanzen verbessern können. Vor dem Hintergrund der empirischen Befunde der Arbeit an video- und textbasierten Unterrichtsfällen (z. B. Gold, Hellermann, Burgula, & Holodynski, 2016) resultieren die folgenden Richtungsvermutungen:

H1: Die Qualität der feedbackbezogenen Performanzen von angehenden Sportlehrenden, welche das videovignettenbasierte Lehr-Lernarrangement absolvieren, ist nach der Durchführung des Interventionsprogramms höher ausgeprägt als vor der Absolvierung der Intervention.

H2: Die Qualität der feedbackbezogenen Performanzen von angehenden Sportlehrenden, welche das textvignettenbasierte Lehr-Lernarrangement absolvieren, ist nach der Durchführung des Interventionsprogramms höher ausgeprägt als vor der Absolvierung der Intervention.

H3: Die Arbeit an videobasierten Unterrichtsfällen erweist sich bez. des feedbackbezogenen Performanzfortschritts im Vergleich zur Arbeit an textbasierten Unterrichtsfällen als effektiver.

Methode

Die Untersuchung basiert auf einer quasi-experimentellen Feldstudie mit den Hauptfaktoren Messzeitpunkt (t0, t1) und Untersuchungsgruppe (Videogruppe: N = 23; Textgruppe: N = 15; Kontrollgruppe: N = 22; Pre-Post-Control-Design mit 2x3-Faktorenstufen). Die Stichprobe bestand aus drei Lehrgängen, welche an einer Hochschule dieselbe Ausbildung absolvierten. Die Kontrollgruppe besuchte ein allgemeines Schulpraktikum (Standardintervention). Bei den beiden Experimentalgruppen fand die Intervention innerhalb von fünf Wochen statt (fünf Sitzungen à zwei Lektionen; Schulpraktikum à 16 Lektionen). Die Qualität der feedbackbezogenen Performanzen (AV) wurde vor und nach den jeweiligen Interventionen mittels Videovignetten fixiert und durch ein Fremdeinschätzungsverfahren beurteilt. Dabei wurde das validierte Diagnoseinstrument Widorski et al. (2012) verwendet, welches für die Messung der Güte des Feedbacks im Sportunterricht adaptiert wurde (vgl. Baumgartner, 2013; Baumgartner, 2017b). Die unabhängige Variable (UV) stellen die Interventionen dar. Die Qualität der Daten wurde u.a. mit der Berechnung des Intraklassenkorrelationskoeffizienten ICC quantifiziert. Zur inferenzstatistischen Datenauswertung wurden zweifaktorielle Varianzanalysen mit Messwiederholung (gemischtes Design) gerechnet und die Effektstärke mittels Populationseffektschätzer  $\omega^2$  (Omega Quadrat) eruiert.

## Ergebnisse

Aus den Ergebnissen zeigt sich u.a., dass sich die Qualität der feedbackbezogenen Performanzen durch die Arbeit an feedbackbezogenen videobasierten Unterrichtsfällen (Videogruppe) mit einem starken Effekt und die Arbeit an textbasierten Unterrichtsfällen (Textgruppe) einen moderaten positiven Effekt erhöht hatte. Die Arbeit an feedbackbezogenen eigenen videobasierten Unterrichtsfällen hat sich im Vergleich zur textbasierten Vorgehensweise in hoher Deutlichkeit als wirksamer erwiesen. Das Schulpraktikum alleine hatte keinen Effekt. Die Resultate weisen u.a. auf die hohe Bedeutsamkeit einer gezielten theorie- und praxisverknüpfenden Ausbildungskultur hin.



## E06

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S212

### **Der Einfluss von dissozialem Schülerverhalten auf die Urteilsgenauigkeit von Lehrkräften: Eine experimentelle Studie**

**Sonja Hense, Friederike Zimmermann**

Institut für Pädagogisch-Psychologische Lehr- und Lernforschung an der CAU Kiel, Deutschland; [shense@jpl.uni-kiel.de](mailto:shense@jpl.uni-kiel.de)

#### Theoretischer Hintergrund

Zu den Hauptaufgaben von Lehrkräften zählt die Leistungsbeurteilung von Schülerinnen und Schülern. Urteilsgenauigkeit beschreibt dabei die Fähigkeit, Personen zutreffend zu bewerten (Schrader, 2010). Die Urteilsgenauigkeit von Lehrkräften ist zwar im Allgemeinen zufriedenstellend, dennoch zeigen empirische Studien, dass systematische Urteilsverzerrungen existieren (z.B. Jussim & Harber, 2005; Südkamp, Kaiser & Möller, 2012). Unterscheidet man zwischen einer direkten (Leistungseinschätzung auf prozentualer Skala) und indirekten (globalere Skala, z.B. Note) Einschätzung, wird insbesondere eine geringere Urteilsgenauigkeit bei indirekten, im Schulalltag üblichen Beurteilungen deutlich (Hoge & Coladarci, 1989). Quellen für Urteilsverzerrungen bei Lehrkrafturteilen liegen z.B. im Selbstkonzept und der Intelligenz der Schülerinnen und Schüler (Schrader & Helmke, 1990). Ein weiteres Schülermerkmal, das sich in längsschnittlichen Feldstudien negativ auf Schulnoten niederschlug, wobei für entsprechende Testleistungen kontrolliert wurde, ist dissoziales Schülerverhalten (z.B. Zimmermann, Schütte, Taskinen und Köller, 2013). Dissoziales Schülerverhalten kann daher als ein weiterer die Urteilsgenauigkeit beeinflussender Faktor vermutet werden. Des Weiteren könnte die Klassenkomposition eine Rolle spielen, da Lehrkräfte auf Grundlage eines klasseninternen Bezugssystems urteilen (z.B. Schrader & Helmke, 1987), bei dem Leistungen und gegebenenfalls auch das Verhalten im Kontext der Mitschülerinnen und Mitschüler eingeschätzt werden.

#### Fragestellung

In dieser Studie untersuchen wir, ob dissoziales Verhalten einen unterschätzenden Einfluss auf die Urteilsgenauigkeit hat, und ob es hinsichtlich des Einbezugs dissozialen Verhaltens eine Diskrepanz zwischen einer Fachnote und einer prozentualen Leistungseinschätzung gibt. Des Weiteren untersuchen wir, ob dissoziales Verhalten im Sinne eines Salienzeffekts insbesondere in die Leistungsbeurteilung einbezogen wird, je geringer dieses Verhalten in der Referenzgruppe der Klasse ausgeprägt ist.

#### Methoden

In einem experimentellen Design unter Nutzung eines computersimulierten Klassenraums untersuchten wir, inwieweit dissoziales Schülerverhalten in Fachnoten und prozentuale Leistungseinschätzungen einfließt und ob der Klassenkomposition im dissozialen Verhalten eine moderierende Rolle zukommt.  $N = 88$  Lehramtsstudierende (66% weiblich; durchschnittlich 23 Jahre alt,  $SD = 3,7$ ) interagierten als Lehrperson mit einer simulierten Klasse. Die Lehrpersonen stellten vorgegebene mathematische Fragen und riefen bei einer Meldung eine Schülerin oder einen Schüler auf. Die Antwort war je nach voreingestelltem Leistungsparameter richtig oder falsch.

Im Anschluss folgten die Leistungseinschätzungen auf einer prozentualen Skala sowie mittels Notenvergabe. Je nach Klassenkomposition bekamen sechs (versus sechs; ausgeglichene Bedingung), neun (versus drei; stark dissoziale Bedingung) oder drei (versus neun; gering dissoziale Bedingung) Schülerinnen und Schüler eine Charakterisierung als dissozial (z.B. schmeißt kleine Papierkügelchen oder Radiergummi durch das Klassenzimmer).

Die statistischen Analysen wurden mit Mplus7 durchgeführt (Muthén & Muthén, 1998–2015). Um für die geschachtelte Datenstruktur zu korrigieren, wurden die Versuchspersonen als Cluster-Variable berücksichtigt. Der Einfluss des dissozialen Verhaltens auf die Urteilsgenauigkeit wurde mittels linearer Regression geprüft. Unterschiede zwischen Effekten hinsichtlich der abhängigen Variablen Notenvergabe und prozentuale Leistungsbeurteilung sowie der Einfluss der Klassenkomposition in einer Mehrgruppenanalyse wurden mit Wald-Tests getestet.

#### Ergebnisse

Unter Kontrolle der tatsächlich gezeigten Leistungen ( $\beta = .46, p < .001$  bzw.  $\beta = .48, p < .001$ ) ergab sich der erwartete unterschätzende Einfluss von dissozialem Verhalten sowohl auf Fachnoten ( $\beta = -.12, p < .001$ ) als auch auf die prozentuale Leistungseinschätzung ( $\beta = -.08, p \leq .01$ ); die Unterschiede zwischen den beiden Beta-Gewichten waren in erwarteter Richtung, verfehlten jedoch die Signifikanz ( $\chi^2(1) = 3.39, p = .07$ ). Es zeigte sich eine Tendenz zu größeren Effekten dissozialen Verhaltens, je weniger Kinder im Klassenraum dissoziales Verhalten zeigten (Notenvergabe bzw. prozentuale Einschätzung: stark dissozial  $\beta$  (SE) =  $-.05$  (.06) bzw.  $-.07$  (.05), ausgeglichen:  $\beta$  (SE) =  $-.14^*$  (.06) bzw.  $-.09^*$  (.05), gering dissozial:  $\beta$  (SE) =  $-.17^*$  (.05) bzw.  $-.11^*$  (.04); \*  $p \leq .05$ ); die erwarteten Unterschiede zwischen den Bedingungen wurden jedoch nicht signifikant.

Die Befunde sind für die Bewertungspraxis in der Schule relevant. Weitere Implikationen für Praxis und Forschung werden diskutiert.

### **Sexualisierte Gewalt Peer to Peer – eine zentrale Perspektive für professionelles Präventionshandeln an Schulen**

**Sabine Maschke<sup>1</sup>, Ludwig Stecher<sup>2</sup>**

<sup>1</sup>Universität Marburg, Deutschland; <sup>2</sup>Universität Gießen, Deutschland; [sabine.maschke@staff.uni-marburg.de](mailto:sabine.maschke@staff.uni-marburg.de), [Ludwig.Stecher@erziehung.uni-giessen.de](mailto:Ludwig.Stecher@erziehung.uni-giessen.de)

Theoretischer Hintergrund. In verschiedenen Studien zur sexualisierten Gewalt wird vor allem die Gewalt von Erwachsenen gegenüber Kindern und Jugendlichen thematisiert. In unserem Beitrag richten wir das Augenmerk auf sexualisierte Gewalt Peer to Peer und hier vor allem auf die Institution Schule als Ort sexualisierter Gewalt. Den theoretischen Hintergrund bildet dabei das Konzept der Vulnerabilität (vgl. u.a. Andresen, 2014; Berzin, 2010). Wir gehen davon aus, dass in der Jugendphase zum einen eine Vielzahl unterschiedlichster und anspruchsvoller Entwicklungsaufgaben in einer relativ kurzen biographischen Zeitspanne zu lösen ist als auch, dass die Jugendlichen in dieser Zeitspanne im Besonderen offen sind für Sozialisationsangebote seitens der Peers. Die Jugendforschung geht davon aus, dass die Gleichaltrigen wichtige Unterstützer für eine positive Entwicklung im Jugendalter sind (Youniss & Smollar, 1985), dies gilt vor allem mit Blick auf die sexuelle Identitätsrollenentwicklung. Virulent wird dieser Zusammenhang aus Entwicklungsaufgaben, Vulnerabilität und Peersozialisation vor allem in Institutionen, in denen Jugendliche zusammenkommen – wie der Schule.

Basierend auf einer aktuellen repräsentativen standardisierten Befragung von 2.700 Schülerinnen und Schülern der 9./10. Jahrgangsstufe zu ihren sexualisierten Gewalterfahrungen und (u.a.) ihrer Wahrnehmung von Schule zeigt der Beitrag, dass vor allem verbale und/oder schriftliche Formen sexualisierter Gewalt unter Gleichaltrigen in der Schule weit verbreitet sind. Wir können hier von einer Art sexualisierter Beschimpfungskultur sprechen. Aber auch Formen körperlicher sexualisierter Gewalt werden in der Schule erlebt. Die Täter/innen sind dabei in der überwiegenden Mehrheit Gleichaltrige, also Mitschüler/innen. Neben der Entwicklungsunterstützung durch Gleichaltrige zeigt sich damit also, dass die Peers mit Blick auf sexualisierte Gewalterfahrungen auch ein erhebliches Sozialisationsrisiko darstellen können.

Fragestellung. Für die Bildungsforschung ist dabei im Besonderen relevant, wie durch sexualisierte Gewalterfahrungen durch Gleichaltrige die Wahrnehmungen schulischer Aspekte der Schülerinnen und Schüler beeinträchtigt werden. Dies bezieht sich einerseits auf die allgemeine Lern- und Schulfreude, das Lern-, oder Klassenklima, die Beziehung zu den Lehrkräften aber auch das Sicherheitsempfinden in der Schule. Mithin also verschiedenen Aspekten, die sich auch auf die Leistungsentwicklung der Schüler/innen auswirken.

Methode/Ergebnisse. Dem Zusammenhang zwischen sexualisierten Gewalterfahrungen und den genannten (und weiteren) schulischen Lernumweltfaktoren geht der Beitrag u.a. mithilfe mehrbenenanalytischer Modelle (Schulebene, Individualebene) nach. Es zeigt sich, dass auch das Erleben vermeintlich leichterer Formen sexualisierter Gewalt wie sexuelle Beleidigungen wesentlich zu einer negativen Wahrnehmung von Schule beiträgt. Das gilt im Übrigen auch für Jugendliche, die an ihrer Schule zwar nicht unmittelbar von sexualisierter Gewalt betroffen sind, aber diese dort beobachtet haben.

Die Befunde zeigen dabei darüber hinaus, dass sich die Einzelschulen mit Blick auf ein sexualisierte Gewalt begünstigendes Schulklima voneinander unterscheiden. Der Beitrag wird auch der Frage nachgehen wie diese Unterschiede (zumindest zum Teil) zu erklären sind.

Konsequenzen für professionelles Handeln. Aus den Befunden werden im letzten Teil des Beitrags Konsequenzen für das professionelle Handeln von Lehrkräften mit Blick auf die Präventionsarbeit in der Schule abgeleitet. Während Präventionsarbeit häufig sich auf sexualisierte Gewalt von Erwachsenen gegen Jugendliche richtet, zeigt Speak!, dass der Blick verstärkt auf die Gleichaltrigen zu richten ist. Ein entsprechendes jugendorientiertes Konzept wird am Ende kurz vorgestellt.

## **Wer mag wen? Profitieren Schülerinnen und Schüler, wenn die Lehrperson deren Peerbeziehungen kennt?**

**Marvin Harks, Bettina Hannover**

Freie Universität Berlin, Deutschland; [m.harks@fu-berlin.de](mailto:m.harks@fu-berlin.de)

Lernerfolg von Schülerinnen und Schülern (SuS) hängt maßgeblich von deren Anstrengungsbereitschaft ab, die wiederum von Erfolgserwartungen, Interesse oder Werten beeinflusst wird (Fredricks et al., 2004). In diesem Beitrag möchten wir diese individuellen Einflussfaktoren auf schulische Anstrengungsbereitschaft um Aspekte sozialer Eingebundenheit erweitern und nehmen an, dass höhere soziale Akzeptanz durch Peers und genauere Kenntnisse der Lehrperson über die Interaktionen im Peernetzwerk die schulische Anstrengungsbereitschaft stärken. Wir thematisieren somit, inwiefern SuS von einer kaum erforschten Facette von diagnostischer Kompetenz von Lehrpersonen profitieren, nämlich der Fähigkeit die Peerbeziehungen der Klasse angemessen wahrzunehmen.

Lehrpersonen, die die Peerbeziehungen ihrer Klasse präzise wahrnehmen, können ein lerngünstigeres Klima schaffen. So fanden Hamm et al. (2011) und Gest et al. (2014) in ihrer Forschung über teacher-attunement, dass SuS sich stärker durch Peers unterstützt und mehr zur Schule zugehörig fühlten sowie motivierter und anstrengungsbereiter waren, wenn ihre Lehrkraft genauer über die Cliquenzusammensetzung der Klasse Bescheid wusste. Zudem kann die Lehrperson präzises Wissen über Peerbeziehungen bei der Konzeption kooperativer Lernformen, für Bullying-Interventionen oder für die gezielte Verbesserung von Peerbeziehungen nutzen, beispielsweise mittels Sitzplatzinterventionen (Van den Berg & Cillessen, 2015).

Allerdings zeigte die teacher-attunement-Forschung, dass Lehrkräfte kaum in der Lage waren, die SuS ihren Cliquen zuzuordnen. Validiert wurde deren Urteilsgenauigkeit an gemittelten Angaben (über alle SuS) zu Cliquen dieser Klasse. In unserer Forschung argumentieren wir hingegen, dass nicht Cliquenzugehörigkeit, sondern die Wahrnehmung individueller Peerbeziehungen, viel nützlicher dafür ist, um etwa kooperative Lernumgebungen zu konzipieren oder sozialer Exklusion vorzubeugen. Im Unterschied zur teacher-attunement-Forschung haben wir daher Lehrpersonen gebeten, soziometrisch erhobene Peerbeziehungen ihrer SuS zu replizieren und fanden, dass Lehrpersonen etwa 30% dieser Peerbeziehungen korrekt wiedergeben konnten (Autor1 und Autor2, 2017).

Darauf aufbauend ergibt sich die Frage, ob sich die von uns gemessene Kompetenz zur Wahrnehmung von Peerbeziehungen ebenso wie teacher-attunement positiv auf die schulische Anstrengungsbereitschaft der SuS auswirkt. Wir nehmen an, dass schulische Anstrengungsbereitschaft höher ausfällt, wenn SuS sozial akzeptierter sind, sich sozial eingebunden fühlen und die Lehrperson präziser die individuellen Peerbeziehungen kennt.

Hierfür haben wir 821 SuS aus 37 Klassen (Klassenstufe 4-10, 17 Grundschulklassen) sowie deren Klassenlehrkraft befragt. Wir haben von allen SuS soziometrisch die drei Peers ermittelt, die diese am liebsten mögen und haben die Lehrperson unabhängig davon gebeten, diese Nominierungen für alle der Klasse zu replizieren. Hierbei haben wir die korrekte Wahrnehmung der Lehrperson individuell pro Schüler/in als Jaccardkoeffizient operationalisiert, der den Anteil der übereinstimmenden Nominierungen im Verhältnis zu allen angegebenen Nominierungen wiedergibt (Neal et al., 2011). Die soziale Akzeptanz im Peernetzwerk haben wir anhand der pro Schüler/in eingehenden soziometrischen Nominierungen im Verhältnis zur Klassengröße ermittelt. Die empfundene soziale Eingebundenheit (Rauer & Schuck, 2003;  $\alpha=.85$ ) und die schulische Anstrengungsbereitschaft (ebd.;  $\alpha=.82$ ) haben wir als Selbstbericht erfasst. Da schulische Anstrengungsbereitschaft vom zurückliegenden Lernerfolg abhängt, haben wir für den Einfluss der Durchschnittsnote in den Hauptfächern (Deutsch, Englisch, Mathematik) kontrolliert.

In Mehrebenen-Regressionsmodellen zeigte sich, dass die schulische Anstrengungsbereitschaft auf Individualebene höher ausfiel, je besser die Durchschnittsnote war ( $\beta=.32$ ,  $p<.001$ ), je stärker der/die Schüler/in sich sozial eingebunden fühlte ( $\beta=.22$ ,  $p<.001$ ) und je genauer die Lehrperson die individuellen Sympathiebeziehungen replizieren konnte ( $\beta=.13$ ,  $p<.01$ ). Die soziale Akzeptanz zeigte unerwarteter Weise keine Vorhersagekraft ( $\beta=-.08$ ,  $p<.1$ ). Auf Klassenebene lag die schulische Anstrengungsbereitschaft (ICC=.10) höher, wenn die SuS eine höhere soziale Eingebundenheit empfanden ( $\beta=.50$ ,  $p<.01$ ). Note, soziale Akzeptanz und Lehrpersonenwahrnehmung von Sympathiebeziehungen waren nicht mehr bedeutsam, nachdem empfundene soziale Eingebundenheit auf Klassenebene einbezogen wurde.

SuS sind somit anstrengungsbereiter, wenn sie sich sozial eingebunden fühlen. Individuell profitieren sie aber zusätzlich, wenn die Lehrperson ihre Peerbeziehungen präziser kennt. Lehrpersonen haben dabei scheinbar keine generell hohe/niedrige Wahrnehmung auf Klassenebene, sondern nehmen die individuellen Peerbeziehungen der Klasse unterschiedlich gut wahr. Implikationen für Unterricht und Lehrkräftebildung werden diskutiert.

## **Giftedness perceived as a resource or vulnerability – Which variables do predict who is considered to be gifted?**

**Jessika Golle<sup>1</sup>, Ulrich Trautwein<sup>1</sup>, Trudie Schils<sup>2</sup>, Lex Borghans<sup>2</sup>**

<sup>1</sup>Universität Tübingen, Deutschland; <sup>2</sup>Maastricht University, Niederlande; [jessika.golle@uni-tuebingen.de](mailto:jessika.golle@uni-tuebingen.de)

Theoretical background:

Teachers' conceptions about giftedness play an important role in the identification and promotion of talented students. This, in turn, provides a basis for transforming the high potential of these students into achievement. In the literature, there is an ongoing debate about the question whether teachers' beliefs about giftedness are in line with the harmony or disharmony hypothesis (cf. Baudson & Preckel, 2016). The harmony hypothesis indicates that gifted students are superior in all domains, i.e. they show high general cognitive abilities, there are highly motivated, creative, and socially able (= giftedness as a resource, see e.g., Endepohls-Ulpe & Ruf, 2005). According to the disharmony hypothesis, gifted students are also considered to show high general cognitive abilities but, in contrast to the harmony hypothesis, they are considered to be emotionally and socially fragile (= giftedness as vulnerability, see e.g., Neihart, 1999). Previous studies revealed a mixed pattern of results, but the majority of research findings support the assumption that giftedness is a resource (e.g., Dai, Moon, & Feldhusen, 1998). In contrast to the empirical findings, teachers' beliefs seem to be more strongly affected by the disharmony hypothesis (cf. Baudson & Preckel, 2016). Thus, there is a meaningful gap between teachers' beliefs and empirical evidence. In this study, we tried to get a deeper insight into this gap. We investigated a representative sample of 6th Grade students and their teachers in the Netherlands. Teachers were asked whether they think that the children in their class were gifted. We predicted this decision by cognitive as well as noncognitive measures. To go one step further than previous studies, we were especially interested in the variation between teachers.

Research questions:

1. Which characteristics explain the teacher judgment about a child's giftedness? According to previous work, we expected to find support for the disharmony hypothesis.

2. How much does the influence of these characteristics vary between teachers? Both hypotheses link giftedness to high cognitive abilities, thus we assumed that for ability measures the variance components should be smaller than the variance components for noncognitive attributes (random slopes).

Methods:

We used data from a longitudinal cohort study that has been conducted in the Netherlands since 2009. Overall, data of 26,720 students and 1,304 teachers were analyzed. To measure general cognitive abilities, we used the data of a non-verbal intelligence test (van Batenburg, 2015) as well as a school achievement test (van Boxtel, Engelen, & de Wijs, 2011). To assess the personality of the students, questions belonging to seven personality dimensions were asked. In addition to the Big five traits, persistence (e.g., "If I want something I continue until I succeed") and performance (e.g., "It is important for me to try my best") were measured. To measure self-concept, questions regarding the academic (e.g., "good at writing an essay") as well as social self-concept (e.g., "good at comforting someone who is sad") were considered. We did confirmatory factor analysis for all noncognitive measures to assess the dimensionality and reliability of the latent variables. Besides these variables, we also included teacher assessments about the student behavior.

Results and Discussion:

We applied logistic multilevel regression analysis. The dependent variable was "being considered to be gifted by the teacher". All remaining variables were used as predictors. Preliminary findings did not support any of the two hypotheses (harmony vs. disharmony). However, as expected, a larger variation in slopes across teachers was found for noncognitive factors compared to cognitive factors. Further analyses should be conducted to figure out whether there are teachers whose beliefs about giftedness are in line with either the harmony or disharmony hypothesis, and which teacher characteristics might explain this variation.

## E07

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H120

### Schulisches Medienhandeln und soziale Ungleichheiten bei digitalen Kompetenzen

Maximilian Weber

Goethe Universität; [m.weber@soz.uni-frankfurt.de](mailto:m.weber@soz.uni-frankfurt.de)

#### Fragestellung

Kinder und Jugendliche in Industrienationen wachsen in einer digitalisierten Welt auf. Für die Teilhabe an der Gesellschaft ist die Nutzung von Informations- und Kommunikationstechnologien (IuK) unabdingbar. Computer- und informationsbezogene Kompetenzen stellen eine Schlüsselkompetenz dar, die vonnöten ist, um in diversen Lebens- und Arbeitssituationen zu bestehen (Eickelmann et al. 2016; DIVSI 2016). Gesellschaftliche Chancengerechtigkeit betrifft damit auch den Bereich des Digitalen (DIVSI 2016, p. 76). Wer Informations- und Kommunikationstechnologien nicht zielgerichtet einsetzen kann oder möchte, ist von bestimmten Prozessen ausgeschlossen (ebd.). Diese Kenntnisse und Fähigkeiten werden in der Literatur häufig als digitale Kompetenzen oder als computer- und informationsbezogene Kompetenzen bezeichnet. Letzterer Begriff wird von Fraillon et al. folgendermaßen definiert: „Computer and information literacy refers to an individual's ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society“ (2013, p. 17).

Im Gegensatz zur gängigen Praxis, bei der Betrachtung von sozialen Disparitäten bei computer- und informationsbezogenen Kompetenzen auf Selbstangaben zurückzugreifen, soll mit Daten eines Kompetenz- und Fähigkeitstests gearbeitet werden. Die bisherigen Analysen in dem Bereich der digitalen Kompetenzen konnten einen Zusammenhang zur sozialen Lage des Elternhauses aufzeigen, jedoch wird bei solchen Analysen kaum auf die Mediennutzung in der Schule und in der Schule erlernten Tätigkeiten zur Informationsgewinnung eingegangen. Es soll daher analysiert werden, wie stark unter der Kontrolle von Drittvariablen der Einfluss der sozialen Herkunft auf die Ausbildung von digitalen Kompetenz wirkt. Konkret wird geprüft:

Findet sich ein Zusammenhang von Merkmalen der sozialen Herkunft auf die computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern auch dann, wenn Merkmale der Mediennutzung in der Schule berücksichtigt werden? Moderiert das Medienhandeln in der Schule den Effekt der sozialen Herkunft auf die computer- und informationsbezogenen Kompetenzen?

#### Theoretischer Hintergrund

Theoretisch lassen sich zwei Wirkzusammenhänge von Medienhandeln in der Schule denken. Erstens könnte der sogenannte Matthäus-Effekt vorliegen, ein Verstärkungsprozess: Jugendliche mit bereits besseren Kenntnissen profitieren von dem Medienhandeln in der Schule in einem höheren Ausmaß als ihre Mitschülerinnen und Mitschüler, die sich auf einem schlechteren Kenntnisstand befinden. Zweitens könnte auch ein Aufholprozess vorliegen: Jugendliche mit weniger guten digitalen Kompetenzen können aufgrund des Erlernens in der Schule ein ähnliches Niveau erreichen wie die Schülerinnen und Schüler, die bereits außerschulisch die Kompetenzen erworben haben. Ein solcher Prozess wird in der Literatur als „ceiling“-Effekt bezeichnet.

#### Methode

Zur Beantwortung der Forschungsfrage wird mittels einer Sekundärdatenanalyse der Einfluss der sozialen Herkunft auf die Ausbildung von computer- und informationsbezogenen Kompetenzen ermittelt. Als Datengrundlage dient der repräsentative Datensatz, der im Zuge der „International Computer and Information Literacy Study“ (ICILS) 2013 entstanden ist. In den Analysen werden die in Deutschland lebenden Teilnehmerinnen und Teilnehmer der Studie berücksichtigt.

#### Ergebnisse

In einem ersten Schritt wird der Einfluss der sozialen Herkunft (höchster ISEI-Wert der Eltern und verfügbares kulturelles Kapital, gemessen über die Anzahl der Bücher im Elternhaus) geschätzt (N=1947). Dieses Modell zeigt einen klaren Zusammenhang zwischen den Ergebnissen des Leistungstests in Bezug auf die computer- und informationsbezogenen Kompetenzen auf. Schülerinnen und Schüler aus sozial besser gestellten Elternhäusern erzielen im Durchschnitt einen höheren Wert auf der Testskala und weisen somit bessere digitale Kompetenzen auf. In einem weiteren Schritt werden die zwei Indizes zur Mediennutzung im Unterricht und der in der Schule gelernten Tätigkeiten zur Informationsgewinnung im Internet in das Modell gegeben. Unter der Kontrolle dieser Variablen bleibt der Effekt der sozialen Herkunft unverändert bestehen. Die beiden Variablen üben zudem keinen signifikanten Einfluss auf die Ausbildung von digitalen Kompetenzen aus. In den Daten kann somit bei der gleichzeitigen Berücksichtigung der sozialen Herkunft kein Zusammenhang zwischen dem schulischen Medienhandeln und der Ausbildung digitaler Kompetenzen gefunden werden.

Warum schafft es die Schule nicht diese Ungleichheit auszugleichen?

### **Können Schulen den Zusammenhang zwischen der sozialen Herkunft und digitalen Kompetenzen von Jugendlichen verringern? Die Moderator-Wirkung des Schulklimas, der technischen Ausstattung und des durchschnittlichen Leistungsniveaus in der Schule**

**Christian Schnautz, Birgit Becker, Maximilian Weber**

Goethe Universität Frankfurt, Deutschland; [Bi.Becker@em.uni-frankfurt.de](mailto:Bi.Becker@em.uni-frankfurt.de), [M.Weber@soz.uni-frankfurt.de](mailto:M.Weber@soz.uni-frankfurt.de)

#### Einleitung und Fragestellung:

Der Zusammenhang zwischen dem sozioökonomischen Status der Eltern und den Lesekompetenzen von Schülerinnen und Schülern ist in vielen Studien demonstriert worden und schon lange bekannt (z.B. Barone 2006). Ein ähnlicher Zusammenhang zeigt sich auch im Bereich der digitalen Kompetenzen (z.B. Hargittai 2010). In diesem Beitrag wird als Hauptfragestellung untersucht, ob ausgewählte Schulmerkmale (Schulklima, technische Ausstattung, Leistungsniveau) diesen Zusammenhang abschwächen und damit digitale Ungleichheiten verringern können (Moderator-Wirkung).

#### Theoretischer Hintergrund:

Der Beitrag schließt an die Forschung zur zweiten digitalen Spaltung der Gesellschaft (second digital divide) an, wobei nicht länger nur Ungleichheiten beim Zugang zu Internet und Hardware, sondern insbesondere Ungleichheiten im Internetnutzungsverhalten und bei digitalen Kompetenzen untersucht werden (z.B. DiMaggio & Hargittai 2001; Zillen & Hargittai 2009). Aus dieser Forschungstradition stammt das Konzept der digitalen Kompetenz (digital literacy), das folgendermaßen

definiert wird: „an individual's ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society“ (Fraillon, Schulz & Ainley 2013: 17).

Wir untersuchen, ob der Zusammenhang zwischen der Bildung der Eltern und den digitalen Kompetenzen von Jugendlichen abgeschwächt wird, wenn sie Schulen mit förderlichen Merkmalen besuchen (für ein ähnliches Vorgehen im Bereich traditioneller Schulleistungen vgl. Gustafsson, Nilsen & Hansen, im Druck). Ein solcher kompensatorischer Effekt kann vermutet werden, da Jugendliche aus Familien mit hohem sozioökonomischen Status wahrscheinlich – ganz unabhängig vom Schulkontext – tendenziell hohe digitale Kompetenzen erwerben, während bei Jugendlichen aus Familien mit geringem sozioökonomischen Status eine förderliche Schulumgebung positiv wirken könnte.

Daten und Methode:

Für die Auswertungen wurden die Daten des „Programme for International Student Assessment“ (PISA) aus dem Jahr 2009 verwendet, wobei die Analysen auf diejenigen westliche Industrieländer beschränkt wurden, die das Modul zu Informations- und Kommunikationstechnologie durchgeführt hatten. Die digitale Kompetenz (abhängige Variable) wurde über einen standardisierten Test erhoben (OECD 2011). Bei den Schulmerkmalen wurden das Schulklima und die technische Ausstattung über Items aus dem Schulfragebogen operationalisiert, zur Messung des Leistungsniveaus einer Schule wurde die individuelle Leseleistung aggregiert (Durchschnittswert auf Schulebene).

Bei der Auswertung wurden Mehrebenenmodelle verwendet, die der geschachtelten Datenstruktur Rechnung tragen, wobei Merkmale der Schüler/innen (Ebene 1) und Schulen (Ebene 2) berücksichtigt wurden. Die dritte Ebene (Länder) wurde über Länder-fixed effects berücksichtigt. Von besonderem inhaltlichem Interesse sind die cross-level Interaktionen zwischen der Bildung der Eltern und den untersuchten Schulmerkmalen (Schulklima, technische Ausstattung, Leistungsniveau).

Ergebnisse:

Es zeigt sich, dass Jugendliche, deren Eltern ein hohes Bildungsniveau aufweisen, eine um durchschnittlich 32 Punkte bessere digitale Kompetenz aufweisen als Jugendliche mit niedrig gebildeten Eltern – diese entspricht etwa einem Drittel einer Standardabweichung. Von den untersuchten Schulmerkmalen zeigt sich beim Schulklima und der durchschnittlichen Leseleistung ein signifikanter Zusammenhang mit den digitalen Kompetenzen, während die technische Ausstattung der Schule keinen Effekt zeigt. Es erweist sich lediglich der Interaktionseffekt zwischen der Bildung der Eltern und der durchschnittlichen Leseleistung in der Schule als signifikant – allerdings nicht in der erwarteten Richtung: Der Zusammenhang zwischen der Bildung der Eltern und den digitalen Kompetenzen der Jugendlichen ist in Schulen mit überdurchschnittlichem Leistungsniveau größer als in Schulen mit geringerer durchschnittlicher Leseleistung. Weitere Analysen ergeben, dass der Zusammenhang zwischen der Bildung der Eltern und den digitalen Kompetenzen zu einem Großteil über die traditionelle „offline“ Lesekompetenz mediiert wird.

### **Einstellungen gegenüber Cybermobbing bei Adoleszenten: Ergebnisse einer Konstruktvalidierung basierend auf der IEA-Studie ICILS sowie der Studie ‚Lebenswelten Ostschweizer Jugendlicher‘**

**Michael Beck<sup>1</sup>, Arvid Nagel<sup>1</sup>, Horst Biedermann<sup>1</sup>, Fritz Oser<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule St.Gallen, Schweiz; <sup>2</sup>Universität Fribourg, Schweiz; [michael.beck@phsg.ch](mailto:michael.beck@phsg.ch)

Cybermobbing ist die Aktualisierung einer negativen Kompetenz. Grundsätzlich kann darunter die Fähigkeit zu Tätigkeiten verstanden werden, die gesellschaftlich nicht erwünscht, vor dem Hintergrund ethischer Normen moralisch fragwürdig, vor dem Gesetz verboten und unter sozialen Standards diskriminierend sind. Um es enger zu fassen: Cybermobbing als negative Kompetenz besteht darin, andere (Opfer) zu schädigen oder zumindest in gesellschaftlich unerwünschter Weise zu beeinträchtigen ohne sich in das Leid dieser zu versetzen, wobei man meistens unerkannt bleiben möchte. Wir wissen noch wenig darüber, wie (junge) Menschen die Lust und den Drang entwickeln, diese Kompetenz immer wieder zu verfeinern (vgl. dazu die Bestandsaufnahme von Petermann & von Marées, 2013). Cybermobbing wird als eine absichtliche, wiederholte, aggressive und schädigende Handlung gegenüber einer anderen Person mithilfe von Informations- und Kommunikationstechnologien bestimmt (vgl. u.a. Kowalski, Morgan & Limber, 2012; Langos, 2012; Petermann & von Marées, 2013; Porsch & Pieschl, 2014). Diesem Tun liegt wohl die Einstellung zugrunde, dass man Macht (über eine andere Person) haben kann, dass man brutal schädigen kann und dabei ungeschoren bleibt, dass man andere verunsichern kann – Einstellungen, die auch beim heutigen Phänomen des Cybermobbings gefunden werden. Obwohl die Thematik Cybermobbing sowohl im nationalen als auch im internationalen Kontext zunehmend an öffentlichem und wissenschaftlichem Interesse gewinnt (z.B. Berne et al., 2013; Fenaughty & Harré, 2013; Festl, 2015; Katzer, 2014; Perren et al., 2012; Pieschl & Porsch, 2014; Sticca, Perren, Ruggieri & Alsaker, 2013; Willard, 2007), so fallen die empirischen Evidenzen bis dato dennoch eher dünn aus. Als ein zentrales Desiderat lässt sich diesbezüglich auch festhalten, dass es bis anhin an geeigneten Instrumenten zur empirisch-quantitativen Erfassung von Einstellungen gegenüber Cybermobbing fehlt.

Diese Forschungslücke bearbeitend liegt ein Ziel des vorliegenden Beitrags in der Präsentation eines eigenentwickelten Instrumentes zur Erfassung von Einstellungen gegenüber Cybermobbing. Konkret werden in einem ersten Schritt die Reliabilität und die Faktorenstruktur des entwickelten Instruments überprüft – basierend auf zwei schweizerischen Datengrundlagen. Daran anschließend wird den Fragestellungen nachgegangen, (1) wie verbreitet sich Einstellungen gegenüber Cybermobbing bei Schweizer Schülerinnen und Schülern (grundsätzlich und getrennt nach Geschlechtern - bei den Opfern, Tätern, Bystander sowie Täter-Opfer) zeigen und (2) wie lassen sich die Einstellungen der Schülerinnen und Schüler gegenüber Cybermobbing erklären?

Im Rahmen der ICILS-Studie (International Computer and Information Literacy Study) der IEA konnten basierend auf einem nationalen Modul insgesamt 2915 Schweizer Schüler/innen der achten Jahrgangsstufen aus 94 Schulen befragt werden. Die Einstellung gegenüber Cybermobbing (Skaleneigenkonstruktion;  $\alpha = .84$ ) wurde anhand von fünf Items erhoben. Das strukturüberprüfende Verfahren der konfirmatorischen Faktorenanalyse bestätigt das theoretisch begründete g-Faktormodell (Modell-Fit:  $\chi^2 = 9.02$ ;  $df = 5$ ;  $p = .000$ ; CFI = .981; RMSEA = .062; SRMR = .020). Gestützt werden diese Ergebnisse durch eine weitere aktuelle Datengrundlage von 496 Jugendlichen (LeOJ: „Lebenswelten Ostschweizer Jugendlicher“; Modell-Fit:  $\chi^2 = 7.83$ ;  $df = 4$ ;  $p = .10$ ; CFI = .99; RMSEA = .04; SRMR = .01; hohe Faktorladungen von  $\lambda = .72$  bis  $\lambda = .84$  für die fünf Indikatoren;  $\alpha = .87$ ). Geschlechterunterschiede lassen sich bei der Einstellung gegenüber Cybermobbing identifizieren, indem sich die Mädchen bedeutsam ablehnender zeigen (ICILS:  $t(2896) = -9.55$ ,  $p < .001$ ; MMädchen = 1.67, SD = .69; MJungen = 1.94, SD = .79 |  $d = .36$ ; LeOJ:  $t(472) = 5.02$ ,  $p < .001$ ; MMädchen = 1.26, SD = .50; MJungen = 1.52, SD = .62 |  $d = .46$ ). Auch zwischen Täter/innen und Nicht-Täter/innen von Cybermobbing unterscheiden sich die diesbezüglichen Einstellungen signifikant ( $t(2913) = -5.99$ ,  $p < .001$ ; MTäter = 2.19, SD = .81; Mkein Täter = 1.79, SD = .75 |  $d = .51$ ). Bedeutsame Prädiktoren zur Erklärung der Einstellung gegenüber Cybermobbing werden im Rahmen des Beitrags mittels multipler Regressionsanalysen präsentiert.

Neben der Darlegung der theoretischen Grundlagen und empirischen Erkenntnisse werden die Befunde auch im Lichte des schulischen Auftrags, Lernende bezüglich eines produktiven Umgangs mit Informations- und Kommunikationstechnologien zu unterstützen, reflektiert und diskutiert.

## **Evaluation von Lernleitern als Strukturierungshilfe im Chemieunterricht**

**Marie-Therese Hauerstein, Helena van Vorst**

Universität Duisburg-Essen, Deutschland; [marie-therese.hauerstein@uni-due.de](mailto:marie-therese.hauerstein@uni-due.de)

Nach wie vor muss sich der mathematisch-naturwissenschaftliche Unterricht dem Vorwurf stellen, den kumulativen Wissensaufbau zu wenig zu fördern, denn häufig gelingt es den Lernenden in diesen Fächern nicht, ihr neu erworbenes Wissen an bereits vorhandenem Vorwissen anzuknüpfen und neu zu strukturieren (BLK, 1997; ISB, 2002). Snow (1989) konnte zeigen, dass Lernen insbesondere bei leistungsschwachen Schülerinnen und Schülern besser gelingt, wenn neue Inhalte strukturiert vermittelt werden. Zudem gilt Strukturierung auch allgemein als wesentliches Merkmal guten Unterrichts (Helmke, 2014; Kounin, 1970; Meyer, 2014). Eine Möglichkeit, eine strukturierte Vorgehensweise im Unterricht sicherzustellen, bietet die Methode der Lernleiter. Bei Lernleitern handelt es sich um eine Strukturierungshilfe, die durch einen sequenzierten Aufbau die Lerninhalte sowie den Unterrichtsgang in sinnvolle Abschnitte unterteilt (Girg, Lichtinger & Müller, 2012).

Neben einem sequenzierten Aufbau zeichnet sich die Lernleiter in dieser Studie zusätzlich durch individuelle Übungsphasen aus, wodurch die Strukturierung mit einer Binnendifferenzierung gekoppelt wird. Die Notwendigkeit einer Differenzierung ergibt sich aus der Heterogenität der Schülerschaft, die sich unter anderem durch Unterschiede im Vorwissen, im Lerntempo und in der kognitiven Begabung äußert. Für einen erfolgreichen Lernprozess der einzelnen Lernenden ist es erforderlich, die verschiedenen Lernvoraussetzungen im Unterricht zu berücksichtigen (Bohl, Bönsch, Trautmann & Wischer, 2012). Eine mögliche Differenzierungsform ist die Einteilung der Schülerschaft nach ihrer Leistungsfähigkeit (Meyer, 2014), zu der bisherige Studien für das Fach Chemie positive Ergebnisse zeigen (z. B. Kallweit, 2014; Anus, 2015). Bislang ist jedoch ungeklärt, welche Wechselwirkungen zwischen der Binnendifferenzierung nach Leistungsfähigkeit und der Strukturierung bestehen, weshalb sich die folgenden Forschungsfragen ergeben:

FF1: Inwieweit wirkt die Strukturierung des Unterrichts durch die Lernleiter auf affektive und kognitive Schülervariablen?

FF2: Inwieweit wirkt die Binnendifferenzierung nach Leistungsfähigkeit auf affektive und kognitive Schülervariablen?

FF3: Inwieweit wirkt die Binnendifferenzierung in Kombination mit der Strukturierung durch die Lernleiter auf affektive und kognitive Schülervariablen?

Es wird eine Prä-Post-Kontrollgruppenstudie im 2x2-Design durchgeführt. Dazu werden drei Interventionsgruppen (SB: Lernleiter-Strukturierung + Binnendifferenzierung; S: Lernleiter-Strukturierung + keine Binnendifferenzierung; B: keine Lernleiter-Strukturierung + Binnendifferenzierung) und eine Kontrollgruppe gebildet (KG: keine Lernleiter-Strukturierung + keine Binnendifferenzierung).

Zum Prä-Messzeitpunkt wird ein Schülerfragebogen als Testinstrument verwendet, welcher die kognitive Fähigkeit (KFT), das Fachwissen, das individuelle Interesse, die Motivation sowie das Selbstkonzept der Lernenden erfasst. Während der Intervention begleitet ein Implementationstagebuch der Lehrkräfte die Studie. Zum Post-Messzeitpunkt wird erneut der Schülerfragebogen der Prä-Testung (ohne KFT) eingesetzt.

In einer Pilotstudie wurden bereits die Gruppen SB und B realisiert, um zunächst den Einfluss der Strukturierung zu untersuchen. Beide Lerngruppen wurden zum Thema Atombau unterrichtet. In der SB-Gruppe (N = 24, 13-14 Jahre, 54 % weiblich) wurde dieses in die Lernleiter eingebettet und enthielt binnendifferenzierende Aufgaben. Die Unterrichtsgestaltung in der B-Gruppe (N = 20, 13-14 Jahre, 60 % weiblich) erfolgte durch die Lehrkraft, die ausschließlich das binnendifferenzierende Material in ihre Unterrichtsgestaltung integrierte. Es zeigen sich in beiden Gruppen signifikante Fachwissenszuwächse mit einer großen Effektstärke (SB:  $t(23) = 9.896$ ,  $p < .001$ ,  $d = 2.29$ ; B:  $t(19) = 6.082$ ,  $p < .001$ ,  $d = 1.69$ ). Ein signifikanter Gruppenunterschied hinsichtlich des Lernzuwachses ist nicht zu beobachten, deskriptiv betrachtet lernen die leistungsschwachen Schülerinnen und Schüler (Chemienoten 4-6) der SB-Gruppe jedoch mehr dazu (1.17 Punkte). Bei den durchschnittlichen Lernenden (Chemienote 3) beider Gruppen zeigen sich kaum Unterschiede (0.23 Punkte), während die leistungsstarken (Chemienoten 1-2) der SB-Gruppe einen höheren Lernzuwachs haben als die der B-Gruppe (2.44 Punkte). Die deskriptiven Ergebnisse könnten ein Hinweis darauf sein, dass eine Strukturierung durch Lernleitern lernförderlicher ist. Auf Grund der geringen Probandenzahl wird im Herbst 2017 die Hauptstudie an zwölf Gymnasien in Nordrhein-Westfalen durchgeführt. Dabei nehmen pro Schule zwei achte Parallelklassen teil, die von derselben Lehrperson entweder mit oder ohne Lernleiter unterrichtet werden (Gruppen SB + S bzw. B + KG).

Erste Ergebnisse der Hauptstudie sollen auf der GEBF 2018 präsentiert werden.

## E08

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H103

### Der Achievement Goal Questionnaire: Replikation der 3x2 Faktorenstruktur und Validierung im schulischen Kontext

**Marion Händel, Barbara Jacob, Stefan Markus**

Universität Erlangen-Nürnberg, Deutschland; [marion.haendel@fau.de](mailto:marion.haendel@fau.de)

#### Theoretischer Hintergrund

Das Konzept der Lern- und Leistungsziele (achievement goals) als wichtiger motivationaler Faktor des Lernens wurde in den letzten Jahren stetig weiterentwickelt. In der aktuellsten Konzeptualisierung nehmen Elliot, Murayama und Pekrun (2011) 3x2 Komponenten an und unterscheiden die Ziele zum einen hinsichtlich einer evaluativen Komponente mit Bezug auf Aufgaben (task), die eigene Person (self) bzw. andere Personen (other), zum anderen bezüglich einer dichotomen Valenz-Komponente hinsichtlich Annäherung (approach) und Vermeidung (avoidance). Der 3x2 Achievement Goal Questionnaire (AGQ) wurde in verschiedenen Studien mit vornehmlich studentischen Stichproben empirisch untersucht (Brondino, Raccanello & Pasini, 2014; David, 2012; Elliot et al., 2011). Ziel unserer Studien ist, eine auf den schulischen Kontext adaptierte deutsche Version hinsichtlich ihrer Faktorenstruktur zu überprüfen und mit konvergenten Maßen zu validieren.

#### Fragestellung

1. Lässt sich die Faktorenstruktur des 3x2 AGQs anhand einer deutschen Schülerversion replizieren?
2. Welche Zusammenhänge zeigen sich für den AGQ mit im deutschsprachigen Raum etablierten Skalen?
3. Welchen Einfluss haben Achievement Goals auf Schulleistungen?

#### Methode

In insgesamt zwei Fragebogenstudien wurde eine deutsche Schüler-Fassung des 3x2 AGQ eingesetzt. Zur Überprüfung der Faktorenstruktur wurden jeweils 10 Modelle mit R (Paket lavaan, Rosseel, 2012) berechnet. In Studie 1 mit 136 Schülerinnen und Schülern der 10. gymnasialen Jahrgangsstufe aus ganz Deutschland (55.9% weiblich) wurden zur Validierung die Skalen zur Erfassung der Lern- und Leistungsmotivation (SELLMO, Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002) erhoben. Diese unterscheiden Annäherungs- und Vermeidungs-Leistungsziele sowie Lernziele.

In Studie 2 wurden neben dem AGQ Zeugnisnoten in den Fächern Mathematik, Deutsch und Englisch erfasst. Daten liegen von 93 Schülerinnen und Schülern (66.3% weiblich) der 10. Jahrgangsstufe aus drei bayerischen Gymnasien vor.

#### Ergebnisse

Die Fit-Indices der konfirmatorischen Faktorenanalysen erwiesen sich für das 3x2 Modell im Vergleich zu neun alternativen Modellen jeweils am besten (Studie 1:  $X^2(120) = 188.00$ , CFI = 0.92, TLI = 0.89, RMSEA = 0.068, AIC= 6367.60, BIC = 6511.02; Studie 2:  $X^2(120) = 209.42$ , CFI = 0.92, TLI = 0.90, RMSEA = 0.090, AIC= 4381.26, BIC = 4510.42).

Zur Prüfung der konvergenten Validität wurden in der ersten Studie die AGQ-Skalen mit denen des SELLMO korreliert. Inhaltlich konvergente Skalen sind AGQ other-approach/other-avoidance und SELLMO Annäherungs-/Vermeidungsleistungsziele sowie AGQ task-approach und SELLMO Lernziele, die jeweils schwach signifikant zusammenhängen ( $.17 \leq r \leq .21$ ,  $p < .05$ ). Darüber hinaus zeigten sich auch signifikante Zusammenhänge mit anderen Skalen des AGQ, die ebenfalls allesamt schwach ausfielen.

In Studie 2 wurde wie bei Elliot et al. (2011) der Einfluss von achievement goals in separaten Regressionen auf die Zeugnisnoten untersucht. Es zeigten sich signifikant positive Effekte der other-approach goals auf die Zeugnisnoten in allen drei untersuchten Fächern ( $\beta = .57$ ,  $p < .01$  für Mathematik,  $\beta = .53$ ,  $p < .05$  für Deutsch und  $\beta = .40$ ,  $p < .05$  für Englisch) und ein negativer Effekt der other-avoidance goals auf die Zeugnisnote im Fach Deutsch ( $\beta = -.51$ ,  $p < .05$ ). Übereinstimmend mit Elliot et al. zeigten sich keine Zusammenhänge mit den Skalen self-/task-approach oder self-/task-avoidance goals; einzig für die Zeugnisnoten in Mathematik zeigte sich ein unerwarteter negativer Effekt der self-approach goals ( $\beta = -.44$ ,  $p < .01$ ).

Die Ergebnisse beider Studien zeigen, dass die von Elliot et al. (2011) postulierte Struktur des 3x2 Modells in Fragebogenerhebungen mit Schülerinnen und Schülern repliziert werden konnte. Other-approach/avoidance goals wirken sich offenbar positiv auf Schulleistungen aus. Die überraschenderweise schwachen Zusammenhänge der inhaltlich übereinstimmenden Skalen des AGQ und des SELLMO werden im Beitrag diskutiert.

### Messinvarianz von Leistungsängstlichkeit über Schulfächer hinweg – Ein Vergleich von Äpfeln mit Birnen?

**Rebecca Schneider<sup>1</sup>, Jörn Sparfeldt<sup>1</sup>, Christoph Niepel<sup>2</sup>, Samuel Greiff<sup>2</sup>**

<sup>1</sup>Universität des Saarlandes, Deutschland; <sup>2</sup>Universität Luxemburg; [rebecca.schneider@uni-saarland.de](mailto:rebecca.schneider@uni-saarland.de)

Theoretischer Hintergrund. Die pädagogisch-psychologische Relevanz von Leistungsängstlichkeit ist unbestritten, wobei insbesondere die Komponenten „Aufgeregtheit“ und „Besorgtheit“ unterschieden werden. Neuere Forschungsarbeiten verweisen zudem auf eine Leistungsängstlichkeitsstruktur entlang der Schulfachgrenzen (z.B. Götz et al., 2007; Sparfeldt et al., 2005). Voraussetzungen für Vergleiche der Leistungsängstlichkeitskomponenten Aufgeregtheit und Besorgtheit über Schulfächer hinweg sind zum einen parallele Formulierungen der Itemstämme und zum anderen ein hinreichender Grad an Messinvarianz. So ist u.a. metrische Messinvarianz zur Untersuchung von Beziehungen zwischen latenten Variablen und skalare Messinvarianz zur Analyse von Mittelwertvergleichen von Noten (Christ & Schlüter, 2012; Geiser, 2011). Ziel der vorliegenden Studie ist die Untersuchung der Messinvarianz von über verschiedene Schulfächer hinweg erfasster Leistungsängstlichkeitskomponenten.

Methode. Insgesamt bearbeiteten N = 570 finnische Schülerinnen und Schüler der sechsten Klassenstufe und N = 482 finnische Schülerinnen und Schüler der neunten Klassenstufe Aussagen zu den mit je 5 Items erfassten Leistungsängstlichkeitsfacetten Aufgeregtheit und Besorgtheit aus dem Prüfungsangstfragebogen (PAF; Hodapp, Rohrman & Ringen, 2011) – jeweils bezogen auf die Fächer Mathematik, Physik, Finnisch (Muttersprache) und Englisch (präsentiert in Gitterform; Sparfeldt et al., 2005). Getrennt für die beiden Klassenstufen wurden Modelle mit schulfachspezifischen Aufgeregtheits- und Besorgtheitsfaktoren spezifiziert. Unter Berücksichtigung der hierarchischen Datenstruktur wurde auf metrische, skalare und strikte Messinvarianz geprüft. Gewählt wurde der Step-Up-Ansatz (Brown, 2006), nach welchem mit der am wenigsten restriktiven Form von Messinvarianz begonnen wird; sollte sich der Modellfit im Vergleich zur weniger restriktiven

Form nicht bedeutsam verschlechtern ( $\Delta CFI < .01$ ; Chen, 2007), wird die jeweils restriktivere Form der Messinvarianz angenommen.

Ergebnisse. Für beide Klassenstufen zeigte sich sowohl metrische (Klassenstufe 6/9 konfigural – metrisch:  $\Delta CFI = .004/.006$ ) als auch skalare (metrisch – skalar:  $\Delta CFI = .003/.003$ ) und strikte Messinvarianz (skalar – strikt:  $\Delta CFI = .001/.002$ ). Alle untersuchten Modelle zeigten mindestens eine akzeptable Modellpassung ( $CFI > .90$ ,  $TLI > .90$ ,  $RMSEA < .08$ ). Diskriminanz-konvergente Validitätshinweise der schulfachspezifischen Leistungsängstlichkeitsmaße mit Noten und schulischen Selbstkonzepten wurden ermittelt.

Diskussion. Die Prüfung hinreichender Messinvarianz ist eine Voraussetzung für die Bearbeitung unterschiedlicher inhaltlicher Fragestellungen der Leistungsängstlichkeitsforschung. Die vorliegende Studie liefert Hinweise auf Messinvarianz für Aufregungs- und Besorgtheitsfaktoren über zentrale Schulfächer der mathematisch-naturwissenschaftlichen (Mathematik, Physik) und sprachlichen (Finnisch, Englisch) Domäne. Die Generalisierbarkeit der Befunde auf weitere Schulfächer (z.B. Geschichte) wird diskutiert. Die Erfassung von Leistungsängstlichkeit in Form eines Gitters ist besonders ökonomisch und erlaubt Vergleiche zwischen den Schulfächern (oder auch anderen Gruppen, z.B. über Klassenstufen) durch die Verwendung gleich formulierter Itemstämme. Weitere Implikationen für Forschung und Praxis werden diskutiert.

## **Schülerkompetenzen im Bereich elementarer Funktionen – Entwicklung und Gütekriterien eines Testinstruments**

**Ute Sproesser<sup>1</sup>, Markus Vogel<sup>1</sup>, Tobias Dörfler<sup>1</sup>, Andreas Eichler<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Heidelberg, Deutschland; <sup>2</sup>Universität Kassel, Deutschland; [sproesser@ph-heidelberg.de](mailto:sproesser@ph-heidelberg.de)

Theoretischer Hintergrund und Fragestellung

In der mathematikdidaktischen Forschung sind Lernschwierigkeiten in Bezug auf (elementare) Funktionen unter Schülerinnen und Schülern seit Jahrzehnten gut dokumentiert (vgl. z. B. Hadjidemetriou & Williams, 2001; Kerlake, 1981; Sproesser et al., 2016; Vogel, 2006). Um Lernende effektiv fördern zu können, ist es notwendig, mit geeigneten Testinstrumenten die Ausprägung ihrer Kompetenzen messen und Fehler feststellen zu können. Für den Inhaltsbereich linearer und quadratischer Funktionen liegt ein entsprechendes Instrument samt Dokumentation von Gütekriterien vor (Nitsch, 2015). Dieses wurde aufgrund der Fokussierung auf lineare und quadratische Funktionen von Nitsch in den Klassenstufen 9 bis 11 eingesetzt. Um Kompetenzen des Umgangs mit elementaren Funktionen in Klassenstufe 7 und 8 zu ermitteln, wurde das Testinstrument im Projekt ProfIL 9 der Pädagogischen Hochschule Heidelberg für diese Klassenstufen adaptiert. Es stellt sich die Frage nach itembezogenen und globalen Gütekriterien für dieses adaptierte Testinstrument. Aufgrund des intendierten Einsatzes in Schulklassen ist insbesondere von Interesse, inwieweit der Test Kompetenzen geschlechterfair abbildet. Die zentrale Forschungsfrage, welche die Studie verfolgt, lautet:

- Inwieweit erfüllt das adaptierte Testinstrument itembezogene und globale Gütekriterien?

Methode

Das Testinstrument wurde insgesamt unter mehr als 800 Schülerinnen und Schülern verschiedener Schularten eingesetzt. Aufgrund der Datenerhebung bis Ende des Schuljahres 2016/17 ist die Codierung der Schülertests zum Zeitpunkt der Abstract-Einreichung noch nicht vollständig abgeschlossen, so dass endgültige Ergebnisse erst auf der Tagung vorgestellt werden können. Die folgenden Angaben beziehen sich auf eine vorläufige Stichprobe von 425 Schülerinnen und Schülern (davon 212 weiblich) aus 9 siebten Gymnasial- und 11 achten Realschulklassen. Die Lernenden waren zum Testzeitpunkt zwischen 11 und 16 Jahren alt ( $M = 13,00$ ;  $SD = 0,84$ ). Um vergleichbare Bedingungen zu schaffen, erfolgten die Testungen in allen Klassen kurze Zeit nach Abschluss der Unterrichtseinheit „lineare Funktionen“.

Der Schülertest umfasst insgesamt 22 Aufgaben, davon zehn im Multiple-Choice-Format und zwölf im offenen Format. Ein von Experten mit mehrjähriger Schulpraxis vorgenommener Abgleich mit den Bildungsstandards (KMK, 2003, 2012; vgl. auch Land Baden-Württemberg, 2016a, b) und mit Schulbüchern aus Gymnasialklasse 7 bzw. Realschulklasse 8 (z. B. Freudigmann et al., 2016; Maroska et al., 2006) charakterisiert die Aufgaben als typisch für diese Klassenstufen.

Die leitende Forschungsfrage soll durch eine IRT-Modellierung der Daten unter Einbezug von Differential Item Functioning (DIF) untersucht werden. Zu diesem Zweck wird die Software ConQuest 2.0 (Wu et al., 2007) verwendet.

Vorläufige Ergebnisse und Diskussion

Analysen mit dem vorläufigen Datensatz deuten auf eine gute EAP-PV-Reliabilität von 0,866 und zufriedenstellende Item-Diskriminationen zwischen 0,35 und 0,68 hin. Die gewichteten MNSQ- und die entsprechenden t-Werte sind nicht für alle 22 Items innerhalb der empfohlenen Grenzen: Vier Items weisen hier leichte Abweichungen auf (Item 2:  $MNSQ = 0,90$ ;  $CI = [0,91; 1,09]$ ;  $T = -2,1$ ; Item 10:  $MNSQ = 0,87$ ;  $CI = [0,91; 1,09]$ ;  $T = -2,8$ ; Item 18:  $MNSQ = 0,89$ ;  $CI = [0,91; 1,09]$ ;  $T = -2,3$ ; Item 20:  $MNSQ = 1,10$ ;  $CI = [0,91; 1,09]$ ;  $T = 2,2$ ). Drei weitere Items weichen bezüglich MNSQ und t-Werten stärker ab (Item 15:  $MNSQ = 0,83$ ;  $CI = [0,90; 1,10]$ ;  $T = -3,7$ ; Item 21:  $MNSQ = 0,81$ ;  $CI = [0,91; 1,09]$ ;  $T = -4,2$ ; Item 22:  $MNSQ = 1,28$ ;  $CI = [0,90; 1,10]$ ;  $T = 5,1$ ). Des Weiteren stellte sich heraus, dass einige Items DIF bezüglich des Geschlechts aufweisen: Bei drei Items gibt es Vorteile für die Mädchen (-0,792 bzw. -0,562 bzw. -0,634 logits Differenz), während bei zwei Items die Jungen besser abschnitten (0,576 bzw. 0,782 logits Unterschied). Vor einer möglichen systematischen Nutzung des Testinstruments zur Ermittlung von Schülerkompetenzen muss auf Grundlage der endgültigen Ergebnisse reflektiert werden, inwieweit der Ausschluss einzelner Items notwendig und sinnvoll ist.

## **Grit vs. Gewissenhaftigkeit: Verortung der Facetten und inkrementelle Validität**

**Fabian T. C. Schmidt<sup>1</sup>, Clemens Lechner<sup>2</sup>, Daniel Danner<sup>2</sup>**

<sup>1</sup>IPN Kiel, Deutschland; <sup>2</sup>GESIS Institut, Mannheim; [fschmidt@ipn.uni-kiel.de](mailto:fschmidt@ipn.uni-kiel.de)

Neben kognitiven Fähigkeiten und domänenspezifischen Kompetenzen erlangen nichtkognitive Fähigkeiten immer mehr Aufmerksamkeit in der Forschung zu Einflussvariablen im Kontext von Bildung (Farrington et al., 2012). Dem Konstrukt Grit kommt dabei neben wissenschaftlichem auch großes Gesellschaftliches Interesse zuteil (Tough, 2012; U.S. Department of Education, 2013). Grit ist eine Persönlichkeitseigenschaft, welche ihren Ursprung in der positiven Psychologie hat und die Fähigkeit beschreibt langfristige Ziele mit Enthusiasmus zu verfolgen (Facette beständiges Interesse) und dabei Rückschläge und Herausforderungen erfolgreich zu bewältigen (Facette Beharrlichkeit; Duckworth & Quinn, 2009). Grit stellte sich in der Vergangenheit als Prädiktor für Erfolg bei der Verfolgung relevanter, selbstgewählter und weit in der Zukunft liegender Ziele dar. Bisher liegt ein breiter Corpus an Befunden bzgl. der Vorhersagekraft und den Zusammenhängen zu einem breiten



Spektrum an Erfolgskriterien aus internationaler Forschung vor (Muenks, Wigfield, Yang & O'Neal, 2007, Strayhorn, 2014). In ihrer Metaanalyse konnten Credé, Tynan und Harms (2016) jedoch zeigen, dass Grit und Gewissenhaftigkeit kaum voneinander trennbar sind und kritisieren die mangelnde Abgrenzung der beiden Konstrukte. Sie vermuten zudem, dass differentielle Zusammenhänge zwischen den Konstrukten auf Subfacettenebene bestehen. Es bleibt zudem offen, inwiefern Grit inkrementelle Vorhersagekraft über Gewissenhaftigkeit hinaus besitzt.

Ziel dieser Studie ist es, die Facetten von Grit erstmalig in die hierarchische Struktur der Gewissenhaftigkeit (Roberts, Chernyshenko, Stark & Goldberg, 2005) auf drei Ebenen (Übergeordneter Gewissenhaftigkeitsfaktor, Industriousness-Faktor, Facettenebene) einzuordnen. Nach der Integration von Grit, sollen Befunde zu der inkrementellen Validität von Grit über Gewissenhaftigkeit hinaus präsentiert werden.

In zwei Studien (NGes = 1235) sollen die Facetten von Grit (BISS-8; Schmidt, Fleckenstein, Retelsdorf, Eskreis-Winkler, Möller, 2017) zum ersten Mal anhand von erweiterten CFA (extension procedure for confirmatory factor analyses; Nagy, Brunner, Lüdtko, & Greiff, 2017) in die Facetten der Gewissenhaftigkeit des NEO-PI-R (Ostendorf & Angleitner, 2004) eingeordnet werden. In einer weiteren Studie (N = 1435) werden zudem Befunde zu der inkrementellen Vorhersagekraft von Grit über Gewissenhaftigkeit hinaus in Bezug auf Lebenszufriedenheit, Einkommen, Gesundheit und Erwerbsstatus untersucht. Zur Kontrolle dieser Zusammenhänge liegen Daten zu Geschlecht, Alter, der kristallinen Intelligenz, und dem Bildungsgrad der Studienteilnehmerinnen und -teilnehmer vor.

Die Ergebnisse der ersten Studien zeigen, dass die Facette Beharrlichkeit 95 % der Varianz mit den übergeordneten Faktoren teilt und eng mit dem Faktor Industriousness auf der mittleren Hierarchieebene verbunden ist. Die Facette beständiges Interesse teilte weniger Varianz mit den Übergeordneten Faktoren (53 %), korrelierte aber mit der Facette Selbstdisziplin auf der unteren Hierarchieebene signifikant. Der globale Gritfaktor zeigt Zusammenhänge ähnlich der Facette Beharrlichkeit. Die Befunde weisen darauf hin, dass Grit und die Facetten von Grit in die hierarchische Struktur der Gewissenhaftigkeit integrierbar sind. Hierbei steht der Befund im Vordergrund, dass Grit gerade den proaktiven Aspekt der Gewissenhaft (Industriousness) widerspiegelt, der gerade in schulischen Kontexten von besonderer Bedeutung ist (Roberts, Lejuez, Krueger, Richards, & Hill, 2014). Anschließend werden Befunde zu der inkrementellen Vorhersagekraft von Grit und Gewissenhaftigkeit präsentiert und Implikationen für die weitere Forschung zu nichtkognitiven Fähigkeiten diskutiert.

## E09

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus S107

### Ist Schulabsentismus ansteckend? Schulische Kontexteffekte im Schuljahr nach PISA 2012

**Christine Sälzer**

TU München, Deutschland; [christine.saelzer@tum.de](mailto:christine.saelzer@tum.de)

#### Theoretischer Hintergrund und Ziel

Schulabsentismus ist ein weit verbreitetes Verhalten bei schulpflichtigen Jugendlichen (z. B. OECD, 2014). Diese Studie untersucht, wie die individuelle Einschätzung des Klassenkontexts mit dem Absentismus einzelner Schüler zusammenhängen kann. Im Fokus steht dabei das absichtliche unerlaubte Fernbleiben vom Unterricht ("Schulschwänzen") im Gegensatz zu gesetzlich erlaubten Absenzen etwa aufgrund von Krankheiten oder Notfällen.

Basierend auf Wallers (2014) Konzept von Schulen als einer sozialen Organisationseinheit nehmen wir an, dass individuelles Schülerverhalten auf deren eigene Einschätzung ihrer Situation in der Schule zurückgeht. Eine solche Einschätzung entsteht aus der Selbstwahrnehmung des Schülers im alltäglichen Kontext der Institution Schule: Schüler schätzen mögliche Handlungsoptionen bezüglich ihrer Chancen und Risiken ein und wählen eine, die ihnen attraktiv erscheint (Bidwell, 2006). Beobachten sie beispielsweise bei Klassenkameraden, dass geschwänzte Schulstunden ungeahndet bleiben und währenddessen angenehme Dinge erlebt werden können, so kann Schulschwänzen durchaus als attraktive Option eingeschätzt werden.

#### Fragestellungen

- (1) Schwänzen Schüler häufiger, je höher der Anteil schwänzender Mitschüler in ihrer Klasse ist?
- (2) Gibt es Subgruppen von Schülern, die erst im Schuljahr nach dem PISA-Test beginnen zu schwänzen und wenn ja, welche Eigenschaften haben diese gemeinsam?
- (3) Sind Schüler in Klassen mit einem hohen Absentistenanteil weniger kompetent in den PISA-Domänen Lesen, Mathematik und Naturwissenschaften?

#### Methode

Unsere Studie nutzt Daten einer Messwiederholung der klassenbasierten Stichprobenerweiterung von PISA 2012 in Deutschland (n = 4100 Schüler in 247 zehnten Klassen in 127 Schulen). Erfasst wurden fachspezifischer und allgemeiner Absentismus zum Zeitpunkt von PISA 2012 (t1) und ein Jahr später (t2) sowie die Kompetenzwerte für Lesen, Mathematik und Naturwissenschaften (beide Messzeitpunkte).

Den drei Fragestellungen folgend wurden die Daten in drei Schritten ausgewertet:

- (1) Die selbstberichtete Absenzenhäufigkeit der Schüler wird mit der ebenfalls selbstberichteten Anzahl schwänzender Mitschüler korreliert. Zur Kontrolle werden Mehrebenenanalysen durchgeführt, um aggregierte Absenzenhäufigkeiten auf Klassenebene als Prädiktor für individuelle Absenzenhäufigkeiten zu verwenden.
- (2) Schüler, die zum Zeitpunkt PISA 2012 noch nicht geschwänzt hatten, ein Jahr später jedoch schon, wurden aus dem Datensatz ausgewählt und ihre Daten deskriptiv ausgewertet (Geschlecht, sozialer Hintergrund, Noten, PISA-Punkte in Lesen, Mathematik und Naturwissenschaften). Odds ratios wurden berechnet, um die Wahrscheinlichkeit zu bestimmen, mit der ein Schüler erst im Schuljahr nach PISA mit dem Schwänzen beginnt. Dabei werden Schüler mit unterdurchschnittlichem Abschneiden in PISA 2012 von solchen unterschieden, die mindestens eine durchschnittliche Kompetenz aufweisen.
- (3) Die Schulstichprobe wird unterteilt in solche mit einem Absentistenanteil von >10 Prozent (n = 42) und <10 Prozent (n = 85). Die Mittelwerte in den PISA-Kompetenzbereichen werden berechnet und zwischen beiden Schulgruppen verglichen.

#### Ergebnisse

Frage 1: Sowohl die Anzahl schwänzender Klassenkameraden (Selbstbericht) als auch der aggregierte Anteil schwänzender Mitschüler pro Klasse korreliert mit der individuellen Absenzenhäufigkeit. Je mehr Klassenkameraden schwänzen, desto eher schwänzt auch der einzelne Schüler selbst ( $r = .38$ ;  $p > .01$ ; Multilevel-Regression:  $B = .29$ ;  $SE = .02$ ).

Frage 2: Besonders Schüler, die in PISA 2012 mindestens in zwei PISA-Kompetenzbereichen unterdurchschnittlich abschneiden, begannen erst im Schuljahr nach PISA zu schwänzen (n = 644; Jungen = 60%). Unter Kontrolle von Geschlecht, sozialem Hintergrund und Noten begannen Schüler, die mindestens eine Standardabweichung unter dem Durchschnitt abschneiden, signifikant häufiger mit dem Schwänzen als ihre Mitschüler (OR = 4.84; 95% CI [3.63, 6.01]).

Frage 3: Schulen mit einer höheren Absenzenrate unterscheiden sich bezüglich der durchschnittlichen Kompetenzwerte in Mathematik (-36 Punkte; SE = 4.5) und Naturwissenschaften (-25 Punkte; SE = 3.6) von Schulen mit einer niedrigeren Absenzenrate, nicht jedoch im Lesen.

Die Ergebnisse der Studie können als Grundlage für die Planung von Interventionen genutzt werden, die die Prävention oder Reduktion von Absentismus in der Sekundarstufe zum Ziel haben.

### Die Rolle der Klassenebene für Hausaufgabenlerntypen von Schülerinnen und Schülern

**Barbara Flunger<sup>1</sup>, Ulrich Trautwein<sup>2</sup>, Benjamin Nagengast<sup>2</sup>, Oliver Lüdtke<sup>3</sup>, Alois Niggli<sup>4</sup>, Inge Schnyder<sup>5</sup>**

<sup>1</sup>Universität Utrecht, Niederlande; <sup>2</sup>Hector-Institut für Empirische Bildungsforschung, Universität Tübingen; <sup>3</sup>IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Universität Kiel; <sup>4</sup>Pädagogische Hochschule Freiburg;

<sup>5</sup>Universität Freiburg; [b.flunger@uu.nl](mailto:b.flunger@uu.nl)

Personenzentrierte Verfahren (z.B. Clusteranalyse oder latente Profilanalyse) werden in der empirischen Bildungsforschung oft angewandt, da mit diesen Methoden potentielle Unterschiede in einer Reihe von Aspekten erforscht und dabei mögliche Interaktionen zwischen verschiedenen Variablen untersucht werden können (z.B., Asendorpf, 2014; Wormington & Linnenbrink-Garcia, 2016). Flunger et al. (2015) konnten unter Anwendung der latenten Profilanalyse fünf unterschiedliche Lerntypen bei der Hausaufgabenbearbeitung identifizieren, die sich bezüglich des Engagements der Schülerinnen und Schülern bei der Hausaufgabenbearbeitung unterscheiden. Auf Basis ihrer Mittelwertsprofile in HA-Zeitaufwand und HA-Compliance wurden sie bezeichnet als „schnelle Lerner“ (hohe HA-Compliance, niedriger Zeitaufwand), „hoch engagierte Lerner“ (hohe HA-Compliance, hoher Zeitaufwand), „Durchschnittsschüler“ (mittlere HA-Compliance, niedriger Zeitaufwand), „sich abmühende Lerner“ (niedrige HA-Compliance, hoher Zeitaufwand) und „Minimalisten“ (niedrige HA-Compliance, niedriger Zeitaufwand).

Unter Kontrolle der Ausgangsleistung zeigte sich, dass Lerntypen, die durch hohe HA-Compliance charakterisiert wurden („schnelle Lerner“ und „hoch-engagierte Lerner“), einen höheren Lerngewinn durch ihr Hausaufgabenverhalten aufwiesen als die anderen Lerntypen.

Das Bildungssystem weist eine hierarchische Struktur auf (Schüler sind genestet in Klassen). Deshalb könnte es zum Beispiel der Fall sein, dass in Klassen von Lehrkräften, die durch HA die Motivation ihrer Schülerinnen und Schüler steigern wollen, seltener sich abmühende Lerner oder Minimalisten zu identifizieren sind, als bei Lehrkräften, die HA als Mittel der Leistungssteigerung einsetzen und weniger stark auf das Interesse der Schülerinnen und Schüler an den gestellten Aufgaben achten. Die latente Profilanalyse kann auf den Mehrebenenkontext erweitert werden und somit wird die Berücksichtigung von Prädiktoren auf unterschiedlichen Analyseebenen ermöglicht. Solche „Multilevel mixture“- Modelle wurden bisher in der erziehungswissenschaftlichen Forschung selten angewandt, obwohl sie das Potential haben die multidimensionale und hierarchische Struktur von Daten zu modellieren und dadurch neue und wichtige Erkenntnisse zu generieren (z.B., Finch & Marchant, 2013).

Die vorliegende Studie ist Teil eines DFG-Forschungsprojekts, das darauf abzielt verschiedene Konfigurationen des HA-Verhaltens und Prädiktoren auf Lehrer- und Schülerebene zu erforschen. Die Datengrundlage umfasst eine längsschnittliche Studie mit 1915 Schweizer Schülerinnen und Schüler der 8. Jahrgangsstufe aus 77 Klassen, die zu ihrem HA-Verhalten im Fach Französisch (September/Oktober und Juni 2004) befragt worden waren. Die Studie untersuchte, ob in manchen Klassen häufiger ein bestimmter HA- Lerntyp auftritt als in anderen Klassen, ob darin spezifische Muster (latente Profile auf Klassenebene) zu unterscheiden sind, und ob sich dieses Befundmuster replizieren lässt. Darüber hinaus wurde untersucht, inwieweit das HA-Vergabeverhalten von Lehrkräften oder spezifische Einstellungen zu HA die Klassifikation ihrer Klassen beeinflussen.

Zu beiden Zeitpunkten konnten in der Verteilung der fünf HA-Lerntypen über die 77 Klassen drei latente Profile auf Klassenebene identifiziert werden. Zwei Profile auf Klassenebene wurden repliziert: das „Durchschnittsklassen“-Profil, in der alle fünf HA-Lerntypen gefunden wurden, und das „geringe Zeitaufwand“- Profil, in dem hoch-engagierte und sich abmühende Lerner wenig wahrscheinlich waren. Das dritte Profil auf Klassenebene unterschied sich zu den beiden Zeitpunkten, wies jedoch jeweils ein theoretisch bedeutsames Muster auf. Zum ersten Zeitpunkt wurde ein „hohes Compliance“-Klassenprofil identifiziert, in dem hoch-engagierte und schnelle Lerner besonders häufig gefunden wurden. Zum zweiten Zeitpunkt wurde ein „niedrige Compliance“-Klassenprofil identifiziert, das durch höhere Anzahlen und Wahrscheinlichkeiten von Durchschnittsschülern und sich abmühenden Lernern charakterisiert wurde. Bei Einbezug des HA-Vergabeverhaltens und Einstellungen zu HA von Lehrkräften als Prädiktoren zeigte sich, dass Klassen von Lehrkräften mit geringerem Ziel, durch Hausaufgaben Übung des Lernmaterials und Leistungssteigerung zu fördern, eine höhere Wahrscheinlichkeit hatten dem „niedrige Compliance“ Klassenprofil zugewiesen zu werden. Zudem wurden Klassen von Lehrkräften, die dem Ziel HA die Motivation ihrer Schülerinnen und Schüler steigern wollen weniger Bedeutung zusprachen, tendenziell eher zu dem „geringe Zeitaufwand“-Klassenprofil zugewiesen. Diese Ergebnisse weisen darauf hin, dass Lehrkräfte und ihre hausaufgaben-spezifischen Einstellungen die Ausprägung eines bestimmten Lernstils von Schülerinnen und Schülern bei den Hausaufgaben beeinflussen können.

### **Kollaboratives Problemlösen in PISA 2015: Deutschland im Fokus**

**Miriam Weis<sup>1</sup>, Fabian Zehner<sup>2</sup>, Vogel Freydis<sup>3</sup>, Leutner Detlev<sup>4</sup>, Reiss Kristina<sup>1</sup>**

<sup>1</sup>Technische Universität München, TUM School of Education, Zentrum für internationale Vergleichsstudien (ZIB), Deutschland; <sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Zentrum für technologiebasiertes Assessment (TBA); <sup>3</sup>Technische Universität München, TUM School of Education, Lehrstuhl für Schulpädagogik; <sup>4</sup>Universität Duisburg-Essen, Fakultät für Bildungswissenschaften, Lehrstuhl für Lehr-Lernpsychologie; [mirjam.weis@tum.de](mailto:mirjam.weis@tum.de), [Fabian.Zehner@dipf.de](mailto:Fabian.Zehner@dipf.de)

Sowohl in der Schule als auch im Arbeits- und Privatleben werden Menschen häufig mit Problemen konfrontiert, für welche die zur Lösung benötigten Informationen oder Fertigkeiten über verschiedene Individuen verteilt sind. In diesen Situationen ist gelingende Kollaboration nicht nur förderlich, sondern eine notwendige Voraussetzung. Speziell in der Berufsausübung sehen sich Arbeitende immer mehr nicht-routinierten und interpersonal zu lösenden Problemen gegenüber (Hermann et al., 2017; Levy & Murnane, 2013), aber auch in der Bildungspraxis findet sich eine verstärkte Nutzung kollaborativer Lernmittel (Mayrberger & Bettinger, 2014). Daher untersucht das Programme für International Student Assessment (PISA) 2015 erstmals die kollaborative Problemlösekompetenz bei fünfzehnjährigen Schülerinnen und Schülern am Ende ihrer Pflichtschulzeit. Kurzgefasst definiert PISA diese Kompetenz als die Fähigkeit, sich effektiv und zielorientiert am Lösungsprozess einer Problemsituation mit mindestens zwei Akteuren zu beteiligen (OECD, 2017). Diese Studie untersucht die kollaborative Problemlösekompetenz von Fünfzehnjährigen im internationalen Vergleich mit einem Fokus auf Deutschland. Außerdem werden deutschlandspezifische Fragestellungen analysiert, wie etwa der Vergleich der Gymnasien mit nicht-gymnasialen Schularten.

In PISA 2015 wurden Fragebögen eingesetzt, die Aspekte kollaborativen Lernens im Unterricht bei Lehrkräften sowie Schülerinnen und Schülern erhoben. Darüber hinaus wurde ein neuer computerbasierter Test verwendet, der eine simulationsbasierte Erhebung am Computer mit programmierten, standardisierten Teammitgliedern darstellt (vgl. Kyllonen, Zhu & von Davier, 2017). Dabei können viele verschiedene Teamkonstellationen (etwa Teamgröße; vgl. Salas, Reyes & Woods, 2017), Randbedingungen und Problemstellungen konfiguriert werden, um die Komplexität möglicher Interaktionen im Testergebnis zu reflektieren (Graesser, Dowell & Clewley, 2017) – so können zum Beispiel über die Aufgaben hinweg die Geschlechter der Teammitglieder wie auch die Teamhierarchie variiert werden. In der Testung agieren die getesteten Schülerinnen und Schüler im Problemraum und interagieren mit simulierten Akteuren zumeist via Chat oder E-Mail, indem sie aus vorgegebenen Antworten wählen. Die mit solchen Tests generierten Daten stellen ein Novum für die Bildungsforschung dar und sind außerdem mit methodischen Herausforderungen verbunden (vgl. von Davier, 2017).

Aktuell entstehen vielzählige Arbeiten zu methodischen Fragen der kollaborativen Problemlösekompetenz (z. B. Liu, Hao, von Davier, Kyllonen & Zapata-Rivera, 2016). Dabei werden Prinzipien der Aufgabenkonstruktion untersucht (z. B. Scoular, Care & Hesse, 2017), und häufig vergleichen die Studien die Kommunikation der Testpersonen in Problemlösesituationen mit simulierten Computeragenten und mit Menschen (z. B. Rosen, 2017; Rosen & Mosharraf, 2016; Scoular et al., 2017). Weiterhin werden Methoden vorgeschlagen, um etwa mit der Komplexität der resultierenden Daten oder mit den Abhängigkeiten der Antworten innerhalb eines Szenarios umzugehen. Dabei werden neu aufkommende Paradigmen wie Computational Psychometrics ins Spiel gebracht (von Davier, 2017), die unter anderem Psychometrie mit Maschinenlernen kombinieren.

Die Präsentation wird Ergebnisse zu den PISA-Schülerinnen und -Schülern in Deutschland im kollaborativen Problemlösen im Hinblick auf ausgewählte Fragestellungen adressieren. Neben den Ergebnissen aus Deutschland im internationalen Vergleich werden außerdem die selbstberichteten Einstellungen zum kollaborativen Arbeiten sowie dessen eingeschätzter Mehrwert präsentiert. Die Befunde differenzieren dabei stets auch nach Geschlecht und besuchter Schulart. Außerdem wird berichtet, wie häufig Lehrkräfte in Deutschland unterschiedliche Formen der Gruppenarbeit nutzen.

PISA 2015 erhob Daten von  $n = 535.791$  Fünfzehnjährigen in 72 Staaten, wovon 51 Staaten die kollaborative Problemlösekompetenz testeten. Aufgrund des Datenembargos der OECD bis zum 21. November 2017 können in diesem Abstract noch keine Ergebnisse berichtet werden. Der Vortrag wird aber neben der Skizzierung des theoretischen Modells und der Darstellung der Operationalisierung in Form des Tests vor allem die empirischen Ergebnisse präsentieren und im Hinblick auf die obengenannten Fragestellungen und methodischen Herausforderungen diskutieren.

## **Die Erfassung von Kontexteinflüssen durch ambulantes Assessment: Wie beeinflusst das Bearbeitungsumfeld die wahrgenommene Hausaufgabenqualität und -bearbeitung von Fünftklässler\*innen?**

**Brigitte Brisson, Désirée Theis**

Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland; [brigitte.brisson@dipf.de](mailto:brigitte.brisson@dipf.de)

### Theoretischer Hintergrund

Hausaufgaben bezeichnen Tätigkeiten, welche Schüler\*innen i.d.R. zuhause für die Schule erledigen (Höhmann & Schaper, 2008). Mit der Hausaufgabenvergabe zielen Lehrkräfte u.a. darauf ab, dass ihre Schüler\*innen wichtige Lerninhalte eigenverantwortlich vertiefen und dadurch ihre Leistung verbessern (Standop, 2013). Ausschlaggebend für eine positive Wirkung von Hausaufgaben sind jedoch u.a. die wahrgenommene Hausaufgabenqualität und individuelle Merkmale der Schüler\*innen: Hausaufgaben unterstützen die Schülerleistung insbesondere dann, wenn Schüler\*innen diese als kognitiv aktivierend und nützlich beurteilen und diese konzentriert und sorgfältig bearbeiten (Trautwein et al., 2006).

Durch den flächendeckenden Ausbau der Ganztagschulen in Deutschland werden Hausaufgaben inzwischen zunehmend im schulischen Kontext bearbeitet. Ein Großteil der Schulen bietet eine Hausaufgabenbetreuung an; die Anmeldezahlen steigen (StEG-Konsortium, 2015). Zudem ersetzen viele Ganztagschulen die traditionellen Hausaufgaben durch Aufgaben, die in sogenannten Lernzeiten im Rahmen des Pflichtunterrichts bearbeitet werden. Diese Lernzeiten finden üblicherweise mit einer Lehrkraft im Klassenverband statt. Es ist anzunehmen, dass entsprechende Veränderungen im Bearbeitungskontext die wahrgenommene Qualität der (Haus-)Aufgaben und Bearbeitungsmerkmale wie Konzentration und Anstrengungsbereitschaft der Schüler\*innen beeinflussen.

Gleichzeitig stammen bisherige Befunde aus der Hausaufgabenforschung häufig aus klassischen Fragebogenstudien (z.B. Trautwein et al., 2006). Eine Befragung über die Smartphones der Schüler\*innen ermöglicht es, ihre Einschätzung zu den (Haus-)Aufgaben im Moment und über einen längeren Zeitraum anhand wiederholter Kurzfragebögen zu erfassen. Neben der Gewinnung von „live“-Daten ermöglichen solch ambulante Erhebungsverfahren auch die Untersuchung von intraindividuellen Verläufen und wertvolle neue Einblicke in die Stabilität der (Haus-)Aufgabenwahrnehmung und -bearbeitung (Fahrenberg et al., 2007).

### Fragestellung

Die vorliegende Studie untersucht, ob sich die wahrgenommene (Haus-)Aufgabenqualität und Bearbeitungsmerkmale unterscheiden, wenn Schüler\*innen ihre Aufgaben zuhause (HA-Schüler\*innen) oder in der Schule im Rahmen von Lernzeiten (LZ-Schüler\*innen) erledigen. Infolge der Eingliederung der Hausaufgaben in die reguläre Unterrichtszeit ist die mittlere wahrgenommene Aufgabenqualität (kognitive Aktivierung, Nützlichkeit) bei LZ-Schüler\*innen möglicherweise höher und stabiler als bei HA-Schüler\*innen. Das erhöhte Ablenkungspotential durch die Klassenkamerad\*innen könnte für eine geringere mittlere Konzentration mit höheren intraindividuellen Schwankungen bei LZ-Schüler\*innen verglichen mit HA-Schüler\*innen sorgen. Der offizielle Charakter der Aufgaben und die schulische Betreuung könnten wiederum zu einer höheren mittleren Anstrengungsbereitschaft und geringeren intraindividuellen Schwankungen bei LZ- als bei HA-Schüler\*innen führen.

### Methode

Im Rahmen eines ambulanten Assessments erhielten hundert Fünftklässler\*innen aus fünf hessischen Integrierten Gesamtschulen vier Wochen lang täglich einen kurzen Fragebogen auf ihr Smartphone. Im Mittel füllten die Schüler\*innen je 8 Fragebögen aus. Der Fragebogen enthielt je 1 Frage zur Interessantheit (kognitive Aktivierung) und Nützlichkeit der (Haus-)Aufgaben sowie zur Konzentration und Anstrengung bei der Aufgabenbearbeitung (4-Punkt-Likert-Skalen). Mittels t-Tests wurde in SPSS untersucht, ob sich Hausaufgaben- und Lernzeit-Schüler\*innen auf diesen Variablen im Mittel (M) und bezüglich intraindividuellen Schwankungen (mittlere individuelle Standardabweichung vom individuellen Mittel, MSD) signifikant unterscheiden. Vertiefende Analysen mit einer erweiterten Stichprobe ( $N = 330$ ) werden ab November 2017 durchgeführt.

### Ergebnisse und Diskussion

Erste Ergebnisse weisen darauf hin, dass sich die untersuchten Schülergruppen sowohl bezüglich der wahrgenommenen Aufgabenqualität als auch der (Haus-)Aufgabenbearbeitung unterscheiden. HA-Schüler\*innen empfanden ihre Aufgaben als interessanter ( $M(HA) = 2.98$ ;  $M(LZ) = 2.68$ ;  $p = .078$ ; randsignifikante Differenz) und nützlicher ( $M(HA) = 3.09$ ;  $M(LZ) = 2.66$ ;  $p = .033$ ) und berichteten eine höhere Konzentration ( $M(HA) = 3.40$ ;  $M(LZ) = 3.10$ ;  $p = .037$ ) bei deren Bearbeitung als LZ-Schüler\*innen. Keine Unterschiede fanden sich in der mittleren Anstrengungsbereitschaft, jedoch schwankte die Anstrengungsbereitschaft bei LZ-Schüler\*innen mehr ( $MSD(HA) = .383$ ;  $MSD(LZ) = .588$ ;  $p = .058$ ; randsignifikante Differenz). Weitere Unterschiede in den intraindividuellen Schwankungen waren nicht signifikant.

Die vorliegenden Befunde liefern erste Hinweise auf Unterschiede bei der Wahrnehmung und Bearbeitung von Aufgaben im schulischen Kontext im Vergleich zur traditionellen Hausaufgabenbearbeitung. Da Schulen zunehmend von traditionellen Hausaufgabenkonzepten abweichen und vermehrt Lernzeiten einführen, ergibt sich ein erhöhter Bedarf an entsprechenden Studien. In einem nächsten Schritt soll untersucht werden, inwiefern die ambulant erhobenen Qualitäts- und Bearbeitungsmerkmale von (Haus-)Aufgaben mit den Leistungen der Schüler\*innen zusammenhängen.

## E10

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 32, H223

### **Klasseneinschätzungen der Klassenführung - ein zulässiges Instrument der Professionalisierung?**

**Georg Krammer, Barbara Pflanzl**

Pädagogische Hochschule Steiermark, Österreich; [georg.krammer@phst.at](mailto:georg.krammer@phst.at)

#### Einleitung

Metaanalysen zeigen, dass Klassenführung ein wesentlicher Bestandteil professionellen Lehrerhandelns ist (Hattie, 2009). Dementsprechend bedeutend ist es, dass die Fähigkeit Klassen zu führen bereits im Lehramtsstudium gefördert wird (Hellermann et al., 2015). Der Vergleich der Selbsteinschätzung der eigenen Klassenführung mit der Fremdeinschätzung durch SchülerInnen kann hierbei eine wichtige Grundlage für den Professionalisierungsprozess darstellen, da SchülerInnen nach mehrjähriger Schulzeit über vielfältige Erfahrungen mit Unterricht verfügen und sie aus der Perspektive der Betroffenen urteilen (Ditton, 2002). Darüber hinaus ist die Wahrnehmung unterrichtlicher Handlungen von Lehrenden durch SchülerInnen ein Prädiktor für das Lernen der SchülerInnen (Helmke, 2012). Um in einem Professionalisierungsprozess die Perspektive der SchülerInnen mit der Selbsteinschätzung der Lehrkraft vergleichen zu können ist es eine Bedingung, dass Selbsteinschätzung und Fremdeinschätzung dasselbe messen. Erst wenn dies gegeben ist, kann der Vergleich der beiden Einschätzungen aussagekräftig sein und zur persönlichen Weiterentwicklung herangezogen werden. Der vorliegende Beitrag greift diese messtheoretische Voraussetzung auf und stellt die Frage: Können Lehramtsstudierende die Selbsteinschätzungen der eigenen Klassenführung direkt mit den Einschätzungen durch SchülerInnen vergleichen, oder sind Unterschiede in den Einschätzungen behaftet mit einem Bias in der Messung?

#### Methode

Um dieser Frage nachzugehen schätzten 4851 SchülerInnen die Klassenführung von 272 Lehramtsstudierenden mittels dem Linzer Diagnosebogen zu Klassenführung (LDK: Lenske & Mayr, 2015) in den Dimensionen Unterrichtsgestaltung, Beziehungsförderung und Kontrolle ein. Zusätzlich schätzten alle Lehramtsstudierende ihre eigene Klassenführung ein. Die Klassenführung wurde auf SchülerInnen-, Klassen- und Lehrendenperspektive mittels Multilevel-Strukturgleichungsmodellen modelliert (vgl. Muthén & Satorra, 1995). Innerhalb der Multilevel-Strukturgleichungsmodelle wurde die Messinvarianz zwischen der Lehrenden- und Klassenebenen geprüft.

#### Ergebnisse & Diskussion

Starke Messinvarianz hielt für die Dimensionen Unterrichtsgestaltung und Beziehungsförderung. Für die Dimension Kontrolle konnte nur partielle schwache Messinvarianz gezeigt werden, da eines der Items einen non-uniformen Bias aufwies. Aufbauend auf dieser partiellen schwachen Messinvarianz, waren die Intercepts aller Items invariant. Alle latenten Interkorrelationen waren signifikant. Obwohl die Interkorrelationen innerhalb der Klassen ( $.875 \leq r \leq .941$ ) augenscheinlich höher ausfielen als innerhalb der Lehramtsstudierenden ( $.566 \leq r \leq .815$ ), unterschieden sie sich nicht signifikant. Die vorliegenden Befunde zeigen, dass es grundsätzlich zulässig ist Klassenführung zwischen Klassen und Lehramtsstudierenden direkt zu vergleichen. Es ist jedoch trotzdem Vorsicht geboten, da einzelne Items besser zwischen Lehrenden differenzieren zu scheinen als zwischen Klassen.

Die latenten Mittelwerte aller drei Dimensionen unterschieden sich: Die Dimensionen Unterrichtsgestaltung und Beziehungsförderung waren höher ausgeprägt für die Klassen als für die Lehramtsstudierenden, während die Dimension Kontrolle niedriger für die Klassen ausgeprägt war als für die Lehramtsstudierenden. Interessanterweise schätzten sich Lehramtsstudierende ungünstiger ein, als ihre Klassen sie einschätzten. Dies steht im Gegensatz zu Studien an im Beruf stehenden Lehrkräften, welche normalerweise eine zu günstigere Selbsteinschätzung der Lehrkräfte zeigen (z.B. Wubbels & Brekelmans, 2005). Eine mögliche Erklärung wäre hierfür, dass die in diesem Beitrag gezeigte Messinvarianz nicht auf im Beruf stehende Lehrkräfte und deren Klassen generalisierbar ist, und Unterschiede dort auf einen Bias in der Messung zurückzuführen sind. Andererseits kann auch nicht ausgeschlossen werden, dass Lehramtsstudierende sich im Laufe der anschließenden Ausübung des Berufs in ihrer Selbsteinschätzung verändern und mit der Zeit dazu tendieren sich besser als ihre Klassen einzuschätzen.

Zusammenfassend kann gesagt werden, dass die von Klassen eingeschätzte Klassenführung mit der Selbsteinschätzung direkt verglichen werden kann. Damit kann dieser Vergleich der Einschätzungen als Basis für einen Professionalisierungsprozess in der Klassenführung herangezogen werden. Allerdings zeigen die vorliegenden Befunde auch, dass dieses SchülerInnenfeedback nicht selbstverständlich mit der Selbsteinschätzung von Lehramtsstudierenden verglichen werden darf.

### **Theorie und Praxis: Veränderungen im Wissen über Classroom Management bei Lehramtsstudierenden**

**Sabine Schlag, Sabine Glock**

Bergische Universität Wuppertal, Deutschland; [schlag@uni-wuppertal.de](mailto:schlag@uni-wuppertal.de)

Classroom Management ist ein Bereich des Unterrichtens und Lehrens, der vielen Studierenden Bedenken bereitet (Evertson & Weinstein, 2006). Studierende fühlen sich oftmals unzureichend auf diese Aufgabe vorbereitet und bemängeln die große Kluft zwischen der theoretischen Universitätsausbildung und der Praxis in den Schulen (Maskan, 2007). Dabei empfinden sie die universitären Veranstaltungen zumeist als zu abstrakt und vermissen konkrete Handlungsvorschläge (Darling-Hammond, 2010). Die Verzahnung von Theorie und Praxis erscheint jedoch enorm wichtig, gerade weil Classroom Management Strategien hauptsächlich durch die eigene Unterrichtspraxis erworben werden (van Tartwijk, den Brok, Veldman & Wubbels, 2009). Allerdings besteht auch die Gefahr, dass in der Praxis Strategien erlernt werden, die nur kurzfristigen, aber keinen langfristigen Erfolg im Klassenraum bieten, und sich somit in der Forschung als wenig effektiv herausgestellt haben. Deshalb ist es wichtig, bereits in Rahmen des Studiums schulpraktische Lerngelegenheiten zu schaffen. Diese werden durch das Praxissemester gewährleistet, das an vielen Universitäten mittlerweile fest im Lehramtsstudium integriert ist. Die Praxisphase wird von bildungswissenschaftlichen Vorbereitungs- und Nachbereitungskursen begleitet und die Studierenden können im schulpraktischen Teil durch die Hospitation und das eigene Unterrichten unter Supervision durch erfahrene Lehrkräfte, Strategien beobachten und testen. Somit kann angenommen werden, dass Studierende über die theoretische Auseinandersetzung mit den Annahmen und Strategien eines effektiven Classroom Managements einen Wissenszuwachs erwerben sollten. Aufgrund der Praxisabhängigkeit des Erwerbs von Classroom Management Strategien sollte der Wissenszuwachs in der Praxisphase jedoch höher sein.

Diese Annahme wurde in einer Längsschnittstudie zu drei Messzeitpunkten überprüft. Die erste Erhebung fand zu Beginn der bildungswissenschaftlichen Vorbereitung des Praxissemesters statt, in der Inhalte zu Classroom Management besprochen wurden. Unmittelbar nach dieser Einheit fand die zweite Erhebung statt. In den Nachbereitungskursen, die am Ende des schulpraktischen Teils stattfanden, wurde zum dritten Erhebungszeitpunkt getestet. Das selbsteingeschätzte Wissen wurde mit dem SEWIK-Fragebogen von Thiel, Ophardt und Piwowar (2013) erhoben (8 Bereiche: z. B. Prozeduren, Konfliktlösung, Regeleinführung, Zeitmanagement). Um eine etwas objektivere Erfassung zu gewährleisten, wurden zusätzlich die Szenarien (3 Bereiche: Allgegenwärtigkeit, Strukturierung, Regeln) von Gold und Holodynski (2015) eingesetzt. Eine messwiederholte ANOVA zeigte eine signifikante Veränderung auf beiden Variablen über die Zeit. Im selbsteingeschätzten Wissen konnte in fast allen Bereichen ein Zuwachs nach der Theorieeinheit und nach dem schulpraktischen Teil festgestellt werden. Davon ausgenommen waren das Wissen über Prozeduren, Konfliktlösung und das Herstellen eines guten Arbeitsbündnisses. Interessanterweise berichteten die Studierenden bereits zum zweiten Erhebungszeitpunkt einen signifikanten Anstieg im selbsteingeschätzten Wissen. Die Übereinstimmung mit Expertenratings der Szenarios (freundlicherweise von Gold & Holodynski zur Verfügung gestellt) zeigte, dass die Studierenden ihr Wissen nur auf den Dimensionen Allgegenwärtigkeit und Strukturierung, aber nicht bei Nutzung und Etablierung von Regeln verbesserten. Die Theorieeinheit konnte hier nicht zu einem Wissenszuwachs führen. Dies zeigte sich vor allem darin, dass die Studierenden zum zweiten Erhebungszeitpunkt die niedrigste Übereinstimmung mit den Experten aufwiesen.

Diese Studie liefert Hinweise darauf, wie während des Studiums Wissen über Classroom Management entwickelt und erworben werden kann. Theorie-basierte Einheiten scheinen eher auf das selbsteingeschätzte Wissen wirken, und somit indirekt auf das Selbstvertrauen in die eigenen Fähigkeiten. Obwohl oftmals selbsteingeschätztes Wissen als anfällig durch soziale Erwünschtheit und anderen Verzerrungen angesehen wird (van de Mortel, 2008), kann die höhere Einschätzung und somit auch das größere Selbstvertrauen als positiv angesehen werden. Das Selbstvertrauen in die eigenen Fähigkeiten kann mit Selbstwirksamkeitserwartungen gleichgesetzt werden, welche positiv mit vielfältigen Bereichen des Unterrichts und Lehrens korreliert. So sind Lehrpersonen mit einer höheren Selbstwirksamkeit im Classroom Management weniger anfällig für Stress und Burnout (Aloe, Amo, & Shanahan, 2014). Der schulpraktische Teil scheint Studierende mit dem Wissen um effektive Classroom Management Strategien auszurüsten. Die Ergebnisse der Studie sprechen für die Relevanz von sowohl Theorie als Praxis während des Studiums, um Studierende auf den Schuldienst vorzubereiten.

## **STUDENT PERCEPTIONS OF CLASSROOM CLIMATE AND ITS RELATIONSHIP TO ACADEMIC OUTCOMES**

**Cansu Atlay<sup>1</sup>, Nicole Tieben<sup>1</sup>, Benjamin Fauth<sup>2</sup>**

<sup>1</sup>Universität Tübingen, Deutschland; <sup>2</sup>Hector Research Institute of Education Sciences and Psychology, Eberhard Karls University Tübingen; [cansu.atlay@lead.uni-tuebingen.de](mailto:cansu.atlay@lead.uni-tuebingen.de)

Prior research has shown that there is a strong association between students' background characteristics (gender, migration and socioeconomic status etc.) and academic achievement. It has also been shown that a positive classroom experience can promote student learning and academic achievement by providing a supportive and safe environment to students from diverse backgrounds. So far, most of the previous research studies examined the influence of student characteristics and classroom environment on student achievement. However, there is lack of research focusing on the impact of student characteristics on their perceptions of the classroom. Based on theoretical and empirical considerations, we argue that: (1) background characteristics are associated with students' perception of the classroom environment, (2) this association can be mediated by prior achievement, and (3) the relationship between student background and academic achievement can be moderated by students' perception of the classroom environment.

We use the longitudinal data from PISA-I-Plus, which is a part of the German national extension to PISA 2003 study (Prenzel and PISA-Konsortium 2006). As an extension international PISA study, PISA-I-Plus contains student answers from the original PISA 2003 study, but includes an additional national follow-up, which was conducted in 2004. The original international PISA sample included 220 schools and 25 students from each of these schools, both randomly drawn with a two-stage stratified sampling approach. For the purposes of the national PISA sample in 2003, which was a grade-based extension to the international study, at least ten grade 9 students were randomly drawn from the selected classrooms in 198 schools providing general education (PISA-9KL). The follow-up took place one year after the initial test, among those students in the PISA-9KL sample, who were finishing the 10th grade (PISA-I-Plus).

Using path analysis, we investigate students' perception of their classroom climate as a function of student background, the role of prior achievement as a potential mediator of this link and student perception as a potential moderator of the relation between student background and academic achievement. Initial results suggest that there socioeconomic status has a negative association with perceived support while migration status has a positive one. Compared to their male peers, female students perceive formal teaching methods to have lower quality. Preliminary analyses signal that prior achievement has a role in shaping these associations.

# E11

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus, Aula 33

## Qualität der Lernunterstützung in Kindergarten und (inklusive) Grundschule. Ergebnisse aus videogestützten Studien

Chair(s): **Elisabeth Moser Opitz** (Universität Zürich, Schweiz)

DiskutantIn(nen): **Marc Kleinknecht** (Leuphana Universität Lüneburg, Deutschland)

Lernen nachhaltig zu unterstützen, ist ein wichtiger Teil der professionellen Kompetenz von Lehrkräften und ist zentral für den Lernerfolg von Kindern (z.B. Decristan et al., 2015, Boonen et al., 2011). Professionelle Kompetenzen beinhalten kognitive und affektive Aspekte, z.B. Wissen in verschiedenen Dimensionen, Einstellungen zum Lernen, Motivation (Bloemke et al., 2015; Kunter et al., 2013).

Eine optimale Lernunterstützung muss bei den Lernenden emotionale und kognitive Aspekte berücksichtigen (Boyer et al., 2008). Pianta und Hamre (2009) nennen drei zentrale Bereiche: Die professionelle Beziehungsgestaltung, die Unterstützung von Lernprozessen sowie die Klassenführung.

Für die Unterstützung von Lernprozessen werden Scaffoldingstrategien als bedeutsam erachtet (z.B. Hmelo-Silver et al., 2007). Bezüglich der Beziehungsgestaltung spielt in Verbindung mit Faktoren wie Motivation, Selbstwirksamkeit und soziale Partizipation das Feedbackverhalten der Lehrkräfte eine zentrale Rolle (z.B. Bunning & Ellis, 2010, Hattie & Timberley, 2007). Als wichtige Aspekte der Klassenführung definieren Pianta und Hamre (2009) Verhaltensweisen der Lehrkräfte zur Steuerung des Verhaltens und der Aufmerksamkeit der Lernenden und der effektiven Nutzung der Unterrichtszeit.

Zielsetzung des Symposiums

Im Symposium werden Ergebnisse aus vier Videostudien zum Mathematik- und zum Sachunterricht vorgestellt, die sich mit verschiedenen Aspekten der Lernunterstützung (Unterstützung von Lernprozessen, Beziehungsgestaltung, Klassenführung) in den Bildungsbereichen Kindergarten und (inklusive) Grundschule befassen. Dabei interessiert zum einen die Qualität der Lernunterstützung und deren Auswirkung auf die Kinder, zum anderen der Einfluss von Faktoren wie Professionswissen, Beliefs, Ausbildung usw. auf die Lernunterstützung. Der Fokus liegt dabei auf dem Mathematik- und Sachunterricht bzw. auf dem naturwissenschaftlichen Lernen.

Beiträge

Im ersten Beitrag geht es um die Lernprozessunterstützung im Kindergarten beim naturwissenschaftlichen Lernen. Vorgestellt wird eine Längsschnittstudie, in der der Redeanteil von Kindern aus 32 Vorschulklassen, dessen Interaktion mit der verbalen Lernunterstützung der Lehrkräfte im Sinne der Scaffolding Means (van de Pol et al., 2010) in tutoriellen Lernsettings und dessen Auswirkungen auf den Lernerfolg untersucht werden.

Im zweiten Beitrag werden Videodaten aus dem Vorschulbereich analysiert (N = 145 frühpädagogische Fachkräfte aus der Schweiz und aus Deutschland), hier geht es um die Lernbegleitung bei mathemathikhaltigen Spielen. Untersucht werden sowohl Aspekte der Beziehungsgestaltung, der Lernunterstützung und der Klassenführung. Zum einen interessiert die Qualität der Lernbegleitung, zum anderen der Einfluss von verschiedenen Ausbildungskontexten.

Die Videodaten aus dem dritten Beitrag stammen aus inklusiven Grundschulklassen, in denen Kinder mit dem Förderschwerpunkt geistige Entwicklung unterrichtet werden. Hier steht die professionelle Beziehungsgestaltung im Vordergrund, operationalisiert durch die Variable „leistungs- und verhaltensbezogenes öffentliches Feedback“. Es wird untersucht, ob und wie sich dieses Feedback auf die soziale Akzeptanz bzw. die Ablehnung der Kinder in der Klasse auswirkt.

Im vierten Beitrag steht der Sachunterricht in der Grundschule im Zentrum des Interesses. In 65 Grundschulklassen wurde untersucht, welche Merkmale das lernunterstützende Handeln der Lehrpersonen aufweist und inwiefern Lehrpersonen sich in der Nutzung eher sozial-regulativer oder eher kognitiv anregender Unterstützungsmuster unterscheiden. Zudem wird der Zusammenhang des Unterstützungshandelns mit Aspekten des Professionswissens analysiert.

Durch die Untersuchung verschiedener Unterstützungsmomente, sowohl in den Studien als auch in deren Zusammenschau, kann das Symposium einen Beitrag leisten, emotional-motivationale und kognitive Aspekte von Lernunterstützung in Beziehung zu setzen und in Facetten ihrer Wirkung wie ihrer Bedingungen zu diskutieren.

*Beiträge des Symposiums*

### Lehrer-Schüler-Interaktion im frühen naturwissenschaftlichen Lernen: Der Redeanteil von Kindern in tutoriellen Settings und dessen Effekt auf den Lernerfolg

**Anika Bürgermeister<sup>1</sup>, Gerlind Grosse<sup>1</sup>, Ueli Studhalter<sup>2</sup>, Miriam Leuchter<sup>3</sup>, Vivien Diebl<sup>1</sup>, Henrik Saalbach<sup>1</sup>, Annette Tettenborn<sup>4</sup>**

<sup>1</sup>Universität Leipzig, Deutschland, <sup>2</sup>ETH Zürich, Schweiz, <sup>3</sup>Universität Koblenz-Landau, Deutschland, <sup>4</sup>PH Luzern

Theoretischer Hintergrund:

Die Lehrer-Schüler-Interaktion ist im deutschsprachigen Unterricht mehrheitlich durch die Lehrperson dominiert (Ackermann, 2011; Seidel, 2003). Konträr dazu wird jedoch zunehmend die Bedeutung einer aktiven Beteiligung der Lernenden an Unterrichtsgesprächen für deren Leistungs- und Motivationsentwicklung diskutiert (Schulz, 2011; Seidel et al. 2003). Nicht geklärt ist dabei aber die Wirkrichtung und damit einhergehend die Frage, ob eine verstärkte Beteiligung zu höherem Lernerfolg führt oder ob sich leistungsstärkere Lernende häufiger am Unterrichtsgeschehen beteiligen (Knobloch, 2013). Außerdem stellt sich die Frage nach der Wechselwirkung mit qualitativen Aspekten des Unterrichtsgesprächs. So gelten verbale Lernunterstützungsmaßnahmen der Lehrperson im Sinne eines Scaffoldings als effektive Methode, das Lernen der Schülerinnen und Schüler zu unterstützen (z. B. Leuchter & Saalbach, 2010; van de Pol et al., 2010). Zudem betrachten bisherige empirische Studien insbesondere das Klassengespräch als Ganzes und fokussieren selten die Lehrer-Schüler-Interaktion in tutoriellen Settings bzw. in der Lernbegleitung von Kleingruppenarbeit.

Davon ausgehend, untersucht die vorliegende Studie den Effekt der Redebeiträge von Kindern in frühen naturwissenschaftlichen Lehr-Lernsettings auf den Lernerfolg. Die Arbeit baut damit direkt auf einer Vorgängerstudie auf, die die Qualität der Redebeiträge von Lehrpersonen in tutoriellen Lernsettings und deren Wirkung auf Lernerfolg untersuchte. In der längsschnittlich angelegten Videostudie wurde dabei die verbale Lernunterstützung der Lehrperson erfasst und mit dem inhaltlichen Lernen der Kinder in Zusammenhang gebracht. Nicht erwartungskonform zeigten sich dabei geringe Effekte der

Scaffolding-Maßnahmen auf den Lernzuwachs. Direkt daran anknüpfend fokussiert der vorliegende Beitrag deshalb auf die quantitative Auswertung der Redeanteile der Kinder und untersucht die Bedeutung der aktiven Teilnahme an den Gesprächen für den Wissenserwerb.

Fragestellungen:

- 1) Welchen Einfluss hat der Redeanteil der Kinder auf deren Lernerfolg in tutoriellen Settings in Kleingruppen?
- 2) Wie beeinflussen sich Scaffolding-Maßnahmen der Lehrkraft und Redeanteile der Kinder gegenseitig?

Methode

Es handelt sich um eine längsschnittlich angelegte Videostudie, bei der in 32 Vorschulklassen eine standardisierte Lerneinheit zum Thema „Sinken und Schwimmen“ durchgeführt wurde. Die Lernleistung der Kinder wurde durch einen standardisierten Wissenstest zu drei Messzeitpunkten erfasst. Die Fachkräfte erhielten eine kurze Einführung in das Material der Lerneinheit.

Die verbale Lernunterstützung der Lehrperson (Scaffolding Means, van de Pol et al., 2010) wurde mittels theoriegeleitetem Kategoriensystem qualitativ und quantitativ erfasst und mit dem inhaltlichen Lernen der Kinder in Beziehung gesetzt. Der Redeanteil der Kinder wird pro Gruppe über Anzahl der Beiträge und Länge des Gesamtredanteils der Kinder absolut und relativ zum Redeanteil der Lehrperson erfasst. Haupteffekte und Interaktionen auf den Wissenszuwachs zu den Messzeitpunkten zwei und drei werden mittels Mehrebenenanalysen (HLM; Raudenbush & Bryk) analysiert.

Ergebnisse:

Entsprechend der Annahme, dass Lernende ein tieferes inhaltliches Verständnis erreichen, wenn sie sich verbal häufiger aktiv mit den Inhalten auseinandersetzen, zeigen erste Befunde positive Effekte von höheren kindlichen Redeanteilen auf die Leistung der Kinder. Es zeigen sich außerdem mehrere komplexe Wechselwirkungen zwischen der Qualität der verbalen Unterstützung durch die Fachkraft, den Redeanteilen der Fachkraft und den Redeanteilen der Kinder.

Die Befunde deuten darauf hin, dass es durchaus sinnvoll ist, die quantitative und qualitative Betrachtung eines Lehr-Lerngesprächs gemeinsam zu betrachten. Sie werden schließlich mit Blick auf die Gestaltung von Lehr-Lernprozessen im Kontext frühkindlicher Bildung diskutiert.

### **Lernunterstützung von mathematikhaltigen Spielen durch frühpädagogische Fachkräfte im Kindergarten**

**Anuschka Meier-Wyder<sup>1</sup>, Andrea Wullschleger<sup>2</sup>, Elisabeth Moser Opitz<sup>2</sup>, Anke Lindmeier<sup>3</sup>, Miriam Leuchter<sup>4</sup>**

<sup>1</sup>Hochschule für Heilpädagogik, Zürich, Schweiz, <sup>2</sup>Universität Zürich, Schweiz, <sup>3</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Kiel, Deutschland, <sup>4</sup>Universität Koblenz-Landau, Deutschland

Die pädagogische Fachkraft (PF) nimmt bei der Planung und Gestaltung von altersgerechten mathematischen Lernsituationen im Kindergarten eine zentrale Rolle ein (z.B. Schuler, 2013). Wichtig sind gemäß Pianta und Hamre (2009) in den Lernsituationen die professionelle Beziehungsgestaltung, die Unterstützung von Lernprozessen sowie die Klassenführung und Organisation. Nach Lindmeier (2011) sind fachspezifische Kognitionen (Fachwissen, fachdidaktisches Wissen) sowie reflexive und aktionsbezogene Kompetenzen bedeutsam. Reflexive Kompetenzen (RC) beinhalten z.B. Fähigkeiten zur Durchführung von Lernstandsdiagnosen oder Planung von Fördersituationen. Aktionsbezogene Kompetenzen (AC) beziehen sich auf die Durchführung von Fördersituationen bzw. die Lernunterstützung. Aspekte wie emotionale Wärme, Klassenmanagement oder fachsprachliche Anregungen können als Indikatoren für AC angesehen werden. Die Anwendung dieser Kompetenzen im beruflichen Alltag ist anspruchsvoll und es wird vermutet, dass die Ausbildung wesentlich zur Kompetenzentwicklung beiträgt (Hepberger et al., 2017). Da für PF verschiedene Ausbildungsformen (akademisch, nicht-akademisch) existieren, interessiert zudem, inwiefern dies zu charakteristischen Qualitätsunterschieden in der Lernunterstützung führt. Diskutiert wird darüber hinaus, welchen Einfluss die Berufserfahrung der PF hat (z.B. Phillips et al., 2009).

Fragestellungen

- In welcher Qualität werden Facetten der Lernunterstützung in den Komponenten AC und RC umgesetzt?
- Welche Unterschiede zeigen sich in der Qualität der Lernunterstützung in der Schweiz und in Deutschland in Abhängigkeit von Ausbildung und Berufserfahrung?

Methode

Die Stichprobe umfasst 145 PF (CH = 68, D = 77; 29% nicht-akademisch ausgebildet; Dienstjahre M = 12,85, SD = 14,13, Min = 1, Max = 41). In den Kindergärten wurden zwei ca. 15-minütige Fördersituationen mit mathematikbezogenen Spielen videografiert. Drei bis vier Kinder spielten das ‚Goldstückspiel‘ mit Unterstützung einer PF. Im Anschluss an die Spielsituation fand ein leitfadengestütztes Interview statt (Überlegungen zur Planung, Reflexion der Lernunterstützung, Auswahl der Kinder, Förderplanung). Die Einschätzung der Qualität der Lernbegleitung bezüglich Aspekten von RC und AC erfolgte mit einem hochinferenten Ratings mit jeweils vier Items (Skala 1 bis 4).

Items für AC: Emotionale Wärme; Klassenmanagement; adaptive Lernprozessanregung; fachsprachliche Anregung. Items für RC: diagnosegeleitete Planung der Spielsituation; Reflexion der Lernunterstützung; Planung der weiteren Förderung.

Mit Generalisierungskoeffizienten g-relativ zwischen 0.81 und 0.95 ist die Interrater-Reliabilität als gut einzuschätzen. Die Reliabilität erwies sich als befriedigend bis gut (AC und RC:  $\alpha = .79$ , AC:  $\alpha = .83$ , RC:  $\alpha = .66$ ).

Ergebnisse

Bezüglich der Qualität der Lernbegleitung zeigen sich bei den PF hohe Werte im Bereich der sozial-emotionalen Wärme (M = 3.07, SD = .53) und Klassenführung (M = 3.02, SD .55). Dagegen fallen die Werte in den Bereichen der adaptiven Lernprozessanregung (M = 1.73, SD = .98) sowie Planung der Spielsituation (M = 1.93, SD = .87) niedriger aus.

Die Schweizer Fachkräfte wiesen sowohl bei AC (M = 2.83, SD = .62) als auch bei RC (M = 2.72, SD = .65) höhere Werte auf als die Fachkräfte aus Deutschland (AC: M = 2.42, SD = .50; RC: M = 2.21, SD = .46).

Die Ergebnisse einer hierarchisch-sequentiellen Regressionsanalyse mit den abhängigen Variablen AC und RC und den Prädiktoren Ausbildung (akademisch/nicht akademisch), Land, Anzahl Dienstjahre und dem Interaktionsterm ‚Land x Ausbildung‘ ergaben für die abhängige Variable AC eine Varianzaufklärung von  $R^2 = .12$  mit einer signifikanten Veränderung von  $R^2$  durch die Aufnahme der Variablen Ausbildung und Land.

Für RC beträgt die Varianzaufklärung .17,  $R^2$  verändert sich signifikant durch die Aufnahme der Variablen Ausbildung und Land.

Diskussion



Unterschiede in der Qualität der Vor- und Nachbereitung von Lernsituationen sowie der Lernunterstützung lassen sich auf die Variablen Ausbildung und Land zurückführen. Die Interaktion der beiden Variablen erscheint jedoch nicht bedeutsam. Auch wenn der Einfluss der Lernunterstützung auf die Kinder mit diesen Daten nicht untersucht werden kann, werfen dies Ergebnisse Fragen auf bezüglich der Qualifikation der Fachkräfte.

### **Der Einfluss von Feedbackverhalten von Lehrkräften auf Aspekte der sozialen Integration in inklusiven Klassen: Ergebnisse einer Längsschnittstudie**

**Andrea Wullschleger<sup>1</sup>, Ariana Garrote<sup>2</sup>, Elisabeth Moser Opitz<sup>1</sup>, Susanne Schnepel<sup>1</sup>**  
<sup>1</sup>Universität Zürich, Schweiz, <sup>2</sup>Universität Autònoma de Barcelona, Spanien

Kompetenzen von Lehrkräften betreffen nicht nur Aspekte des Unterrichts, sondern auch die Gestaltung von Interaktionen. Eine häufige Form von Interaktion im Unterricht ist das Feedback der Lehrkräfte an die Schülerinnen und Schüler. Leistungsbezogenes Feedback kann zum einen eine Rückmeldung hinsichtlich der Korrektheit einer Antwort resp. einer Aufgabenlösung von Lernenden beinhalten, zum anderen Hinweise zum Arbeits- und Lernverhalten (Lipowsky, 2009). Lehrkräfte-Feedback zählt zu den wirkungsvollsten Faktoren für das Lernen von Schulkindern (Hattie & Timperley, 2011; Hattie, 2013). Sie haben zum Ziel, die Schülerinnen und Schüler in ihrem Lernen zu unterstützen (Bunning & Ellis, 2010). Rakoczy, et al. (2008) zeigten auf, dass positiv evaluierendes Feedback die intrinsische Motivation von Schülerinnen und Schülern zu steigern vermag.

Feedback von Lehrkräften kann aber auch die sozialen Prozesse im Klassenzimmer beeinflussen. Huber (2011) argumentiert unter Bezugnahme auf die Theorie sozialer Referenzierungsprozesse (Feinman, 1992), dass Schülerinnen und Schüler, die von der Lehrperson positives Feedback erhalten, sozial besser in die Klasse integriert sind, als Lernende, die negatives Feedback erfahren. In PC-gestützten Experimentalstudien (Huber, 2013; Huber, Gebhardt & Schwab, 2015) konnten diese Zusammenhänge bestätigt werden. Noch wenig untersucht ist, ob und wie sich diese Prozesse im realen Unterricht und bezogen auf die Entwicklung von verschiedenen Aspekten der sozialen Integration zeigen.

Forschungsfrage

Wie wirkt sich das Feedback von Lehrkräften zum Sozialverhalten und zur fachlichen Leistung der Schülerinnen und Schüler unter Einbezug der Variablen Geschlecht und Mathematikleistung auf die Entwicklung von verschiedenen Aspekten der sozialen Integration aus?

Methode

Untersucht wurde das Feedbackverhalten von 33 Lehrkräften in inklusiven Klassen (N = 546, Klasse 1–3; 49% Mädchen). Das Datenmaterial besteht aus Videoaufnahmen einer nicht standardisierten Mathematiklektion. Mithilfe eines Codiermanuals wurden Vorkommen, Empfänger, Bezug, Beurteilung, Form und affektive Tönung des Lehrkräfte-Feedbacks codiert. 25% der Daten wurden doppelcodiert. Dabei zeigte sich eine sehr gute Beurteilerübereinstimmung (Cohens Kappa 0.9).

Seitens der Kinder wurden Informationen zu den Freundschaftsbeziehungen (Nominationsmethode), zur sozialen Ablehnung (Ratingmethode) sowie zur Kooperation und Hilfsbereitschaft (Einschätzung der Mitschülerinnen und Mitschüler) aller Kinder zu Beginn und am Ende des Schuljahres erhoben. Zur Erfassung der Mathematikleistungen wurde ein eigenes Instrument eingesetzt (Cronbachs Alpha .86 bis .91).

Ergebnisse

Die Häufigkeit von Lehrkräfte-Feedback während einer Unterrichtslektion variiert stark (M = 92; SD = 27, Min. = 14, Max. = 162). Feedbacks bezüglich abweichendem Sozialverhalten korrelieren negativ mit Feedbacks auf eine korrekte fachliche Leistung ( $r = -.81, p < .01$ ). Feedbacks auf korrekte fachliche Leistung korrelieren negativ mit Feedbacks auf falsche fachliche Leistung ( $r = -.51, p < .01$ ).

Feedbacks zu abweichendem Sozialverhalten korrelieren negativ mit der Kooperation ( $r = -.25, p < .01$ ), der Hilfsbereitschaft ( $r = -.12, p < .01$ ) und der Anzahl Freundschaften ( $r = -.13, p < .01$ ) und positiv mit der Ablehnung ( $r = .18, p < .01$ ). Feedbacks zu korrekten fachlichen Leistungen korrelieren negativ mit Ablehnung ( $r = -.25, p < .01$ ) und positiv mit Kooperation ( $r = .15, p < .01$ ). Feedbacks zu falschen fachlichen Leistungen korrelieren positiv mit Kooperation ( $r = .13, p < .01$ ), Hilfsbereitschaft ( $r = .17, p < .01$ ) und Freundschaften ( $r = .19, p < .01$ ).

Erste Ergebnisse von Mehrebenenanalysen unter Berücksichtigung des Geschlechts und der Mathematikleistungen der Schülerinnen und Schüler zeigen einen signifikanten Einfluss von Feedback zu abweichendem Sozialverhalten auf die Veränderung des Kooperationsverhaltens der Kinder im Verlaufe eines Schuljahres. Feedback zur fachlichen Leistung ist hingegen kein Prädiktor für die Veränderung des Kooperationsverhaltens. Die Hilfsbereitschaft, die soziale Ablehnung und die Freundschaftsbeziehungen werden entgegen den Erwartungen nicht durch das Feedback der Lehrkräfte beeinflusst.

Fazit

Die Ergebnisse werfen Fragen auf, einerseits hinsichtlich der Wirkungskraft von Lehrerfeedback auf Aspekte der sozialen Integration in inklusiven Klassen bzw. auf deren Entwicklung, andererseits hinsichtlich der gewählten methodischen Herangehensweise.

### **Lernunterstützung zwischen sozial-regulativen und kognitiv anregenden Funktionen – Ergebnisse einer Videostudie im naturwissenschaftlichen Sachunterricht der Grundschule**

**Stefanie Schnebel<sup>1</sup>, Franziska Vogt<sup>2</sup>**  
<sup>1</sup>PH Weingarten, Deutschland, <sup>2</sup>PH St. Gallen, Schweiz

Theoretischer Hintergrund

Mit dem Ziel, naturwissenschaftliches Lernen zu fördern, wurde in der Studie „Innovation naturwissenschaftlich-technischer Bildung in der Primarstufe“ ein Lernarrangement entwickelt und in 65 Grundschulklassen eingesetzt. Ziel des Forschungsprojektes war es u.a. Erkenntnisse über die Qualität des unterstützenden Lehrpersonenhandelns und der dem Handeln zugrunde liegenden Facetten professionellen Wissens zu erlangen.

Dem unterstützenden Lehrpersonenhandeln wird im naturwissenschaftlichen Lernen in der Primarstufe eine große Bedeutung beigemessen (u.a. Hushman et al. 2015; Decristan et al. 2015). Kognitiv aktivierende und strukturierende Aktivitäten zeigen sich dabei als besonders relevant für das Lernen der Schülerinnen und Schüler (Hardy et al. 2006; Leuchter & Saalbach 2014). Untersuchungen zeigen allerdings auch die Bedeutung motivationaler und regulativ-organisationaler Aktivitäten, um die Basis für fachliches Lernen der Schülerinnen und Schüler zu schaffen (Boyer et al. 2008, Hmelo-Silver et al. 2007). Soll individuell

unterstützt werden, muss die Tiefenstruktur der angestrebten Lernprozesse beachtet werden, was die Relevanz fachdidaktischen Wissens verdeutlicht. Fachliches und fachdidaktisches Wissen wird als notwendige Voraussetzung für fachdidaktisch professionelles Lehrpersonenhandeln betrachtet (u.a. Kleickmann et al. 2010; Leuchter & Saalbach 2014). Die Zusammenhänge sind jedoch erst in Ansätzen empirisch überprüft.

#### Fragestellung

Der Beitrag geht folgenden Fragestellungen nach:

Welche Merkmale weist das lernunterstützende Handeln der Lehrpersonen auf?

Inwiefern lassen sich Lehrpersonen einem eher sozial-regulativen oder eher kognitiv anregenden Unterstützungsmuster zuordnen?

Welche Zusammenhänge bestehen zwischen Aspekten des Professionswissen und dem gezeigten Unterstützungshandeln der Lehrpersonen?

#### Design und Methode

In der internationalen, interdisziplinären Interventionsstudie wurde ein naturwissenschaftliches Lernarrangement entwickelt und über sechs Unterrichtsstunden in 75 dritten bzw. vierten Klassen im Schuljahr 2013/14 eingesetzt.

Das professionelle Wissen der Lehrpersonen (n=65) wurde zu mehreren Messzeitpunkten über Tests und standardisierte Befragung erfasst. Von Interesse waren dabei u.a. Beliefs zum Lehren und Lernen, die über die Skalen zur didaktischen Orientierung von Kleickmann (2008) erfasst wurden. Zusammenhänge zwischen Beliefs und Unterstützungshandeln wurden mittels von Korrelationsberechnungen ermittelt.

Das videografierte Unterstützungshandeln der Lehrkräfte wurde mit Hilfe eines Codierschemas mit 16 deduktiv und induktiv entwickelten Codes in sechs Facetten analysiert und quantitativ ausgewertet (xxx, 2016). Das Codiersystem bildet in den Facetten Kognitive Aktivierung und Strukturierung die Tiefenstruktur problemlösenden Lernens ab. Als weitere Facetten konnten Klassenmanagement, Evaluation, Transmission und Motivierung identifiziert werden. Mithilfe einer latenten Klassenanalyse wurden typische Unterstützungsmuster identifiziert.

#### Ergebnisse

Als ein Ergebnis der Studie konnten drei Muster von Lehrerhandeln identifiziert werden (LogL= -488.9; AIC = 1029.3;  $\chi^2 = .503$ ), deren Differenzierung insbesondere auf die Häufigkeit von kognitiv anregenden, evaluativen und sozial-regulativen Unterstützungsimpulsen zurückzuführen ist. Interessante Zusammenhänge zeigen sich zwischen den Facetten kognitive Aktivierung und Strukturierung ( $r = .369$ ,  $p < .01$ ) sowie kognitiver Aktivierung und Klassenmanagement ( $r = -.592$ ,  $p < .01$ ).

Zwischen professionellem Wissen und dem Unterstützungshandeln konnten teilweise Zusammenhänge identifiziert werden. Auf Ebene der Unterstützungsmuster zeigen sich keine Zusammenhänge. Auf Ebene einzelner Codes und Facetten zeigen sich einzelne Korrelationen. So hängt das fachliche Wissen der Lehrpersonen mit den Codes inhaltsbezogene Rückmeldung ( $r = .314$ ,  $p = .028$ ) und Transmission von Wissen ( $r = .303$ ;  $p = .034$ ) positiv sowie mit Organisation von Material ( $r = -.344$ ;  $p = .016$ ) negativ zusammen. Die fachdidaktischen Beliefs conceptual change ( $r = .377$ ;  $p = .009$ ) und Fehlkonzepte ( $r = .296$ ;  $p = .043$ ) hängen mit dem Unterstützungshandeln der Facette kognitive Aktivierung zusammen. Damit können einige erwartungskonforme Zusammenhänge zwischen Wissen bzw. Beliefs und Unterstützungshandeln identifiziert werden. Größtenteils zeigen sich die Dimensionen des Wissens und des Handelns als unabhängig voneinander. Diese Befunde widersprechen den Ergebnissen der Studie von Kleickmann et al. 2010, korrespondieren allerdings mit Befunden anderer Studien. Mögliche Erklärungen könnten sein, dass die Gestaltung des Lernarrangements einen starken Einfluss auf das Handeln der Lehrpersonen ausübt. Zudem ist das Handeln der Lehrpersonen ein multidimensional determiniertes Konstrukt, das nicht nur mit einem Voraussetzungsaspekt zusammenhängt.

## E12

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H119

### **Erfassung und Förderung der Beratungs- und Gesprächsführungskompetenz von (Lehramts-)studierenden**

Chair(s): **Jennifer Paetsch** (Otto-Friedrich-Universität Bamberg), **Daniela Sauer** (Otto-Friedrich-Universität Bamberg)

DiskutantIn(nen): **Martin Gartmeier** (Technische Universität München)

Der Berufsalltag von Lehrpersonen ist geprägt von einer Vielzahl an kommunikativen Interaktionsprozessen mit Schülerinnen und Schülern, Eltern oder auch Kolleginnen und Kollegen (Gartmeier, 2017; Hertzsch & Schneider, 2013). Obwohl die Beratungs- und Gesprächsführungskompetenz als wichtige Facette der professionellen Handlungskompetenz von Lehrkräften angesehen wird (z.B. Baumert & Kunter, 2013, KMK,2004; Wild, 2004), finden sich vor allem in der ersten Phase der Lehrerausbildung bislang nur wenig evidenzbasierte Lehrangebote, welche sich explizit deren Entwicklung und Förderung widmen (Gartmeier, 2017). Zu verschiedenen Aspekten der Beratungs- und Gesprächsführungskompetenzen von Lehrkräften liegen bereits Studien vor (z.B. Aich, 2011; Hertel, 2009; Klug, Bruder, Keller & Schmitz, 2012; Sauer, 2015). Unklar ist bislang, welche Kompetenzfacetten in der universitären Ausbildung durch die Studierenden erworben werden können und sollten. Weiterhin ist wenig darüber bekannt, welche Einstellungen und Überzeugungen Lehramtsstudierende in Bezug auf Beratung aufweisen. Das Symposium verknüpft Beiträge, die sich mit der Erfassung und Förderung von Beratungs- und Gesprächsführungskompetenzen von (Lehramts-)studierenden befassen.

Im ersten Beitrag stellen Horn et al. Ergebnisse aus einer Studie zum Beratungsverständnis angehender Lehrkräfte vor, in der Assoziationen mit Beratungshandlungen und -haltungen erfasst wurden. Der zweite Beitrag von Sauer et al. widmet sich der Förderung von Beratungskompetenzen von Lehramtsstudierenden. Vorgestellt wird ein für die universitäre Phase der Lehrerausbildung neu konzipiertes Peer-Beratungs-Training sowie das Evaluationskonzept zur Prüfung der Wirksamkeit. Im dritten Beitrag stellen Deistler et al. die Implementation eines umfangreichen Lehrbausteins für Lehramtsstudierende an der Universität Kassel vor. Ein wichtiger Bestandteil des Lehrbausteins ist dabei ein simuliertes Elterngespräch mit Schauspielenden, in dem die Studierenden sich in einem Beschwerdegespräch erproben können. Im vierten Beitrag von Weinhardt et al. liegt der Fokus auf der Erfassung von Beratungskompetenz. Vorgestellt werden Ergebnisse einer Videostudie, in der unterschiedlich erfahrene Studierende der Erziehungswissenschaft Beratungsgespräche mit ausgebildeten SimulationsklientInnen führen. Abschließend werden die Beiträge mit dem Fokus auf Beratungskompetenzen von angehenden Lehrkräften diskutiert.

#### *Beiträge des Symposiums*

### **Was verstehen Lehramtsstudierende unter Beratung? – Entwicklung eines Fragebogens zur Messung des Beratungsbegriffs von angehenden Lehrkräften**

**Dorothea Horn, Barbara Drechsel, Daniela Sauer**  
Otto-Friedrich-Universität Bamberg

Aufgrund der wachsenden Vielfalt an Beratungsanlässen im Kontext Schule (Hoffmann, 2016), u.a. durch gesellschaftliche Herausforderungen wie die Zunahme der Heterogenität der Schülerschaft durch Migration und der damit verbundenen Forderung nach Bildungs- und Erziehungspartnerschaften zwischen Elternhaus und Schule (Kultusministerkonferenz [KMK], 2013; Sacher, 2014) gewinnen Beratungskompetenzen von Lehrkräften an Relevanz (Schnebel, 2012). Neben spezifischen Dimensionen der Beratungskompetenz von Lehrpersonen (z.B. Bruder, 2011; Hertel, 2009) zählen Überzeugungen und Werthaltungen sowie motivationale Orientierungen zu den zentralen Aspekten der professionellen Handlungskompetenz von Lehrkräften (Baumert & Kunter, 2006; Krapp & Hascher, 2009). Während zu verschiedenen Aspekten der Beratungskompetenzen von Lehrpersonen (z.B. Hertel, 2009; Klug, Bruder, Keller & Schmitz, 2012) bereits Erkenntnisse vorliegen, sind Studien, welche sich der Beschaffenheit von beratungsbezogenen Überzeugungen sowie deren Bedeutung für das Beratungshandeln im schulischen Kontext widmen, rar (Fives & Buehl, 2012). Da zudem bislang ein quantitatives Instrument zur Erfassung von Überzeugungen zum Thema Beratung fehlt, setzt der vorliegende Beitrag an dieser Forschungslücke an und geht der Frage nach, wie das Beratungsverständnis von Lehrkräften und Lehramtsstudierenden erfasst werden kann und welche typischen Profile sich differenzieren lassen. Hierzu wurde ein Fragebogen entwickelt, der das Beratungsverständnis angehender Lehrkräfte über Assoziationen mit Beratungshandlungen und -haltungen erfasst. Die Konstruktion des Fragebogens basiert auf Ähnlichkeitseinschätzungen semantischer Bedeutungsrepräsentationen, die bereits erfolgreich zur Messung von Lernüberzeugungen eingesetzt wurden (Bauer et al., 2010; Thoma & Prenzel, 2009; Drechsel, 2001). Vorgegeben werden Begriffsstimuli, die im Zusammenhang mit dem Begriff Beratung typische Handlungen abbilden und hinsichtlich der Ähnlichkeit mit den eigenen subjektiven Vorstellungen einzuschätzen sind. Die 25 Items dieser Ähnlichkeits-Skala Beratungsbegriffe gehen auf die Ergebnisse einer qualitativ-rekonstruktiven Interviewstudie mit Lehrkräften zu ihrem Beratungsverständnis (Sauer, 2015) zurück, in der verschiedene homologe und kontrastierende handlungsleitende Orientierungen von Lehrkräften bei der Beratung von Eltern rekonstruiert und die sich darin abbildenden Beratungsverständnisse typisiert werden konnten (ebd.). Die entwickelte Ähnlichkeits-Skala Beratungsbegriffe wurde mit Lehramtsstudierenden pilotiert, hinsichtlich der Skalengüte analysiert, interpretiert und modifiziert und schließlich von Lehramtsstudierenden zu Beginn des erziehungswissenschaftlichen Studiums (NTypenanalyse = 266) bearbeitet. Latente Profilanalysen identifizierten drei positiv geprägte Überzeugungsprofile, von denen zwei Profile auf einen Novizen-Beratungsbegriff (nProfil 1 und 3 = 187; 70,3 %) und das dritte Profil auf einen ausdifferenzierteren Beratungsbegriff (nProfil 2 = 79; 29,7 %) zu Beginn des Studiums hinweisen (Horn, 2017). Analog zu den Befunden von Sauer (2015), die mit praktizierenden Lehrkräften gearbeitet hatte, verweisen die ersten Ergebnisse zum Beratungsbegriff von Lehramtsstudierenden auf Vorstellungen von Beratung, die eher einem Alltagsverständnis von Beratung (Grewe, 2005) entsprechen und bspw. Beratung als direktiven Ratschlag verstehen (Sauer, 2015) und sich damit deutlich vom professionstheoretischen Diskurs zur Beratungsaufgabe von Lehrkräften (z.B. Hertel, 2009; Mutzeck, 2008; Schwarzer & Posse, 2005; Strasser & Gruber, 2003) unterscheiden.

Allerdings wurde aus den Analysen auch deutlich, dass bestimmten Facetten des Beratungsbegriffs mit dem Instrument nicht erfasst werden konnten. Um subjektive Beratungsbegriffe auf allen Expertisestufen (Studierende bis Beratungsexperten mit langjähriger Berufserfahrung) noch differenzierter erfassen zu können, wurde das Instrument erneut überarbeitet und in einer Studierendenbefragung eingesetzt (N = 350), deren Ergebnisse Teil des Beitrags sein werden. Außerdem werden der Prozess der Skalenkonstruktion sowie die jeweiligen Analysen berichtet und die Ergebnisse hinsichtlich ihrer Bedeutung für die Professionalisierung der subjektiven Begriffe (beispielsweise durch Rollenreflexion in der Ausbildung von Lehrkräften) sowie

möglicher Implikationen für die Gestaltung adäquater Lernbedingungen und Reflexionsgelegenheiten in der Lehrerbildung diskutiert.

## **Evidenzbasierte Förderung der Beratungskompetenz von Lehramtsstudierenden. Das Bamberger Peer-Beratungstraining**

**Daniela Sauer, Jennifer Paetsch, Barbara Drechsel, Jörg Wolstein  
Otto-Friedrich-Universität Bamberg**

Die Beratung von Schülerinnen und Schülern, deren Eltern sowie von Kolleginnen und Kollegen zählt seit den 1970er Jahren zu den ausgewiesenen Aufgabenfeldern von Lehrkräften (Baumert & Kunter, 2013; Deutscher Bildungsrat, 1972; Hertel, 2017; KMK 2004; Kunina-Habenicht et al., 2012). Besonders durch die zunehmende Relevanz der Erziehungs- und Bildungspartnerschaft zwischen Elternhaus und Schule (Ostermann, 2016; Sacher, 2014) gewinnt die Beratungstätigkeit von Lehrpersonen in den vergangenen Jahren an Bedeutung. Vorliegende empirische Befunde (z.B. Aich & Behr, 2015; Bruder, 2011; Bülow 2016; Gerich et al., 2015; Hertel, 2009; Weinhardt, 2014) verweisen darauf, dass Beratungskompetenzen sowohl von Lehrpersonen in der dritten Phase als auch von Studierenden erlernbar sind bzw. erfolgreich gefördert werden können. Trotz der Anerkennung der Relevanz des Themas und einer Anzahl von Studien dazu ist der Aufbau und die Förderung von Beratungskompetenzen im Bereich der universitären Lehrkräfteausbildung bislang noch wenig verankert (Gerich, 2016; Strasser & Gruber, 2003; Widulle, 2016). Für Lehramtsstudierende in der ersten Phase der Ausbildung ist es oft schwierig, die Relevanz des Themas Beratung zu erkennen, da sie selbst in dieser Phase noch deutlich auf die eigene Schulerfahrung rekurrieren und Beratung im schulischen Kontext häufig nicht als Schwerpunkt der Arbeit einer Lehrperson wahrgenommen haben. Eine Möglichkeit, Beratungskompetenzen mit hoher Authentizität zu vermitteln, sind Peer-to-Peer Ansätze. Während diese im Bereich der Prävention und Gesundheitsförderung schon länger verbreitet sind (Mellanby, Rees & Tripp, 2000), gibt es im Bereich der Hochschullehre noch wenig Erfahrung dazu (Benz & Kunz, 2007). Von den verschiedenen Formaten dieser Ansätze erscheint ein Training auf Basis der Peer-Education am geeignetsten (Kempen, 2007). Hier arbeiten einzelne Peers mit der Zielgruppe, der sie zum Beispiel bezüglich des Alters oder / und in der institutionellen Zugehörigkeit gleichen. Deshalb wurde an der Universität Bamberg ein entsprechendes Lehr-Lernformat entwickelt, bei dem erfahrene Studierende Studienanfänger in den Lehramtsstudiengängen basale Beratungskompetenzen vermitteln (Spätgens, Wolstein & Drechsel, submitted). Zusätzlich soll ein Bewusstsein für diese Facette des professionellen Handelns geschaffen und die Motivation und Offenheit für künftige Lernprozesse in diesem Bereich gefördert werden. Ein weiterer Vorteil des Peer-Trainings ist, dass so auch in großen Kohorten, wie sie in der bildungswissenschaftlichen Lehramtsausbildung häufig sind, theoriebezogenes und praxis- bzw. trainingsbasiertes Lernen stattfinden kann.

Ziel des Beitrags ist die Vorstellung des Seminarkonzepts und Befunde aus der formativen und summativen Evaluation des Trainings. Dabei wird folgenden Forschungsfragen nachgegangen:

-Sehen Lehramtsstudierende das Thema Beratung als wichtig und persönlich relevant für ihren Beruf an?

-Ist der Peer-Ansatz geeignet, um Lehramtsstudierenden grundlegende Kompetenzen zum Thema Beratung zu vermitteln?

-Wie kann man die Veränderung von Wissen, Motivation, Beliefs und Handlungskompetenzen sowohl bei Peer-Edukatoren als auch bei Peers messen?

Neben den theoretisch-konzeptionellen Darstellungen des Seminardesigns präsentiert der Beitrag das Evaluationskonzept und die dafür neu entwickelten oder adaptierten Messinstrumente (z.B. nach Hertel, 2009; Bruder, 2011), beispielsweise ein Wissenstest, Fallvignetten, Fragebögen zur Erfassung selbsteingeschätzter Kompetenz und beratungsbezogener Überzeugungen sowie simulierte Beratungsgespräche zur Erfassung handlungsbezogener Beratungskompetenzen der Studierenden. Die Befunde der Pilotierungsstudien weisen darauf hin, dass es gelingt, Beratung als professionelles Handlungsfeld für Studierende zu erschließen, dass das Peer-Education-Format eine vielversprechende Methode zur Vermittlung von Beratungskompetenzen ist, und dass die für die Evaluation vorgesehenen Messinstrumente dazu geeignet sind, die genannten Indikatoren zu erfassen.

Das Seminarkonzept und die Befunde werden mit Blick auf die Weiterentwicklung des Lehr-Lernformats und auch auf ihre Übertragbarkeit auf weitere Aspekte des Beratungshandelns im schulischen Kontext, diskutiert.

## **Wie kann die Elterngesprächsführungskompetenz von Lehramtsstudierenden wirksam trainiert werden? Das Projekt „ProfKom-L“**

**Anja Deistler<sup>1</sup>, Natalie Fischer<sup>1</sup>, Martin Gartmeier<sup>2</sup>  
<sup>1</sup>Universität Kassel, <sup>2</sup>TU München**

Theoretischer Hintergrund

Das Führen von Elterngesprächen gewinnt in der schulischen Praxis zunehmend an Bedeutung (u.a. Hertel, 2009). Gute Elterngespräche können nicht nur positiven Einfluss auf Schulleistungen, Lernklima sowie Wohlbefinden in der Schule für die Schüler\_innen haben, auch Eltern und Lehrpersonen profitieren von einem guten Lehrpersonen-Eltern-Kontakt (zsf. Sacher, 2012). Lehrkräfte unterscheiden sich allerdings hinsichtlich ihrer Gesprächsführungskompetenz. So kann u.a. davon ausgegangen werden, dass Persönlichkeitsmerkmale der Lehrkraft Einfluss auf die Gestaltung der Elterngespräche nehmen (Sauer, 2015; Skaalvik & Skaalvik, 2012). Auch Selbstwirksamkeitserwartungen und Überzeugungen gelten als wirksam bezüglich des Lehrerhandelns (zsf. Baumert & Kunter, 2013; Lipowsky, 2006). Diese könnten bereits in der Lehramtsausbildung gefördert werden.

Gleichwohl sind universitäre Lehrangebote, die Gesprächsführungskompetenz im Lehramtsstudium gezielt fördern, selten (Gartmeier, Bauer, Noll & Prenzel, 2012). Bisher entstandene Ansätze wurden in kontrollierten Interventionsstudien überprüft (Aich, 2011; Hertel, 2009; Gartmeier, Bauer, Fischer, Karsten & Prenzel, 2011). Eine Implementation in die Ausbildungspraxis steht aber bisher weitgehend aus. So entstand auf Grundlage des Münchener Modells zur Gesprächsführungskompetenz in Lehrer-Elterngesprächssituationen (MMG-E) (Gartmeier et al., 2011) im Projekt „ProfKom-L“ (Professionelle Kommunikation – Lehrbaustein) (gefördert von der Vodafone-Stiftung) ein umfangreicher Lehrbaustein für Lehramtsstudierende (Gartmeier et al., 2015), der verschiedene Verhaltensweisen von Lehrkräften für Elterngespräche gezielt thematisiert. Ein Bestandteil dieses Lehrbausteins ist auch ein simuliertes Elterngespräch mit Schauspielenden, in dem die Studierenden ein Beschwerdegespräch erproben. Diese praktische Einheit trägt u.a. der Forderung nach Erprobungs- und Anwendungsphasen in Weiterbildungen Rechnung (Lipowsky & Rzejak, 2012).

Ziele der vorgestellten Studie sind die Implementation des ProfKom-Bausteins in eine universitäre Lehrveranstaltung in einem Basismodul der Lehramtsausbildung an der Universität Kassel und die Evaluation seiner Wirksamkeit. Betrachtet werden

neben persönlichen Merkmalen der Studierenden auch Überzeugungen und Fähigkeitsselbsteinschätzungen hinsichtlich (zukünftiger) Elterngespräche. Darüber hinaus werden die videographierten, simulierten Elterngespräche auf ihre Qualität untersucht.

#### Methode

Der Studie liegt ein Prä-Post-Kontrollgruppen-Design zugrunde. Es wurden 81 Lehramtsstudierende der Universität Kassel im Sommersemester 2016 und 2017 jeweils längsschnittlich mittels Online-Fragebogen u.a. zu ihrem Hintergrund, ihrer Persönlichkeit sowie ihren Überzeugungen und Fähigkeitsselbsteinschätzungen hinsichtlich der Gesprächsführung befragt. Die Untersuchungsgruppe (n=41) nahm zwischen den Befragungen am Lehrbaustein teil. Die Teilnahme der Wartekontrollgruppe (n=40) am Lehrbaustein erfolgte erst nach Abschluss der Befragungen.

Zudem nahmen die Studierenden beider Gruppen an simulierten Elterngespräch mit Schauspielenden zur Messung der Performanz des Gelernten jeweils zwischen den Befragungen teil (nUG=47; nKG=46). Neben den Fragebogendaten aus den Sommersemestern 2016 und 2017 liegen derzeit mittels Manual hochinferent-ausgewertete Videodaten von Studierenden aus dem Sommersemester 2016 vor.

Alle Analysen erfolgen mittels Varianzanalysen mit Messwiederholungen sowie längsschnittlichen Regressionen in SPSS 24.

#### Ergebnisse

Erste Analysen der Fragebogendaten zeigen auf deskriptiver Ebene leichte Veränderungen der Selbsteinschätzung der Fähigkeiten zur Gesprächsführung, der Selbstwirksamkeitserwartungen sowie der Überzeugungen der Lehramtsstudierenden unabhängig von der Gruppenzugehörigkeit. Zudem erweisen sich persönliche Merkmale der Studierenden für Einschätzungen zum zweiten Messzeitpunkt als bedeutsam, wie zum Beispiel die Abiturnote in Deutsch für die Selbsteinschätzung der Ausdrucksfähigkeit in Gesprächen ( $R^2 = .123$ ,  $F(1,70)=10.824$ ,  $p=.002$ ) oder die Ausprägung in der Extraversion auf die Fähigkeitseinschätzung zur Gesprächsführung allgemein ( $R^2 = .157$ ,  $F(1,77)=15.561$ ,  $p=.000$ ).

Erste Analysen der Videos aus dem Sommersemester 2016 zeigen auf deskriptiver Ebene eine höhere Qualität der Gespräche zugunsten der Untersuchungsgruppe über alle erfassten Facetten des MMG-E – Beziehungsgestaltung, Problemlösung und Gesprächsstrukturierung – hinweg, erste Analysen deuten auf statistisch bedeutsame Unterschiede hin.

Somit gibt es Hinweise auf die Wirksamkeit des Lehrbausteins von ProfKom-L und die erfolgreiche Implementation in die Praxis der Lehramtsausbildung an der Universität Kassel.

### **Messung von Beratungskompetenz**

**Marc Weinhardt<sup>1</sup>, Augustin Kelava<sup>2</sup>**

<sup>1</sup>Evangelische Hochschule Darmstadt, School of Professional Studies, <sup>2</sup>Universität Tübingen, Hector-Institut für empirische Bildungsforschung

Der Beitrag stellt die Erfassung von Beratungskompetenz im Rahmen einer Videostudie vor, in der unterschiedlich erfahrene Studierende der Erziehungswissenschaft Beratungsgespräche mit ausgebildeten Simulationsklient\*innen führen. Die durch Ratingskalen erfassten Kompetenzen werden durch Strukturgleichungsmodelle beschrieben (Weinhardt und Kelava, 2016) und geben Auskunft über Struktur und zeitliche Entwicklung des Beratungskompetenzerwerbs in berufsbiographisch frühen Stadien, wie sie prototypisch in der Hochschulbildung zu finden sind.

## E13

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H118

### Diagnosekompetenzen von Lehrpersonen als zentrale Komponente professionellen Handelns

Chair(s): **Jens Möller** (Universität Kiel), **Stefan Daniel Keller** (PH FHNW, Schweiz)

DiskutantIn(nen): **Anna-Katharina Praetorius** (Universität Zürich)

In der Diskussion über Qualität in der Lehrerbildung und ist kaum eine Komponente des professionellen Handelns so prominent diskutiert worden wie die diagnostische Kompetenz (DK; Südkamp et al., 2014; Baumert und Kunter, 2006). Studierende sollten am Ende der Ausbildung in der Lage sein, so lautet die Empfehlung in Deutschland, „fachbezogene Formen der Lern diagnose und Lernförderung zu kennen, zu beurteilen und einzusetzen“ (KMK, 2004). DK ist damit Voraussetzung einer objektiven und fairen Leistungsbewertung und zugleich unerlässlich für einen Unterricht, welcher sich adaptiv an den Kompetenzentwicklungen der Lernenden orientiert und diese in Planungs- und Evaluationsprozesse aktiv miteinbezieht (Spinath, 2005; McElvany, Schroeder et al., 2009).

Das Ziel dieses Symposiums ist es, diesen zentralen Aspekt von Lehrerprofessionalität empirisch zu fokussieren und dabei unterschiedliche Handlungs-, Fach- und Forschungsperspektiven transdisziplinär zu integrieren. Die darin vorgestellten Arbeiten beschäftigen sich alle aus verschiedenen Blickwinkeln mit dem „Einschätzen“ von fachlich geprägten Situationen und Leistungen durch Lehrpersonen. Dabei rücken sowohl Unterrichtsvorbereitung, -durchführung wie auch Leistungsbeurteilung ins Blickfeld.

Für das Fach Mathematik stellen Jeschke, Lindmeier und Heinze eine Studie zur aktionsbezogenen Kompetenz von Lehrkräften vor. Diese ist definiert als die Fähigkeit, professionelles Wissen unter Zeitdruck auf die typischen Anforderungen während des Unterrichtens anzuwenden, etwa um Lernenden eine passende Rückmeldung zu geben. In einer videobasierten Fallstudie wird nachgewiesen, dass die meisten Lehrpersonen weniger Schwierigkeiten bei der Identifizierung mathematischer Probleme, jedoch häufig Defizite bei der Formulierung passender Rückmeldungen für die Lernenden aufweisen.

Für das Fach Biologie fragen Fischer, Walter, Möller und Harms, ob Lehramtsstudierende in der Lage sind, zwischen verschiedenen evolutionsbiologischen Fehlvorstellungen von Lernenden zu differenzieren. Mit Hilfe des computerbasierten Messinstruments „student inventory“ werden Erkenntnisse über den Zusammenhang von deklarativem Wissen (fachdidaktisch, fachwissenschaftlich) und prozeduralem Wissen bei der DK der Biologie-Lehrkräfte gewonnen.

Für das Fach Geschichte legen Waldis, Marti, Nitsche und Gollin eine Untersuchung zu Aufgabenformaten, Beurteilungskriterien und Feedbackstrategien von Geschichtslehrpersonen im gymnasialen Unterricht vor. Dabei werden die diagnostischen Kompetenzen deskriptiv erfasst und zudem erforscht, ob die Beteiligten im Rahmen einer gezielten Schreibintervention eine an Kompetenzdimensionen angelehnte Diagnostik zu entwickeln beginnen.

Für das Fach Englisch schliesslich stellen Keller und Möller eine Studie zur DK von Lehrpersonen bei authentischen Englisch-Essays auf der gymnasialen Oberstufe vor. Darin werden Einflussfaktoren auf Text- und Schülerebene auf die Beurteilungsqualität der Aufsätze gemessen und Implikationen für die Lehrerbildung diskutiert.

Die Beiträge im Symposium werden diskutiert von Anna-Katharina Praetorius, welche selber sowohl zur Unterrichtsqualität wie auch zur diagnostischen Kompetenz von Lehrkräften intensiv geforscht und publiziert hat.

*Beiträge des Symposiums*

### Was hindert Mathematiklehrkräfte daran kompetent zu handeln?

**Colin Jeschke, Anke Lindmeier, Aiso Heinze**

IPN Kiel

Theoretischer Hintergrund und Fragestellung

Zur Verbesserung der Lehramtsausbildung untersuchen aktuelle Ansätze der Professionsforschung zunehmend handlungsnahe Komponenten professioneller Kompetenz (Blömeke, Gustafsson & Shavelson, 2015). So wählte Lindmeier (2011) bezogen auf Fachunterricht einen handlungsnahen Zugang und definiert aktionsbezogene Kompetenz (AC) von Lehrkräften als die Fähigkeit, professionelles Wissen unmittelbar und unter Zeitdruck auf die typischen Anforderungen während des Unterrichtens anzuwenden. Somit beinhaltet AC nicht nur die Fähigkeit beispielsweise eine Schüleräußerung unter Zeitdruck zu diagnostizieren, sondern auch daraus spontan eine passende Handlung – z. B. eine schülergerechte Erklärung – abzuleiten (Lindmeier, Heinze & Reiss, 2012). Dieses Kompetenzkonstrukt wurde für das Fach Mathematik bereits erfolgreich operationalisiert und untersucht (u. a. Knievel, Lindmeier & Heinze, 2015).

Aufbauend hierauf werden in der vorliegenden Studie Voraussetzungen, die zur Bewältigung solcher komplexen instruktionalen Anforderungen benötigt werden, untersucht. Aus theoretischer Perspektive gelten dabei sowohl Fachwissen (CK) und fachdidaktisches Wissen (PCK) (vgl. Knievel et al., 2015) als auch die allgemeine Situationswahrnehmung (Situation Awareness; Miller, 2011) als notwendige Voraussetzungen für die Beurteilung fachspezifischer Unterrichtssituationen und somit für professionelles Handeln. Weitere Einflussgrößen für AC sind bislang unklar.

Vor diesem Hintergrund wird in der vorliegenden Studie folgende Forschungsfrage adressiert: Welche weiteren Einflussfaktoren für AC lassen sich zusätzlich zum fachspezifischen Wissen und Situationswahrnehmung empirisch identifizieren?

Methode

Die vorliegende Fallstudie basiert auf den Daten von 74 angehenden Mathematiklehrkräften (41 Lehramtsstudierenden, Universität Kiel; 33 Lehrkräfte im Vorbereitungsdienst aus Nord- und Süddeutschland). Als Indikator für mathematikspezifisches AC dienen 8 videobasierte Items (Antwortformat: direkte Sprachaufnahme als Lehrerhandlung; Kodierung: Fleiss  $\kappa = .65-.89$ ). Unterschiede in der Situationswahrnehmung wurden durch 8 neuentwickelte video-basierte Items mit Fokus auf Klassenführungsproblemen (z.B. problematisches Schülerverhalten) erfasst. Für mathematikspezifisches CK und PCK wurden bereits erprobte Instrumente verwendet (Loch, Lindmeier & Heinze, 2013).

Mit Hilfe dieser Instrumente wurden diejenigen angehenden Lehrkräfte identifiziert, die trotz guter Wissensbasis und Situationswahrnehmung nur schwache Leistung in AC erreichen können, um Hinweise auf weitere Einflussfaktoren zu erlangen. Entsprechend wurden die Fälle ausgewählt, die (1) überdurchschnittliche Leistungen in sowohl dem fachspezifischen Wissen (CK und PCK) als auch der Situationswahrnehmung aufweisen und zusätzlich (2) unterdurchschnittliche Leistungen in

AC zeigten. Diese Auswahlkriterien erfüllten N = 7 Fälle (4 weiblich; 4 Studierende; Durchschnittsalter 27,0 Jahre). Diese Fälle lieferten für AC insgesamt 50 Aufgabebearbeitungen, wobei 6 erfolgreiches Lehrerhandeln abbilden und deswegen nicht in der Analyse berücksichtigt wurden. Entsprechend wurden 44 Aufgabebearbeitungen einer tiefergehenden Inhaltsanalyse unterzogen.

#### Ergebnisse

Von den 44 Antworten enthielten n = 21 Angaben keine Hinweise auf korrekte Situationsdiagnose (z. B. Antworten, die nicht auf die Aufgabe eingehen; fachliche Fehleinschätzung) oder keine zur Situation passende Handlung und wurden für die weitere Analysen verworfen. Die übrigen n = 26 Handlungen zeigten dagegen eine korrekte Situationsdiagnose und Hinweise auf einen passenden Lösungsansatz, jedoch erhebliche Defizite bei Ableitung passender Handlungen. Diese Defizite lassen sich anhand von zwei Kategorien charakterisieren: 1) Defizite in der Formulierung passender Rückmeldungen (n = 21), z. B. Hilfestellung nach diagnostizierter Schwierigkeit bei einer Mathematikaufgabe. 2) Defizite bei der Präzision der gegebenen Rückmeldung (n = 5), u. a. durch Hinzunahme irrelevanter Zusätze. Die Ergebnisse zeigen, dass die Formulierung passender Rückmeldungen das am häufigsten auftretende Defizit darstellt. Insgesamt trat dieses Defizit bei drei der sieben betrachteten Fälle dominant auf, während die weiteren vier Fälle zusätzlich Defizite bei Situationsdiagnose und fachlichen Einschätzung zeigten.

#### Diskussion

Passend zu den Fallauswahlkriterien zeigten die meisten der betrachteten Handlungen weniger Probleme bei der Problemidentifizierung oder Problemlösung, jedoch häufig Defizite bei der Formulierung passender Rückmeldungen für die Lernenden. Es kann somit davon ausgegangen werden, dass die betrachteten angehenden Lehrkräfte zwar die nötige Fähigkeiten zur Beurteilung der Problemsituation besaßen, jedoch nicht die nötige Fähigkeit, das Wissen adäquat und fallbezogen anzuwenden. Diese Hinweise bestätigen aktuelle Kompetenzmodellierungen, in denen neben der Situationswahrnehmung und -diagnose weiterführende situationspezifische Fähigkeiten zur Wissensanwendung notwendig sind.

### Diagnostisches Wissen von Lehramtsstudierenden im Fach Biologie messen

**Julian Fischer<sup>1</sup>, Julia Walter<sup>2</sup>, Jens Möller<sup>2</sup>, Ute Harms<sup>1</sup>**  
<sup>1</sup>IPN Kiel, <sup>2</sup>Universität Kiel

#### Theoretischer Hintergrund

Zentrales Ziel des Biologieunterrichts ist es, dass Schülerinnen und Schüler konzeptuelles biologisches Wissen erwerben. Entsprechende Lernprozesse werden maßgeblich durch Fehlvorstellungen der Lernenden, unvollständige oder falsche Erklärungen wissenschaftlicher Konzepte, behindert, die aufgrund von Erfahrungen im Alltag bzw. im vorangegangenen Unterricht entwickelt wurden (sog. misconceptions; u. a. Smith, Disessa, & Roschelle, 1994; Vosniadou & Skopeliti, 2013). Bleiben diese Fehlvorstellungen im Unterricht unerkannt, verhindern sie Konzeptwechselprozesse und damit das Erlernen des wissenschaftlich korrekten Konzepts (Bransford et al., 2000). Lehrkräfte benötigen diagnostisches Wissen, um Fehlvorstellungen ihrer Schülerinnen und Schüler zu identifizieren und bestmögliche Bedingungen zum Erreichen der Lernziele zu schaffen (Rogalla & Vogt, 2008). Voraussetzung hierfür ist, dass das im Lehramtsstudium erworbene deklarative diagnostische Wissen (z. B. zu wissen, dass es zum Thema Evolution spezifische Fehlvorstellungen gibt) nicht träge bleibt, sondern in prozedurales Wissen umgesetzt wird (vgl. Blömeke, Kaiser, & Lehmann, 2008; Renkl, 1996; z. B. Identifizierung einer Fehlvorstellung in einer Schülerantwort). Diagnostisches Wissen wird hier nach Südkamp et al. (2012) als Urteilsgenauigkeit bezüglich des vorliegenden Fehlkonzepts verstanden. Während das deklarative Professionswissen von Lehrkräften in den letzten Jahren intensiv beforscht wurde, stellt die Analyse handlungsnahen prozeduralen Wissens ein Forschungsdesiderat dar. Erst die Erfassung von deklarativem und prozeduralem Wissen ergibt ein vollständiges Bild der Entwicklung professioneller Kompetenzen im Lehramtsstudium. Im vorliegenden Beitrag wurden mit Hilfe des Schülerinventars deklaratives und prozedurales diagnostisches Wissen im biologiedidaktischen und pädagogisch-psychologischen Wissensbereich sowie deren Zusammenhänge erhoben. Als fachlicher Inhalt der Studie wurde die biologische Evolution (Industriemelanismus des Birkenspanners) gewählt, da hierzu bereits eine Vielzahl von Schülerfehlvorstellungen beschrieben wurde (vgl. Gregory, 2009).

#### Forschungsfragen:

Sind die Lehramtsstudierenden in der Lage, Fehlvorstellungen zur Evolution in Schülerantworten zu identifizieren?

Sind die Lehramtsstudierenden in der Lage, zwischen anthropomorphen, teleologischen und essentialistischen Fehlvorstellungen zu differenzieren?

Wie hängen das deklarative evolutionsbiologische und biologiedidaktische Wissen mit dem prozeduralen diagnostischen biologiedidaktischen Wissen zusammen?

Wie hängen das prozedurale diagnostische biologiedidaktische und das prozedurale pädagogisch-psychologische diagnostische Wissen zusammen?

#### Methode

Beim Schülerinventar (Kaiser et al., 2015) handelt es sich um ein Instrument, mit dessen Hilfe diagnostisches Wissen von Lehramtsstudierenden im experimentellen Setting am Computer gemessen werden kann. Es ermöglicht urteilsrelevante (hier die jeweilige Schülerfehlvorstellung) und -irrelevante Informationen (z. B. Noten, kognitive Fähigkeiten) zu Schülerinnen und Schülern zusammenzustellen. Zur Beantwortung der Forschungsfragen werden nachgewiesene evolutionsbiologische Fehlvorstellungen (u.a. Baalman et al., 2004), die bereits in der Literatur beschrieben wurden, in schriftliche Erläuterungen der virtuellen Schülerinnen und Schüler integriert (urteilsrelevante Informationen). Diese Schülertexte wurden auf Basis des Argument-Pattern (Toulmin, 2003) konzipiert und unterscheiden sich in der Artikulation von anthropomorphen, teleologischen und essentialistischen Fehlvorstellungen sowie wissenschaftlich korrekten Erklärungen zum Industriemelanismus des Birkenspanners. Ein Interrating der Texte durch drei Experten hinsichtlich der verschiedenen Fehlvorstellungen ergab eine Übereinstimmung von  $Rel = .94$ . Die Urteilsgenauigkeit (prozedurales diagnostisches Wissen) der Probanden (N=80 Studierende, Studiengang Master of Education, Fach Biologie, erstes Semester) wird über die Identifizierung der Fehlvorstellungen und der jeweiligen Fehlvorstellungsvarianten gemessen. Für die Berechnung von Interaktionseffekten zwischen irrelevanten Informationen und Urteilsgenauigkeit erhalten die Probanden zu jeder virtuellen Schülerin / jedem virtuellen Schüler Angaben zu Geschlecht, Alter, kognitiven Fähigkeiten sowie die letzte Zeugnisnote in Biologie, Deutsch und Mathematik. So wird der Zusammenhang zwischen prozeduralem diagnostischem Wissen im biologiedidaktischen (Genauigkeit der Beurteilung der Schülerfehlvorstellungen) und pädagogisch-psychologischen Wissensbereich (Beeinflussung der

Beurteilung durch urteilsirrelevante Informationen) ermittelt. Zur Messung des deklarativen fachwissenschaftlichen und fachdidaktischen Wissens werden geschlossene Items aus dem KiL-Projekt des IPN eingesetzt.

Ergebnisse

Die Studie liefert erste Hinweise zum diagnostischen Wissen von Biologie-Lehramtsstudierenden hinsichtlich der Identifizierung evolutionsbiologischer Fehlvorstellungen. Auch werden Erkenntnisse über den Zusammenhang von deklarativem Wissen (fachdidaktisch, fachwissenschaftlich) und prozeduralem Wissen erlangt. Über die gezielte Beeinflussung der Probanden durch urteilsirrelevante Informationen werden erste Hinweise auf den Zusammenhang zwischen diagnostischem Wissen im biologiedidaktischen und pädagogisch-psychologischen Wissensbereich herausgearbeitet.

## **Beurteilungskriterien und Feedbackstrategien zum historischen Schreiben bei Geschichtslehrpersonen**

**Monika Waldis, Philipp Marti, Martin Nitsche, Kristine Gollin**  
PH FHNW

Theoretischer Hintergrund und Fragestellung

Die Fähigkeit des Schreibens von Sachtexten wird als zentrale Kompetenz in einer Wissens- und Kommunikationsgesellschaft und als Teilaspekt der allgemeinen Studierfähigkeit betrachtet. Zugleich kann Schreiben in höheren Schulstufen als Werkzeug für die Strukturierung fachlicher Lernprozesse genutzt werden. Das Potential von materialbasierten Schreibaufgaben im Geschichtsunterricht zur Initiierung von historischen Denkprozessen ist in der internationalen Literatur mehrfach beschrieben worden (van Drie et al. 2015; Monte Sano & De La Paz 2012; Voss & Wiley 1997). Im deutschsprachigen Raum fehlen bislang erprobte Konzepte zur Schreibförderung im Fach und die Vorstellungen davon, welchen Qualitätskriterien historische Schülererzählungen genügen sollen, variieren in der rezenten Literatur beträchtlich (Barricelli 2005; Hartung 2013; Pandel 2010). Hingegen bestehen Modelle zu Kompetenzen historischen Denkens (Körber et al. 2007; van Drie & van Boxtel, 2008), die als Heuristik bei der Entwicklung kompetenzorientierter Aufgaben und der Beurteilung von Schülertexten herangezogen werden können. Darauf aufbauend können sowohl in summativer als auch in formativer Hinsicht fachlich begründete Beurteilungskriterien von Schreibprodukten und Feedbackstrategien zur Förderung des historischen Denkens abgeleitet werden (Nitsche & Waldis 2016). Die Interventionsstudie "Schülerinnen und Schüler schreiben Geschichte" erforscht die Entwicklung der Diagnosekompetenz zur Förderung von Kompetenzen historischen Denkens bei Gymnasiallehrkräften. Im Folgenden wird danach gefragt, a) welche Aufgabenformate, Beurteilungskriterien und Feedbackstrategien im gymnasialen Unterricht gegenwärtig Anwendung finden und b) ob im Rahmen einer gezielten Schreibintervention in Gymnasialklassen, welche die Lehrpersonen in ein Kompetenzmodell sowie in die Entwicklung von Beurteilungskriterien und Feedbackstrategien einbindet, die Beteiligten eine an Kompetenzdimensionen angelehnte Diagnostik zu entwickeln beginnen.

Methode

Im Rahmen der Interventionsstudie "Schülerinnen und Schüler schreiben Geschichte" wurden in der Vorstudie mit 15 Gymnasiallehrkräften (Auswahl: theoretical sampling) ein teilstrukturiertes Interview zur Schreibpraxis im Geschichtsunterricht, zur Ausgestaltung von Prüfungen sowie zu Beurteilungskriterien von Schülertexten durchgeführt. Parallel dazu wurde eine onlinebasierte Schreibplattform entwickelt, die materialbasierte Schreibaufgaben für Schülerinnen und Schüler der 10. und 11. Gymnasialstufe zu den beiden Themen Migration und Nationalismus bereit hielt. Die Lehrkräfte (n = 16) wurden bei deren Implementation im Unterricht unterstützt. Zentraler Teil der professionellen Weiterbildung stellte die Einführung in ein Kompetenzmodell historischen Denkens (Körber et al. 2007) sowie darauf aufbauend die gemeinsame Erarbeitung von Textbeurteilungskriterien und Feedbackstrategien dar. Zur Erfassung der diagnostischen Kompetenz wurde den beteiligten Geschichtslehrpersonen vor und nach der Intervention zwei Textvignetten vorgelegt mit der Aufforderung, die Schülertexte zu korrigieren, zu beurteilen sowie ein Feedback zur Textüberarbeitung zu formulieren.

Die Auswertung der Lehrerinterviews in der Vorstudie erfolgte mittels qualitativ-interpretativen Zugangs. Für die Analyse der Textvignetten wurde ein "Scoring rubric" entwickelt, mit dessen Hilfe die Nennung fachlicher Beurteilungskriterien wie Umgang mit Zeit, Materialbezug, Quellenkritik, Kontextualisierung, Umgang mit Fachbegriffen, Begründung der historischen Orientierung etc. und die Nennung formal-sprachliche Aspekte erfasst werden konnte. Die entsprechenden Aspekte wurden auch bei der Analyse des Feedbacks erfasst.

Ergebnisse

Die Lehrerinterviews der Vorstudie ergaben, dass die Lehrpersonen nur ansatzweise über Förderkonzepte des historischen Denkens durch Schreiben verfügen. In Geschichtsprüfungen werden zumeist Essays oder Erörterungen gefordert. Die angewandten Beurteilungskriterien umfassten Praktiken wie die Auszählung inhaltlicher Statements oder die Globalbeurteilung der in den Schülertexten aufscheinenden Argumente. Die Arbeit mit Materialien, deren sorgfältige Auswertung und Inbeziehungsetzung wurde z.T. erwartet, jedoch nicht systematisch in die Bewertung einbezogen. Vereinzelt wurden fachliche Beurteilungskriterien genannt; meistens standen formal-sprachliche Aspekte im Vordergrund. Diese Vorgehensweise dominierte auch die Beurteilung der Textvignetten im Vortest durch die Interventionsgruppenlehrpersonen. Ihr Feedback umfasste zumeist eine Globalaussage zur Textqualität ("gut geschrieben"), gepaart mit inhaltlichen Kommentaren. Für die geplante Präsentation wird ein Pre-/Post-Vergleich der Beurteilungskriterien und Feedbackstrategien in der Interventionsgruppe durchgeführt. Es wird erwartet, dass im Verlauf der durchgeführten Intervention die Liste der Beurteilungskriterien theoretisch angereichert und Sicherheit in der Einschätzung von deren Ausprägung erworben wird. Zudem wird erwartet, dass qualitative und quantitative Veränderungen beim Feedback sichtbar werden.

## **Diagnosekompetenz von Englischlehrkräften bei Essays aus der gymnasialen Oberstufe**

**Stefan Keller<sup>1</sup>, Jens Möller<sup>2</sup>**  
<sup>1</sup>PH FHNW, <sup>2</sup>Universität Kiel

Theoretischer Hintergrund

Eine der zentralen Aufgaben von Lehrkräften ist die Leistungsbeurteilung von Schülerinnen und Schülern (Kaiser, Südkamp & Möller, 2016). Sie beeinflusst nicht nur die Selbsteinschätzung der Schülerinnen und Schüler (Möller & Trautwein, 2015), sondern auch die Unterrichtsgestaltung, da Lehrkräfte ihren Unterricht an den wahrgenommenen Kompetenzen der Schülerinnen und Schüler orientieren. Die meisten experimentellen Studien zu dem Thema untersuchen den Einfluss von leistungsirrelevanten Schülermerkmalen auf relativ globale Urteilsmaße wie Noten; fachspezifische Urteile werden dabei oft vernachlässigt.



Die vorliegende Arbeit stellt deshalb einen fachlichen Aspekt von Diagnosekompetenz in den Fokus, welcher für Englisch-Lehrkräfte auf der Oberstufe zentral ist: Die Beurteilung authentischer und komplexer argumentativer Essays. Dabei wird die Auswirkung des Geschlechts der Schülerinnen (Studie 1) sowie der lexikalischen Qualität der Texte (Studie 2) auf die Lehrerbeurteilungen untersucht, so dass sich folgende Forschungsfrage ergibt: Wie wirken sich die Variablen „Geschlecht“ und „lexikalische Qualität“ auf die Beurteilungen von englischen Schülertexten aus der gymnasialen Oberstufe durch Lehrkräfte aus? Diese Studien sind Teil des internationalen Forschungsprojekts ASSET (Assessing Students' English Texts).

#### Methode

Die Untersuchungen werden mit dem „Schülerinventar: ASSET“ (SIA) durchgeführt (vgl. dazu auch Beitrag von Fischer et al. in diesem Symposium). Mit SIA können bei englischsprachigen Schüleressays leistungsrelevante und -irrelevante Einflussmerkmale auf die Textbeurteilung unabhängig voneinander manipuliert und ihr Einfluss auf die Beurteilung durch Lehrkräfte untersucht werden. Die Beurteilung der Texte durch Probanden erfolgt dabei nach fachdidaktisch begründeten Kriterien (Keller, 2013, Culham, 2003), welche in der Form von Rating-Skalen im SIA abgebildet werden.

In Studie 1 beurteilten N = 27 bzw. N = 25 Probanden jeweils vier argumentative englische Schülertexte zu zwei verschiedenen Themen: „Einfluss von Technologie auf das moderne Leben“, und „Guter Unterricht in der Schule“. Jeder Text war entweder von hoher oder niedriger Qualität und wurde randomisiert entweder einem Mädchen- oder einem Jungennamen zugeordnet. Es sollte untersucht werden, ob sich Interaktionseffekte zwischen dem Thema und dem Geschlecht der Lernenden zeigen, weil Lehrpersonen bei Themen zu Technologie jungen Männern oft eine höhere Sachkompetenz zutrauen (Jackson, 2016).

In der zweiten Studie wurden ebenfalls zwei bessere und zwei schlechtere Texte eingesetzt, dabei aber jeweils die Qualität des Vokabulars systematisch variiert („lexical complexity“). Die Texte wurden von N = 39 Lehramtsstudierenden beurteilt. Hier sollte untersucht werden, wie ein linguistischer Faktor wie Vokabular sich auch die Gesamtbeurteilung der Texte auswirkt, und ob Transfereffekte auf andere, nicht-linguistische Aspekte von Beurteilung zu beobachten sind. Als Analyseverfahren wurde jeweils eine multivariate Varianzanalyse mit Messwiederholung durchgeführt.

#### Ergebnisse

In beiden Studien zeigte sich, dass die Probanden eindeutig zwischen Texten mit hoher und niedriger Gesamtqualität unterscheiden konnten. Bei der Studie „Geschlecht“ ergaben sich Geschlechterunterschiede bei einzelnen negativeren Bewertungen der Texte mit technologischem Thema. Diese erhielten signifikant tiefere Bewertungen, wenn die Lehrkräfte den Text einem Mädchen zuordneten. Bei Texten zum Thema Schule ergaben sich keine signifikanten geschlechtsspezifischen Unterschiede.

Bei der Studie „Vokabular“ ergab sich ein Halo-Effekt der Dimension Vokabular auf die Einschätzung der Struktur der Essays: Texte mit schlechterer Vokabularqualität wurden auch von der Struktur (Aufbau) her schlechter eingeschätzt, obwohl diese Dimensionen textlinguistisch völlig unabhängig sind.

#### Diskussion

Die vorliegenden Studien geben fachspezifische Einblicke in die Diagnosekompetenz von Lehramtsstudierenden im Fach Englisch. Diese waren fähig, bessere von schlechteren Texten zu unterscheiden, es zeigten sich in beiden Fällen aber auch Halo-Effekte, die auf systematische Urteilsfehler hindeuten. Die hier entwickelten Untersuchungsinstrumente können auch als Trainingsinstrumente in der Aus- und Weiterbildung von Lehrpersonen eingesetzt werden.

## E14

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H117

### Professionelles Handeln bei der Beurteilung von Schülertexten - Empirische Befunde

Chair(s): **Cristina Vögelin** (Pädagogische Hochschule FHNW, Schweiz)

DiskutantIn(nen): **Jan Retelsdorf** (Universität Hamburg, Deutschland)

Die Beurteilung von schriftlichen Arbeiten gehört zu den zentralen Kompetenzen von Lehrkräften in unterschiedlichen Fachgebieten. Sie spielt eine wichtige Rolle in der Auswahl von passendem Unterrichtsmaterial und liefert Informationen über den Fortschritt der Lernenden (Baumert & Kunter, 2006; Schrader, 2013). Ferner bildet die Textbeurteilung die Basis für Notenentscheidungen, Klassenübertritte und Diplome (Kronig, 2007).

Eine schriftliche Arbeit zu verfassen, sei es fiktive Erzählliteratur oder ein argumentativer Aufsatz, erfordert von den Lernenden eine Vielzahl an linguistischen, rhetorischen und strategischen Fähigkeiten. Um diese komplexen und vielseitigen Texte fair beurteilen und den Lernenden eine hilfreiche Rückmeldung geben zu können, müssen Lehrkräfte in ihrer Ausbildung ausreichend darauf vorbereitet werden. Jedoch kommt die Beurteilung von schriftlichen Arbeiten in der Lehrerbildung oft zu kurz (Porsch, 2010). In Folge dessen fühlen sich Lehramtsstudierende in Deutschland und in der Schweiz am Ende ihrer Ausbildung oft nicht genügend qualifiziert (Rauin & Meier, 2007). Somit bedarf es einer Verbesserung der Lehrerbildung im Hinblick auf die Beurteilung von schriftlichen Schülerarbeiten. Um dieses Defizit zu überwinden, müssen wir mehr über die genauen Beurteilungskompetenzen der Lehrkräfte sowie über die Einflussfaktoren im Bewertungsprozess wissen. Wenig empirische Studien haben sich mit spezifischen Faktoren und deren Einfluss auf die Beurteilung von Schülertexten befasst.

Das folgende Symposium knüpft an diesem Forschungs- und Entwicklungsdefizit an. Es werden vier empirische Studien vorgestellt, welche genauer untersuchen, wie angehende und erfahrene Lehrkräfte im Fach Deutsch, Englisch und Ethik Schülertexte beurteilen. Folgende Fragen werden adressiert: Wie genau oder wie gut können Lehrkräfte Texte beurteilen und den Lernenden rückmelden? Welche Einflussfaktoren können dabei identifiziert werden und wie beeinflussen diese die Beurteilung von anderen Textmerkmalen? Dabei kommen Genauigkeit der Textbeurteilungen, der Einfluss bestimmter Faktoren sowie die schriftliche Rückmeldung von Lehrkräften in den Blick. Es wird sowohl die Beurteilung von schriftlichen Arbeiten aus der Primar- wie auch aus der Sekundarschule empirisch erfasst und es wird diskutiert, wie diese objektiver und fairer werden kann.

Im ersten Beitrag stellen Lindauer, Sommer und Sturm einen Teil aus dem Projekt NoviS „Novizinnen und Novizen im Schreibunterricht“ vor, in welchem sie näher auf das Beurteilen und Kommentieren von Schülertexten im Fach Deutsch auf der Primarstufe eingehen. Im zweiten Beitrag wird die Expertise der Lehrkräfte in den Fokus gerückt, indem Hennes die Genauigkeit der Lehrkräfteurteile im Fach Deutsch auf Primarstufe mithilfe des Experten-Novizen-Paradigmas untersucht. Im Fach Englisch auf der Sekundarstufe präsentieren Jansen und Vögelin eine Studie aus dem Forschungsprojekt ASSET „Assessing Students' English Texts“, in welcher einerseits der Einfluss von Rechtschreibung auf die Beurteilung von anderen Textmerkmalen erforscht wird und andererseits untersucht wird, ob sich dieser Effekt durch einen Prompt reduzieren lässt. In einem ähnlichen Untersuchungsdesign analysiert Machts, wie die Textbeurteilung von angehenden Lehrkräften im Fach Ethik durch Textoberflächen- und Texttiefermerkmale beeinflusst wird.

#### Beiträge des Symposiums

### Schülertexte fachlich angemessen beurteilen und lernförderlich kommunizieren – was (angehende) Lehrpersonen hierzu mitbringen

**Nadja Lindauer, Tim Sommer, Afra Sturm**  
Pädagogische Hochschule FHNW, Schweiz

Schülertexte angemessen beurteilen und auf dieser Basis ein schriftliches Feedback an die SchülerInnen zu formulieren, das sie bei der Überarbeitung ihrer Texte wirksam unterstützt, setzt auf Seiten der Lehrperson hohes fachdidaktisches Wissen voraus: Sollen die SchülerInnen ihren Text nicht nur auf der lokalen Ebene, z.B. bezüglich Rechtschreibung oder Handschrift, revidieren, sondern dazu befähigt werden, auch Überarbeitungen auf der Textebene vorzunehmen, so müssen sie auch sogenannte höhere Aspekte wie Inhalt oder Struktur wahrnehmen können. Dies wiederum bedingt, dass auch die Lehrpersonen globale Aspekte zu erkennen und adäquat zurückzumelden vermögen. Wie etwa die Studie von Parr und Timperley (2010) zeigt, fokussieren die schriftlichen Kommentare von Lehrpersonen jedoch häufig die lokale Ebene. Des Weiteren sind sie oftmals nicht im jeweiligen Schülertext verankert, fallen ambig und teilweise widersprüchlich aus (Cho & MacArthur, 2010). Dies kann durch fehlendes fachdidaktisches Wissen bedingt sein: Lehrpersonen dürften dann dazu tendieren, in ihrem Feedback lediglich allgemeine Prinzipien zu formulieren, die für die SchülerInnen nicht operationalisierbar sind (Myhill, Jones & Watson, 2013). Nicht zuletzt haben direktiv formulierte Kommentare im Vergleich zu dialogisch angelegten einen geringeren Effekt auf die Textüberarbeitung der SchülerInnen (MacArthur, 2015).

Vor diesem Hintergrund wurde im SNF-Projekt "NovizInnen im Schreibunterricht (NoviS)" danach gefragt, welches fachdidaktische Wissen angehende Lehrpersonen in Bezug auf das Beurteilen und Kommentieren von Schülertexten mitbringen und inwiefern sich dieses im Verlauf des ersten Berufsjahres verändert. Zur Klärung dieser Fragen wurden Primar-Studierenden der PH FHNW am Ende der Ausbildung (N=99) sowie nach dem ersten Berufsjahr (N=51) zwei offene Aufgaben zu einem Schülertext vorgelegt: Im Rahmen der ersten Aufgabe sollte die Qualität des Schülertexts beurteilt werden. Die zweite Aufgabe bestand darin, einen schriftlichen Kommentar an den Schüler zu schreiben und darin auch Tipps zu formulieren, wie die Geschichte noch besser geschrieben werden könnte. Die Beurteilungen und Kommentare wurden inhaltsanalytisch dahingehend kodiert, ob sie eher die lokale oder globale Ebene fokussieren, ob sie den Blick eher auf Gelungenes oder nicht Gelungenes legen und ob sie im Schülertext verankert sind. Eine Einschätzung der Qualität erfolgte bei den Beurteilungen durch einen Vergleich mit einem Expertenurteil, und zwar im Hinblick auf die lokale und globale Ebene. Bei den Kommentaren wurde die Qualität mit den Kriterien "formulierte Tipps" (im Schülertext verankert vs. nicht verankert) und "Form" (eher direktiv vs. eher dialogisch) ermittelt. Die Interrater-Reliabilität wurde mit Cohen's Kappa berechnet und lag bei Werten über 0.6.

Sowohl für die Textbeurteilungen als auch für die Kommentare zeigte sich kein Schwerpunkt in Bezug auf die Ebene (eher lokal, eher global). Weniger als ein Drittel der (angehenden) Lehrpersonen vermochte allerdings eine angemessene Textbeurteilung zu verfassen, wie aus dem Vergleich mit dem Expertenurteil hervorgeht. Ähnliches gilt für die Kommentare: Sie enthielten selten nachvollziehbare Tipps für die SchülerInnen und wiesen einen hohen Anteil an direktiv formulierten Rückmeldungen auf. Textbeurteilung und Kommentar korrelieren nur gering (bspw. in Bezug auf Ebene und nur zu t1:  $r_s=.21$ ,  $p=.035$ ). Inhaltlich korrespondierten sie ebenfalls nur bedingt miteinander: In den Kommentaren manifestierte sich eine deutlich

stärkere Tendenz, Positives hervorzuheben – ungeachtet der tatsächlichen Qualität des Schülertextes – und lokale und globale Aspekte gleichermaßen zu berücksichtigen. Dies kann anschliessend an Bouwer et al. (2016) als eine Frage des Feedback-Stils interpretiert werden. Hinsichtlich der zweiten Forschungsfrage ergaben sich weder bei den Textbeurteilungen noch bei den Kommentaren signifikante Veränderungen nach einem Jahr Berufserfahrung (bspw. in Bezug auf Ebene Textbeurteilung:  $Z = -2.19$ ,  $N=49$ ,  $p=.095$ ; sowie Kommentar:  $Z = -.18$ ,  $N=49$ ,  $p=.857$ ).

## **Textbewertung – Die Expertise der Lehrkräfte im Fokus**

**Ann-Kathrin Hennes, Barbara Schmidt, Alfred Schabmann**  
Universität zu Köln, Deutschland

### Hintergrund

Kinder und Jugendliche mit mangelnder Schreibkompetenz sind in ihrer gesellschaftlichen Teilhabe beeinträchtigt (Becker-Mrotzek, 2014). Sie benötigen daher eine möglichst frühe und gezielte Förderung. Dies setzt jedoch voraus, dass Lehrpersonen in der Lage sind, Kinder und Jugendliche mit Problemen zu erkennen. Lehrpersonen müssen dazu die Schreibprodukte von Schüler\*innen entsprechend beurteilen können (Hennes et al., 2017).

Allerdings sind Texte aufgrund ihrer Komplexität generell schwierig zu bewerten (Feenstra, 2014). Zudem fehlen eindeutige Textbewertungskriterien für den schulischen Kontext. In der Forschungsliteratur (z.B. Knopp et al., 2012) werden zwar Kriterien guter Texte benannt (z.B. adäquater Informationsgehalt, Adressatenorientierung und Kohärenz), diese beziehen sich aber zumeist auf nicht direkt beobachtbare Dimensionen eines Textes, was ihre tatsächliche Anwendung erschwert.

Bisherige Untersuchungen zur Expertise von Lehrkräften bei der Textbewertung zeigen, dass zwischen den Urteilen von Lehrkräften zu dem gleichen Schülertext eine große Varianz besteht (Birkel & Birkel, 2002). Lehrpersonen stimmen nicht miteinander überein, wenn es darum geht, schlechte/gute Texte zu identifizieren. Mögliche Gründe für die bestehende Varianz zwischen den Lehrkräftenurteilen sind bisher kaum erforscht. Es ist daher unklar, ob die Komplexität der Textbewertung an sich dafür verantwortlich ist, oder ob es Lehrkräften an Wissen zu Textbewertungskriterien mangelt (während andere Personengruppen, wie beispielsweise Lektoren oder Autoren, möglicherweise über Kenntnisse in diesem Bereich verfügen und übereinstimmende Urteile abgeben).

### Forschungsfragen

Anknüpfend an den beschriebenen theoretischen und empirischen Forschungsstand lautet die Forschungsfrage der vorliegenden Studie:

- Wie akkurat sind Lehrkräftenurteile im Kontext der Textbewertung (im Vergleich zu den Urteilen anderer Personengruppen)?

### Methode

Im Gegensatz zu den meisten Studien mit Fokus auf die diagnostischen Kompetenzen von Lehrkräften können im Kontext der Textbewertung keine standardisierten Diagnoseinstrumente als Vergleichsmaßstab herangezogen werden, da kein deutschsprachiges Diagnoseinstrument zur Erfassung der Schreibkompetenz existiert.

In der vorliegenden Studie wird daher ein neuartiger Forschungsansatz verwendet, welcher das Experten-Novizen-Paradigma der Psychologie (Bromme, 2001) mit einem methodischen Ansatz der Kreativitätsforschung (consensual assessment technique; Kaufman, Baer, Cropley, Reiter-Palmon & Sinnett, 2013) kombiniert. Beide Ansätze gehen davon aus, dass Personen mit hoher Expertise auf ihrem jeweiligen Gebiet zuverlässige Bewertungen vornehmen und hierin mit anderen Experten des relevanten Gebietes übereinstimmen.

Übertragen auf den Kontext der Textbewertung wurden im Rahmen dieser Studie 20 Texte von 19 professionellen Lektoren (Experten1), X Grundschullehrkräften (Experten2), 15 Wissenschaftlern im Bereich der Schreibforschung (Experten3) und 36 Studierende ohne Hintergrundwissen (Novizen) bewertet. Alle Rater haben die gleichen 20 Texte als Ganzes auf einer Skala von 1 (sehr schlecht) bis 6 (sehr gut) bewertet. Nach einer Pause von mindestens drei Tagen wurden dieselben 20 Texte im Hinblick auf die aus der Literatur bekannt Textbewertungskriterien bewertet.

### Auswertung

Die Gruppe der Novizen und die verschiedenen Expertengruppen werden hinsichtlich 1) ihrer Interrater-Reliabilität (IRR) und 2) der verwendeten Textbewertungskriterien miteinander verglichen.

In Anlehnung an die bisherigen Forschungsbefunde zu Textbewertung durch Lehrkräfte und Ergebnisse aus der Kreativitätsforschung zur Relevanz von Expertise in Bezug auf die Reliabilität von Urteilen wird die folgende Rangfolge angenommen: IRR Lektoren/Wissenschaftler > IRR Lehrer > IRR Studierende.

### Ergebnisse

Die verschiedenen Personengruppen unterscheiden sich bei der globalen Bewertung von Texten nicht in ihren IRR. Sowohl die Expertengruppen als auch die Novizen erzielten gute Werte (.62-.77). In Bezug auf die kriteriengeleitete Bewertung ergibt sich ein Unterschied zwischen den IRR der verschiedenen Gruppen. Die angenommene Reihenfolge (IRR Lektoren/Wissenschaftler > IRR Lehrer > IRR Studierende) kann für diesen Teil der Bewertung bestätigt werden. Die IRR der Lehrkräfte liegen für die verschiedenen Kriterien zwischen .28 und .52 und sind damit als eher niedrig zu bewerten.

Diese Befunde führen zu der Schlussfolgerung, dass Lehrkräfte entlang der bekannten Textbewertungskriterien trainiert werden müssen, damit sie diese zu Bewertung nutzen und mit ausreichender Expertise im diagnostischen und didaktischen Bereich (expliziter Schreibunterricht) agieren können.

## **„Beurteilen Sie nicht nur die Rechtschreibung!“ Ein Experiment zur Textbeurteilung aus dem Projekt ASSET**

**Thorben Jansen<sup>1</sup>, Cristina Vögelin<sup>2</sup>, Nils Machts<sup>1</sup>, Jens Möller<sup>1</sup>, Stefan Keller<sup>2</sup>**

<sup>1</sup>Christian-Albrechts-Universität zu Kiel, Deutschland, <sup>2</sup>Pädagogische Hochschule FHNW, Schweiz

Eine der zentralen Aufgaben von Lehrkräften ist die Einschätzung der Leistungen von Schülerinnen und Schülern (Kaiser, Südkamp, & Möller, 2016). Nur bei adäquater Beurteilung der Leistungen der Lernenden kann eine angemessene Förderung geboten werden (Helwig, Anderson, & Tindal, 2001). Jedoch zeigen Lehrkräfte substantielle Schwierigkeiten Lernertexte adäquat zu beurteilen (Birkel & Birkel, 2002; Ingenkamp & Lissmann, 2008; Kaiser et al., 2016).

In sprachlichen Fächern ist die Beurteilung von Texten zentral. Ein dabei häufig genannter Einflussfaktor ist die Qualität der Rechtschreibung (Birkel & Birkel, 2002; Rezaei & Lovorn, 2010). Birkel und Birkel (2002) zeigten, dass die Rechtschreibung für

7% der Notenvarianz bei der Gesamtnote von deutschen Grundschultexten verantwortlich ist. Rezaei und Lovorn (2010) fanden, dass bei der Beurteilung verschiedener Textmerkmale der Inhalt bei Texten hoher Qualität mit vielen sprachlichen Fehlern negativer beurteilt wurden als bei Texten niedriger Qualität mit wenigen sprachlichen Fehlern.

Im Projekt ASSET „Assessing Students' English Texts“, das als Kooperationsprojekt von englischer Fachdidaktik und Pädagogischer Psychologie durchgeführt wird, untersuchen wir experimentell die Qualität der Beurteilung von argumentativen englischen Lernertexten mit fachdidaktisch begründeten Skalen. In der hier vorgestellten Studie wurde der Einfluss von Qualität und Rechtschreibung auf die Beurteilung verschiedener Textmerkmale untersucht. Außerdem wurde geprüft, ob ein möglicher Effekt durch den Prompt, sich bei der Bewertung anderer Kriterien nicht durch die Rechtschreibleistung beeinflussen zu lassen, reduziert werden kann. Es ergaben sich folgende Forschungsfragen:

1. Werden Lernertexte höherer Qualität auf allen Beurteilungsskalen positiver bewertet als Texte niedriger Qualität?
2. Beeinflusst die Rechtschreibung die Beurteilung von anderen Textmerkmalen?
3. Lässt sich dieser Effekt durch einen Prompt reduzieren?

#### Methode

Die Stichprobe bestand aus N = 66 Lehramtsstudierenden (66% weiblich). Das Durchschnittsalter war M = 24.31 (SD = 4.90). Die Probanden beurteilten mithilfe einer Erweiterung des Schülerinventars (Kaiser, Möller, Helm, & Kunter, 2015) jeweils vier authentische Lernertexte auf einer holistischen und analytischen Skala. Aus 15 argumentativen Essays von Lernenden aus der 11. Klasse (Gymnasium) wurden von Experten mithilfe der NAEP Skala (Board, 2010) zwei Texte auf der Qualitätsstufe „hoch“ und zwei Texte auf der Qualitätsstufe „niedrig“ ausgewählt und in der Länge angeglichen. Außerdem wurde die Rechtschreibung in den Texten variiert, sodass dieselben Texte einigen Probanden mit guter (ein Fehler je 100 Wörter) und anderen mit schlechter (sieben Fehler je 100 Wörter) Rechtschreibleistung präsentiert wurden. Diese Variation entspricht den typischen Rechtschreibfähigkeiten von Lernenden in der 11. Klasse, welche im Rahmen einer Vorstudie evaluiert wurden.

Unmittelbar vor der Beurteilung las die Experimentalgruppe den folgenden Prompt „Empirische Studien haben gezeigt, dass viele Lehrkräfte sich zu sehr von der Anzahl von Rechtschreibfehlern in Aufsätzen beeinflussen lassen, wenn sie die Qualität von Texten beurteilen.“ Die Kontrollgruppe sah den Prompt „Achten Sie darauf, ausgewogen und fair zu beurteilen.“ Die Probanden bewerteten die Texte auf einer Globalskala und auf den Detailskalen „Gesamtstruktur“, „interne Organisation der Abschnitte“, „Evidenzbasierung der Argumentation“, „Sprachmechanik“, „Grammatik“, „Vokabular“ und „Aufgabenerfüllung“. Als Analyseverfahren wurde eine multivariate Varianzanalyse mit Messwiederholung mit dem Zwischensubjektfaktor Prompt und den zwei Innersubjektfaktoren Textqualität und Rechtschreibung durchgeführt.

#### Ergebnisse

Bezüglich der ersten Forschungsfrage zeigten sich erwartete Effekte der Textqualität auf die Textbewertung. Texte höherer Qualität wurden auf allen Skalen besser bewertet als Texte niedriger Qualität. Texte mit schlechterer Rechtschreibung wurden zudem auf allen analytischen Kriterien negativer beurteilt. Zudem zeigten Interaktionseffekte, dass in der Experimentalgruppe mit Prompt die Rechtschreibung einen geringeren Einfluss auf die Beurteilung der analytischen Skalen hatte.

#### Diskussion

In Zukunft können mit dem hier eingesetzten Instrument weitere Einflussfaktoren der Textbeurteilung systematisch variiert und empirisch untersucht werden. Für Lehramtsstudierende im Fach Englisch stellt dies eine praxisrelevante Lerngelegenheit dar und es kann zu Ausbildungszwecken eingesetzt werden. Durch entsprechende Trainings könnten systematische Urteilsverzerrungen verringert und Bewertungen von Schülertexten und Aufsätzen fairer und objektiver gestaltet werden.

### **Einfluss von Textoberfläche und Texttiefe auf die Beurteilung von Schülertexten im Fach Ethik. Ein Experiment.**

**Nils Machts<sup>1</sup>, Thorben Jansen<sup>1</sup>, Cristina Vögelin<sup>2</sup>, Jens Möller<sup>1</sup>, Stefan Keller<sup>2</sup>**

<sup>1</sup>Christian-Albrechts-Universität zu Kiel, Deutschland, <sup>2</sup>Pädagogische Hochschule FHNW, Schweiz

Von Schülern im Unterricht oder in Klausuren erstellte Texte gehören in sprachlichen Fächern wie Deutsch oder Englisch und in inhaltsbezogenen Fächern, wie Ethik, Wirtschaft oder Politik, zu den wichtigsten Beurteilungsgenständen von Lehrkräften. Dabei können Schülertexte sich auf einer Reihe von Textdimensionen unterscheiden. In Studien zur Textbeurteilung wurde ein starker Fokus auf den Einfluss der Rechtschreibung in Schülertexten auf deren Benotung gelegt. Verschiedene Autoren konnten diesbezüglich bereits zeigen, dass die Rechtschreibung sich auf die Gesamtbeurteilung von Schülertexten auswirkt, auch wenn die Beurteilung in erster Linie inhaltlichen Textkriterien folgen sollte (Birkel & Birkel, 2002; Rezaei & Lovorn, 2010).

Im Rahmen des Kooperationsprojektes ASSET (Assessing Students' English Texts) unter Beteiligung von englischer Fachdidaktik an der Pädagogischen Hochschule FHNW in Basel und Pädagogischer Psychologie der CAU Kiel wurde eine Reihe von Studien begonnen, in denen die differenzierte Beurteilung von Textmerkmalen untersucht wird. In der vorliegenden Studie wird dieser Ansatz auf das Fach Ethik ausgeweitet.

#### Methode

Unsere Stichprobe bestand aus N = 60 Studierenden im Master des gymnasialen Lehramts mit unterschiedlichen Fächerkombinationen. 54.8 % der Teilnehmer waren weiblich, das Durchschnittsalter betrug M = 26.7 Jahre.

Die Schülertexte wurden aus einem Klassensatz Klausurbearbeitungen im Fach Ethik der 10. Klasse eines Gymnasiums in Schleswig-Holstein gewonnen. Anhand von Expertenurteilen wurden vier Texte ausgewählt, deren inhaltliche Güte den Schulnoten 1, 2, 3 und 4 entspricht. Alle Texte hatten eine Länge von etwa 600 Wörtern. Für eine systematische Variation der Textoberfläche wurden zunächst alle Texte von Rechtschreibfehlern und grammatikalischen Fehlern bereinigt. Anschließend wurde für jeden der vier Texte eine Variante mit unauffälligen Oberflächenmerkmalen und eine Variante mit negativeren Oberflächenmerkmalen erstellt. Die unauffällige Variante enthielt zwei Rechtschreibfehler, zwei grammatikalische Fehler und zwei Kommasetzungsfehler. Die negativere Variante enthielt 11 Rechtschreibfehler, acht grammatikalische Fehler, acht Kommasetzungsfehler und 24 Negativanpassungen der verwendeten Wörter und Satzstrukturen. Bei den Negativanpassungen wurde besonders darauf geachtet, dass weder der Inhalt, die Argumentationslinie, noch die Textstruktur beeinflusst wird.

Für die Präsentation der Schülertexte und die Beurteilung verwendeten wir eine Anpassung des Schülerinventars (Kaiser, Möller, Helm & Kunter, 2015). Jede Versuchsteilnehmerin erhielt über unsere Computeranwendung Zugang zu den vier Texten. Die Zuweisung der Textvarianten erfolgte randomisiert unter den Voraussetzungen, dass jede Probandin zwei Texte mit unauffälligen Oberflächenmerkmalen und zwei Texte mit negativeren Oberflächenmerkmalen erhielt und dass nie die zwei inhaltlich stärksten bzw. die zwei schwächsten Texte der gleichen Variante angehören. Alle Texte sollten von unseren

Probanden mit einer Gesamtnote und mit Teilnoten für die Oberflächenmerkmale Rechtschreibung, Grammatik und Wortschatz, sowie für die Tiefenmerkmale Inhalt, Aufbau und Argumentation beurteilt werden.

Für die Auswertung verwendeten wir eine multivariate zweifaktorielle Varianzanalyse (2x4). Die Mehrebenenstruktur der Daten konnte dabei auf Grundlage sehr niedrigerer ICCs für die aggregierten Urteile der Versuchspersonen vernachlässigt werden.

#### Ergebnisse

Multivariate Analysen mit allen Beurteilungsskalen zeigten signifikante Effekte sowohl für die Textoberfläche als auch für die Texttiefe. Univariate Analysen ergaben Effekte sowohl der Textoberfläche als auch der Texttiefe auf alle Beurteilungsskalen. Dabei wurden die Tiefenmerkmale bei schwächerer Textoberfläche im Durchschnitt eine halbe Note negativer und die Oberflächenmerkmale pro Abstufung der Texttiefe im Durchschnitt eine Viertelnote negativer beurteilt. Interaktionseffekte zeigten sich für keine der Beurteilungen.

Wir konnten zeigen, dass angehende Lehrkräfte sowohl bei der Beurteilung von Textoberflächenmerkmalen als auch bei der Beurteilung von Texttiefeigenschaften in erheblichem Maße über alle tatsächlichen Textmerkmale generalisieren. So scheinen Lehrkräfte einerseits mehr Fehler in inhaltlich schwächeren Texten zu sehen, als auch inhaltliche Komponenten in oberflächlich schwächeren Texten negativer zu beurteilen.

## E15

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H116

### Innovative Methoden zur Erfassung und Förderung der Selbstreflexion von Kognition, Metakognition und Emotion beim Lernen

Chair(s): **Charlotte Dignath** (Goethe-Universität-Frankfurt), **Manuela Benick** (Universität des Saarlandes, Deutschland), **Laura Dörrenbächer** (Universität des Saarlandes)

DiskutantIn(nen): **Franziska Perels** (Universität des Saarlandes)

Die Erfassung verschiedener Komponenten selbstregulierten Lernens mittels Selbstberichtsverfahren wird zunehmend kritisch diskutiert. Niedrige Korrelationen mit Leistungsmaßen sprechen dabei für eine geringe Validität dieser Verfahren. Ebenso die Annahme, dass sowohl der Gebrauch (meta-)kognitiver Strategien als auch eigene Emotionen eher situativ als State operationalisiert werden sollten. Diese Überlegungen haben zur Erprobung innovativer Methoden zur Messung der Selbstreflexion von Metakognition und Emotion geführt. Ziel dieses Symposiums ist es, verschiedene Methoden der Erfassung von Selbstreflexion (mit Blick auf Kognition, Metakognition und Emotion) zu kontrastieren, um einen umfassenden Überblick über das aktuelle Forschungsfeld zu geben. Allen Beiträgen gemein ist, dass sie Methoden einsetzen, mittels derer die genannten Variablen situativ erfasst werden. Zwei der Beiträge nutzen dabei Selbstbericht, während die beiden anderen die Selbstreflexion über Kodierungen erfassen. In zwei Beiträgen wird das Selbstreflexionsinstrument auch als Interventionsmaßnahme genutzt.

Die folgenden Fragestellungen werden in den Beiträgen dieses Symposiums übergreifend bearbeitet:

- Wie kann die Selbstreflexion (von Kognition, Metakognition oder Emotion) valide und reliabel erfasst werden?
- Wie hängen diese Maße mit Traitmaßen zusammen?
- Kann die situative Erfassung der Selbstreflexion im Sinne einer Intervention auch Selbstreflexion fördern?

Das Symposium soll einen breiten Überblick geben über unterschiedliche Methoden, die auf eine valide, situative Erfassung von Metakognition und Emotion als State-Variable abzielen. In diesem Zusammenhang kommen innovative Auswertungsmethoden zum Einsatz, die darüber hinaus in einigen Beiträgen als Interventionsmaßnahme zur Förderung von Selbstreflexion genutzt werden.

Die folgenden Einzelbeiträge tragen zur Beantwortung der Fragestellungen des Symposiums bei:

Im Beitrag von NN wird die Methode des Experience Sampling im Rahmen einer Studie zu alltagsbezogenen Lern- und Leistungsemotionen von Studierenden vorgestellt. Basierend auf Selbstberichtsfragebögen zeigen sich Unterschiede zwischen den allgemeinen, negativen und situationsspezifischen Lern- und Leistungsemotionen.

NN stellt in ihrem Beitrag eine Studie vor, in deren Rahmen Interaktives Ambulantes Assessment als eine Weiterentwicklung des Ambulanten Assessments getestet wird. Mittels längsschnittlichem Kontrollgruppendesign werden Studierende zu verschiedenen Aspekten der Metakognition sowie zu kognitiver Ressourcenaufwendung befragt. Die Interventionsgruppe erhält zusätzlich ein automatisiertes Feedback und wird dazu angeregt, individuelle Lernfragen zu generieren. Die mehrbenenanalytische Auswertung zeigt eine hypothesenkonforme Zunahme für die Nutzung metakognitiver Strategien sowie für die Aufwendung kognitiver Ressourcen.

NN stellt in ihrer Studie ein im Rahmen einer Interventionsstudie zur Trainierbarkeit von Selbstreflexion neu entwickeltes Instrument zur Ermittlung der Selbstreflexionsfähigkeit vor. Im Gegensatz zu den beiden vorherigen Beiträgen basiert dieses Instrument nicht auf Selbstbericht, sondern es handelt sich um einen vierstufigen Beurteilungsbogen zur Erfassung der Reflexionstiefe. Die Ausprägung der einzelnen Stufen wird auf einer siebenstufigen Likert-Skala abgetragen. Ebenso wie im Beitrag von Simone Löffler wird neben der Testung des Instruments auch dessen Wirksamkeit als Interventionstool überprüft.

NN stellen zwei Studien vor, die die Wirksamkeit von Rubrics und Portfolio als Methoden der Selbstreflexion für kognitive und emotional-motivationale Aspekte des Lernens testen. Während sich für das Portfolio signifikante Interaktionseffekte zugunsten der Interventionsgruppe zeigen, insbesondere für leistungsschwächere Schülerinnen und Schüler, scheinen Rubrics für die untersuchten Variablen keinen Zugewinn zum Kontrollunterricht zu liefern.

NN wird die Beiträge diskutieren und Implikationen für mögliche Forschungsdesiderate liefern.

#### Beiträge des Symposiums

### Experience Sampling in der Empirischen Bildungsforschung

**Ulrike Nett, Anna-Lena Rottweiler**  
Universität Augsburg

Eine Vielzahl von Fragen der empirischen Bildungsforschung bezieht sich auf das Erleben und Handeln von Lernenden in einer bestimmten Situation. Dennoch wurde der Aspekt der aktuellen Situation in der Empirischen Bildungsforschung lange Zeit kaum berücksichtigt (vgl. Goetz, Bieg, & Hall, 2016). Dies geschieht erst seit kurzem, unter anderem mit Hilfe der Methode des Experience Sampling (z.B. Becker, Goetz, Morger, & Ranellucci, 2014; Bieg, Goetz, & Hubbard, 2013).

Die Methode des Experience Sampling umfasst all diejenigen Methoden, mit denen das Erleben von Studienteilnehmenden in einer alltäglichen Situation über kurze Selbstberichtsfragebögen erfasst wird (vgl. Goetz et al., 2016; Larson & Csikszentmihalyi, 1983; Csikszentmihalyi & Larson, 2014). Inzwischen werden, meist mit Hilfe digitaler Geräte (z.B. Smartphones), die Studienteilnehmenden zu einem (entweder bestimmten oder aber zufälligen) Zeitpunkt über ein Signal dazu aufgefordert, einen kurzen Selbstberichtsfragebogen auszufüllen (vgl. Goetz et al., 2016). Ein bedeutender Vorteil von Selbstberichten via Experience Sampling gegenüber allgemeinen Fragebögen wird dabei vor allem darin gesehen, dass sie das aktuelle Erleben im Moment und nicht verzerrt erfassen (vgl. Ebner-Priemer & Trull, 2009).

Ziel des vorliegenden Beitrags ist es, anhand des Beispiels einer Studie zu Lern- und Leistungsemotionen von Studierenden die Methode des Experience Sampling vorzustellen. Ein definierendes Merkmal von Emotionen ist ihre situationsspezifische Variabilität (vgl. Frijda, 2007; Spielberger et al., 1970). Allerdings haben Lern- und Leistungsemotionen auch relativ große stabile Anteile, sogenannte Traits (vgl. Nett, Bieg, & Keller, 2017). Nun stellt sich die Frage, welche Anteile mit welcher Erfassungsmethode optimal erfasst werden. Sind Traitfragebögen zu wenig situationsspezifisch und valide? Werden flüchtige Emotionen durch die Erfassung über Experience Sampling unterschätzt? Diese Fragen und ihre Implikationen für das Design von Studien, die situationale Aspekte berücksichtigen, werden mit Hilfe einer beispielhaften Studie diskutiert.

Insgesamt 211 Studierende der Psychologie und Wirtschaftswissenschaften (75 % weiblich; MAlter= 21,4 Jahre; Compliance = 91%) nahmen an einer Studie zum Erleben von Lern- und Leistungsempfindungen mit Bezug auf eine bedeutsame Klausur teil. Während zwei Experience Sampling Phasen von jeweils einer Woche wurden die Studierenden in ihrem Alltag zu ihren Lern- und Leistungsempfindungen (Freude, Hoffnung, Angst und Ärger, jeweils Einzelitems, fünfstufige Likertskala) mit Bezug auf die Klausur befragt. Zu sieben Messzeitpunkten am Tag wurden sie zunächst gefragt, ob sie in der vorangegangenen Stunde an die Klausur gedacht haben und, falls sie diese Frage bejahten wurden sie zu ihren Lern- und Leistungsempfindungen während der vergangenen Stunde und im Anschluss daran nach ihren aktuellen Lern- und Leistungsempfindungen befragt. In einem Nachtest wurden ihre allgemeinen Lern- und Leistungsempfindungen (Freude, Hoffnung, Angst und Ärger, jeweils Skalenitems, fünfstufige Likertskala, adaptiert nach Pekrun et al., 2011) erfasst. Es wurden ausschließlich Situationen (42%), zu denen die Studierenden an die Klausur dachten, in die Analysen einbezogen. Die Mehrebenenstruktur der Daten wurde in den Analysen berücksichtigt. Es zeigt sich, dass insbesondere die Mittelwertunterschiede zwischen den allgemeinen negativen Lern- und Leistungsempfindungen und den situationsspezifischen deutlich sind (Freude: Mretrospektiv = 1.38/Maktuell = 1.34/Mallgemein = 1.49; Hoffnung: Mretrospektiv = 1.96/Maktuell = 1.93/Mallgemein = 2.07; Angst: Mretrospektiv = 1.53/Maktuell = 1.53/Mallgemein = 2.37; Ärger: Mretrospektiv = 0.96/Maktuell = 0.92/Mallgemein = 2.59).

Die Zusammenhänge zwischen der kurzzeitig retrospektiven und der aktuellen Erfassung von Lern- und Leistungsempfindungen sind sehr hoch ( $r_{\text{Freude}} = .98/r_{\text{Hoffnung}} = .98/r_{\text{Angst}} = .99/r_{\text{Ärger}} = .99$ ). Die Zusammenhänge zwischen den situationsspezifischen und den allgemeinen Lern- und Leistungsempfindungen sind hingegen relativ niedrig (retrospektiv und allgemein:  $r_{\text{Freude}} = .52/r_{\text{Hoffnung}} = .47/r_{\text{Angst}} = .61/r_{\text{Ärger}} = .44$ ; aktuell und allgemein:  $r_{\text{Freude}} = .52/r_{\text{Hoffnung}} = .45/r_{\text{Angst}} = .60/r_{\text{Ärger}} = .44$ ).

Diese Ergebnisse zeigen: kurzzeitig retrospektive und aktuelle Messungen erfassen nahezu dasselbe Konstrukt. Durch Traitfragebögen und situationsspezifische Selbstberichte hingegen werden substantiell unterschiedliche Konstrukte erfasst. Es wird diskutiert, worin diese Unterschiede bestehen können und Implikationen für zukünftige Forschung gefolgert.

### **Interaktives Ambulantes Assessment (IAA) zur Erfassung und Förderung von Selbstreguliertem Lernen in natürlichen Lernkontexten**

**Simone Löffler**

Karlsruher Institut für Technologie

Bei der Mehrheit der vorhandenen empirischen Arbeiten erfolgte die Erfassung von Selbstreguliertem Lernen (SRL) mittels Fragebogen (vgl. Montalvo & Torres, 2004). Aufgrund der retrospektiven Selbstberichte sind auf diese Weise erhobene Daten jedoch in der Regel mit gravierenden Verzerrungen behaftet (vgl. Shiffrin, Stone & Hufford, 2008). Innerhalb kontrollierter Laborstudien wurden alternative Verfahren, wie Thinking Aloud Technik, Eye-Tracking, Logfile-Registrierung oder Beobachtung verwendet. Diese Verfahren bieten zwar den Vorteil einer zeitnahen und verhaltensbasierten Erfassung, problematisch ist jedoch, dass es sich bei Laborstudien um künstliche Settings handelt und durch die genannten Verfahren entweder Denkprozesse unterbrochen oder Motive und dem Verhalten unterliegende Denkprozesse vernachlässigt werden (vgl. Loeffler, Myrtek & Peper, 2013; Veenman, 2011). Ambulantes Assessment (AA; unter Verwendung elektronischer Tagebücher) erlaubt eine zeitnahe Erfassung und ermöglicht darüber hinaus eine dynamische, situationsspezifische und verhaltensnahe Konzeptualisierung von SRL sowie die Registrierung der interessierenden Konstrukte direkt im Alltag der Probanden. In der Regel werden standardisierte Kurzskalen zur Erfassung der unterschiedlichen Komponenten von SRL verwendet. Kritisch ist jedoch anzumerken, dass es sich auch hierbei um subjektive Angaben handelt.

IAA stellt eine Weiterentwicklung des AA dar. Die interaktiven Komponenten ermöglichen einerseits eine bessere Kontrollierbarkeit der Alltagserhebungen und folglich eine Erhöhung der internen Validität sowie der Compliance in ambulanten Studien (z.B. durch das Senden von Erinnerungsnachrichten auf die Studiensmartphones). Andererseits kann IAA die Integration neu erlernter Strategien in den Alltag unterstützen. Desweiteren können auf Basis der in den ambulanten Erhebungen gewonnenen Daten individuelle smartphone-gestützte Interventionen unmittelbar im (Lern-)Alltag durchgeführt werden.

Im Folgenden soll anhand einer Studie im akademischen Kontext illustriert werden, wie IAA eingesetzt werden kann, um metakognitive Strategien zu fördern.

85 Studierende (2. Semester, Maschinenbau) wurden randomisiert Kontroll (KG)- und Interventionsgruppe (IG) zugeordnet. IG und KG wurden während der Vorbereitung auf eine Klausur (1. Assessmentphase; 14 Tage) mittels elektronischer Tagebücher vor bzw. nach jeder Lernphase zu verschiedenen Aspekten der Metakognition sowie der kognitiven Ressourcenaufwendung befragt. Die KG absolvierte in der Vorbereitungsphase (14 Tage) auf eine zweite vergleichbare Klausur das identische Assessment wie in der 1. Phase. Für die IG erfüllten die elektronischen Tagebücher in der 2. Assessmentphase zusätzlich zur Erfassung (entsprechend der 1. Phase) eine intervenierende Funktion (IAA), indem automatisiertes Feedback und das Generieren individueller Lernfragen den Einsatz metakognitiver Strategien und die Ressourcenaufwendung unterstützten sollten. Es wurde erwartet, dass die IAA-Phase (2. Phase) Selbst-Monitoring- bzw. Selbstreflexionsprozesse (als relevante Aspekte der Metakognition; z.B. Schmitz & Wiese, 2006) bei der IG anregte und sich somit förderlich auf die Anwendung metakognitiver und ressourcenbezogener Strategien auswirkte. Erwartungsgemäß zeigten Mehrebenenanalysen für die 2. Phase eine hypothesenkonforme Zunahme der Nutzung metakognitiver Strategien und der Aufwendung kognitiver Ressourcen bei der IG (im Vergleich zur KG und 1. Phase).

Weitere Anwendungsmöglichkeiten von IAA im Kontext der Förderung von SRL werden anhand von empirischen Beispielen skizziert. Chancen und Risiken des Einsatzes von IAA in Studien zu SRL werden diskutiert.

### **Der Einfluss von Selbstreflexion auf das Wohlbefinden und die Leistungsfähigkeit bei Lehramtsstudierenden**

**Josephine Berger**

Technische Universität Darmstadt

In verschiedenen Untersuchungen konnte bereits gezeigt werden, dass die Fähigkeit, sich selbst zu reflektieren langfristig mit einem gesteigerten Wohlbefinden zusammenhängt (Lyubomirsky, 2001) und einen positiven Einfluss auf die Leistung haben kann (Prochaska et al., 2011). Eine besondere Relevanz von Selbstreflexion wird bei Personen aus dem Lehrkontext gesehen, da eine konstante Reflexion zu einer professionellen Weiterentwicklung

beitragen kann (Malthouse & Roffey-Barentsen, 2013). Daher hat die Förderung der Reflexionsfähigkeit für die Lehrerbildung eine besondere Bedeutung, wird jedoch bisher nicht ausreichend in den Lehrerbildungsprogrammen abgebildet (Abels, 2010).

Selbstreflexion findet nicht immer mit der gleichen Intensität statt (Malthouse & Roffey-

Barentsen, 2013). Angenommen wird, je tiefgreifender die Reflexion durchgeführt wird, desto größer ist der Lernzuwachs. Dies führt wiederum zu einem gesteigerten Wohlbefinden. Bei einer intensiven Reflexion beschäftigt man sich tiefgreifend mit dem Thema und untersucht einzelne Aspekte, wie positive oder negative Ursachen genauer. Weiter werden Handlungsalternativen abgeleitet und präzisiert. Dafür werden Bedingungen für ein geändertes Verhalten definiert und Effekte und Konsequenzen bedacht (Christmann,

2003). Die Trainierbarkeit von Reflexion ist bereits bekannt, jedoch soll überprüft werden, ob ein spezielles Training sich auch auf die Intensität der Selbstreflexion auswirken kann

(Schmitz, 2016).

Fragestellung

1. Die Überprüfung der Güte des entwickelten Reflexionsinstruments
2. Wie beeinflusst die Reflexionskompetenz die Leistungsfähigkeit und wie wirkt sie sich langfristig auf das Wohlbefinden aus?
3. Wie wirken sich die verschiedenen Interventionen aus?

Methode

Vorstudie: Erstellung Reflexionsstufen

Der Beurteilungsbogen zur Erfassung der Reflexionstiefe umfasst vier Stufen. Dabei stellt die erste Stufe noch kein richtiges Reflektieren dar; die zweite Stufe eine oberflächliche Reflexion und erst die weiteren Stufen können als vertiefte Reflexion bezeichnet werden. Die Ausprägung der einzelnen Stufen wird auf einer sieben-stufigen Likert-Skala abgetragen.

1. Deskriptiv

o Ein Berichten/Aufzählen von Ereignissen. Kann als Anfang bzw. Rahmen für das eigentliche reflektierte Schreiben gesehen werden.

2. Analyse der möglichen Ursachen

o Die möglichen Ursachen für ein Ereignis/Verhalten werden analysiert

3. Entwicklung Handlungsalternativen

o Entwicklung von Handlungsalternativen unter Bezugnahme der Realisierungsmöglichkeiten

4. Präzisierung der Handlungsalternativen

o Die Handlungsalternative wird konkretisiert unter Bezugnahme von Rahmenbedingungen, wie zeitliche und räumliche Faktoren oder anderen Voraussetzungen (Hilfsmittel, beteiligte Personen, Materialien, etc.).

Hauptstudie

Anhand der Lehrveranstaltung „Pädagogische Psychologie für Lehramtsstudenten“ soll

die Reflexionskompetenz der Teilnehmenden erfasst und mit der Referatsleistung

(Qualitätsmerkmal) in Bezug gesetzt werden. Weiter ist der Zusammenhang zwischen

Reflexionsfähigkeit und Wohlbefinden von Interesse. Die Erhebung umfasst voraussichtlich circa 120 Studenten und wird im Januar 2018 abgeschlossen sein. Im Vordergrund steht hierbei die geeignete Erfassung von Reflexion in einem experimentellen Setting. Dafür werden die Teilnehmenden in drei Seminargruppen eingeteilt und erhalten zu Beginn der Studie unterschiedliche Interventionen zum Thema Selbstreflexion. Zu dem Beispiel des Modells für selbstreguliertes Lernen sollen die Teilnehmenden ihr eigenes Lernverhalten reflektieren. Die Differenzierung der einzelnen Interventionen liegt dabei in der Intensität der Reflexion. Gruppe 1 reflektiert auf Stufe eins und zwei; Gruppe 2 auf den Stufen eins bis drei und Gruppe 3 auf allen Stufen.

Ergebnisse

Das Kategoriensystem der Reflexionsstufen wurde bereits anhand von 48 Reflexionsberichten getestet und einer Beobachterübereinstimmung unterzogen. Die Interrater-Reliabilität ( $\kappa_{\omega} = .89$ ) kann als sehr gut bezeichnet werden.

Bisherige Studien konnten einen hohen korrelativen Zusammenhang von  $r = .80$  zwischen Selbstreflexion, welche mit Tagebüchern gemessen wurde und dem subjektiven Wohlbefinden zeigen (Schmitz, 2016).

## **Potential von Selbstreflexions- und Selbsteinschätzungsformaten – Rubrics- und Portfolioformate im Vergleich**

**Florian Hofmann, Michaela Gläser-Zikuda**

Friedrich-Alexander-Universität Erlangen-Nürnberg

Die Fähigkeit, selbstreguliert zu lernen, gilt als eine Schlüsselkompetenz in Schule, aber auch Hochschule und Weiterbildung (Friedrich & Mandl, 1997). Selbstreguliert Lernende planen eigenständig, welche Ziele sie erreichen möchten, sie wählen geeignete Lernstrategien, um die Lernhandlung durchzuführen, sie überwachen ihren Lernprozess fortwährend, und nach Abschluss der Lernhandlung bewerten sie, inwiefern sie die gesteckten Ziele erreicht haben (Friedrich & Mandl, 1997; Schmitz & Schmidt, 2007; Zimmerman, 2002).

Die Förderung selbstregulierten und selbstbestimmten Lernens sowie vielfältiger Kompetenzen stellt eine zentrale Aufgabe schulischen Unterrichts dar (Deci & Ryan, 2000; Tillmann & Wischer, 2006). Seit geraumer Zeit existieren vielfältige pädagogische Konzepte und Instrumente, die auf unterschiedlichen Ebenen diese Zielsetzungen adressieren. Das Portfolio wird beispielsweise als Bestandteil einer schülerorientierten Lernumgebung verstanden, die individuelles und selbstreguliertes Lernen anzuregen und zu fördern vermögen (Gläser-Zikuda, 2007; Häcker, 2006). Des Weiteren sammeln sich unter dem Stichwort „formative assessment“ zahlreiche innovative Diagnostikkonzepte, die Informationen für die Unterstützung individueller Lernprozesse sowie die Gestaltung von Lehr-Lernprozessen gewinnen (Black & William 1998; 2009). Selbsteinschätzungen in unterschiedlichsten Formaten (z.B. Rubrics, Kriterienraster; vgl. Andrade & Wang, 2008) besitzen eine Reihe von positiven Effekten auf metakognitive Prozesse und Motivation (Sadler & Good 2006) sowie Schulleistung (Kingston & Nash 2011; Gläser-Zikuda et al., 2006). Mit Blick auf diese Charakteristika und Funktionen von Selbstreflexions- und Selbsteinschätzungskonzepten bzw. -instrumenten werden in Ausschnitten zwei Interventionsstudien vorgestellt.



## Fragestellungen

Die Interventionen beider Studien sahen unter anderem eine Trainingsphase für die Lehrkräfte sowie eine methodisch-didaktische Standardisierung des Unterrichts vor. Beide Studien widmen sich folgenden Forschungsfragen: (1) Welche Effekte auf kognitive und affektive Lernfaktoren und die Kompetenzentwicklung besitzen die Selbstreflexions- und Selbsteinschätzungskonzepte? (2) Worin besteht das besondere Potential dieser Selbstreflexions- und Selbsteinschätzungskonzepte?

## Methode und Ergebnisse

### Interventionsstudie 1 mit Rubrics:

In einer quasi-experimentellen Studie im Deutschunterricht wurden Rubrics eingesetzt, um die Selbsteinschätzung der Schüler/innen anzuregen und die Kompetenzentwicklung zu fördern. Insgesamt nahmen N = 235 Schüler/innen (46,8 % Mädchen) aus neun Klassen der neunten und zehnten Jahrgangsstufe einer Realschule teil. Der Untersuchungsablauf sah vor, dass alle Schüler/innen zunächst eine Kurzzusammenfassung zu einem journalistischen Text verfassten. Anschließend wurden die Schüler/innen in eine Treatment- (n = 125) und Kontrollgruppe (n = 110) geteilt. Die Treatmentgruppe führte nun anhand eines Beurteilungsbogens (Rubrics) eine Kompetenz-Selbsteinschätzung durch. Die Kontrollgruppe bearbeitete eine alternative Aufgabe mit Kompetenzbezug. Die Kompetenz der Schüler/innen „Inhalte zusammenfassen“ wurde zu fünf Messzeitpunkten erfasst. Die Ergebnisse basierend auf einer Varianzanalyse mit Messwiederholung deuten auf einen signifikanten Lernzuwachs für alle teilnehmenden Schüler/innen hin; die Selbsteinschätzungen hatten jedoch keinen bedeutsamen Effekt auf die Kompetenzentwicklung (Hofmann, 2017).

### Interventionsstudie 2 mit Portfolio:

Im Rahmen einer quasi-experimentellen Interventionsstudie im Physikunterricht an vier Gymnasien wurden daher in einem Treatment- Kontrollgruppenvergleich n = 201 Schüler/innen aus vier Klassen der achten Jahrgangsstufe von vier Lehrer/innen über ein Schulhalbjahr hinweg zum Thema Elektrizität unterrichtet. Zur Ermittlung der postulierten Effekte der Portfolioarbeit hinsichtlich der kognitiven und affektiven Lern- und Leistungsresultate wurden mithilfe standardisierter Instrumente in Prä- Post- und Follow-Up-Erhebungen folgende Variablen erfasst: Selbstregulation, soziale Kompetenzen, Lernmotivation, Lern- und Leistungsemotionen, Problemlösekompetenzen sowie Fachleistungen auf zwei Kompetenzniveaus (orientiert an PISA). In diesem Beitrag werden mit Blick auf das Potential des Portfolios die Ergebnisse zu differentiellen Effekten (Leistungs- und Kompetenzniveau, Gender) referiert, da insbesondere bei den leistungsschwächeren Schüler/innen der Treatmentgruppe positive Effekte hinsichtlich der Kompetenzentwicklung und Lernemotionen belegt werden konnten (Gläser-Zikuda et al., 2011; Hagenauer et al., 2016). Die Wirksamkeit des Portfolio-Ansatzes im Sinne einer selbstregulationsfördernden Lernumgebung ließ sich, wie bereits in einer vorhergehenden Interventionsstudie (Gläser-Zikuda, Lindacher & Fuß, 2006), bestätigen.

Im Rahmen der Präsentation werden Rubrics und Portfolio hinsichtlich ihres Potentials für die Unterstützung von Selbstregulation und Kompetenzentwicklung in zwei Unterrichtsfächern diskutiert und verglichen.

## E16

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H115

### Peerinteraktionen in kulturell und sprachlich diversen Settings

Chair(s): **Dominique Rauch** (DIPF, Deutschland), **Michel Knigge** (Universität Potsdam, Deutschland)

DiskutantIn(nen): **Linda Juang** (Universität Potsdam, Deutschland)

Beziehungen zwischen Gleichaltrigen eröffnen vielfältige Erfahrungs- und Bildungsgelegenheiten. Peers sind Bezugspunkt für schulische Integration (Schulleistung, Leistungsmotivation), aber auch die allgemeine Lebenszufriedenheit. Neben dem Deutschen als Verkehrs- und Unterrichtssprache spielen auch die in den Familien gelernten Herkunftssprachen bei der Interaktion unter Gleichaltrigen eine lebhaftere Rolle. Die Nutzung der Herkunftssprachen ist dabei zumeist auf Interaktionen außerhalb des Unterrichts beschränkt. Im Sinne einer Aufwertung und Anerkennung der Herkunftssprachen als Ressource für schulisches Lernen werden Möglichkeiten gesucht Mehrsprachigkeit in den Unterricht einzubeziehen, ohne Lehrkräfte und Schüler\*innen zu überfordern. Dabei stehen sowohl Lernfortschritte, wie auch soziale Aspekte unter der Bedingung der Mehrsprachigkeit im Fokus. Zwei Beiträge des Symposiums fokussieren auf Peer-Learning-Trainings, die explizit die Mehrsprachigkeit von Schüler\*innen aufgreifen (Rauch, Decristan, Reitenbach & Schastak; Gantefort & Maahs). Ein weiterer Beitrag stellt Möglichkeiten vor (mehr)sprachliche Kompetenzen durch Peers einschätzen zu lassen. Diese Studie hat als Validierung eines Instruments zur Fremdeinschätzung hohe methodische Relevanz für die Mehrsprachigkeitsforschung, da diese bislang vor allem auf Selbsteinschätzungen angewiesen war. Der vierte Beitrag des Symposiums beleuchtet die Frage, wie sich der Zusammenhang von Sozialisierung, Identität und Anpassung bei Jugendlichen ohne Migrationshintergrund im Gegensatz zu Jugendlichen mit Migrationshintergrund gestaltet. Das Symposium deckt eine Altersspanne vom Grundschulalter bis zur Adoleszenz und gleichberechtigt mehrsprachigkeits- wie kulturbezogene Aspekte von Peer-Interaktionen ab.

#### Beiträge des Symposiums

### Förderung des Deutsch-Leseverstehens bilingualer Grundschüler durch Peer-Learning (BiPeer)

**Dominique Rauch<sup>1</sup>, Jasmin Decristan<sup>2</sup>, Valentina Reitenbach<sup>1</sup>, Martin Schastak<sup>1</sup>**

<sup>1</sup>DIPF, Deutschland, <sup>2</sup>Bergische Universität Wuppertal, Deutschland

Kinder mit Migrationshintergrund aus der Türkei verfügen bereits in der Grundschule über niedrigere Lesekompetenzen als ihre Mitschüler/-innen ohne Migrationshintergrund (Haag, Böhme, & Stanat, 2012). Das Projekt BiPeer untersucht Möglichkeiten, das Deutsch-Leseverständnis (LV) Türkisch-Deutsch bilingualer Grundschüler/-innen durch Peer-Learning (PL) zu unterstützen. Interventionsstudien haben gezeigt, dass PL schulisch relevante Kompetenzen wie LV fördert und Kinder mit Migrationshintergrund hiervon besonders profitieren (Adesope et al., 2011). Forschung zur Tandem-Zusammenstellung fokussierte bislang hauptsächlich die Variablen Geschlecht, Alter und Leistungsniveau der Peers (Robinson, Schofield, & Steers-Wentzell, 2005). Obwohl kommunikative Fähigkeiten für erfolgreiches PL essentiell sind, wurden die Tandemzusammensetzung hinsichtlich des Sprachhintergrunds der Peers und der Sprachgebrauch bilingualer Peers während des PL bisher kaum erforscht. Ein bilinguals Kind könnte mit einem monolingualen zusammen lernen und dabei vom größeren Vokabular des monolingualen Kindes profitieren (Limbird et al., 2014). Lernen zwei bilinguale Kinder gemeinsam unter dem Angebot, beide Sprachen in der Interaktion zu nutzen, könnte dies die Kommunikation während des PL erleichtern (Hopewell, 2011) und so zu höherem LV führen. Im BiPeer Projekt werden der Sprachhintergrund der Peers und die Sprachnutzung während des PL variiert.

#### Hypothesen

1. Das PL-Training verbessert das Leseverstehen.
2. Bilinguale Kinder verbessern sich in monolingual-bilingualen Tandems stärker im Leseverstehen als in bilingual-bilingualen Tandems, wenn während des PL nur Deutsch gesprochen wird.
3. Bilinguale Kinder verbessern sich in bilingual-bilingualen Tandems stärker im Leseverstehen, wenn Sie während des PL beide Sprachen nutzen können.

#### Methode

Das Projekt nutzt ein Treatment-Kontrollgruppendesign: Es werden drei Lese-Interventions-Gruppen (LG1-3) mit drei Rechen-Kontroll-Gruppen (RG1-3) verglichen. In LG1 und RG1 lernen ein monolinguales und ein bilinguals Kind, in allen anderen Gruppen zwei bilinguale Kinder zusammen. In LG2 und RG2 arbeiten die Kinder ausschließlich auf Deutsch, in LG3 und RG3 auf Deutsch und Türkisch. Die Kinder wurden nach Vorleistungen parallelisiert den Gruppen zugewiesen.

Die Kinder nahmen an einem sechswöchigen PL-Training (12 Sitzungen à 45 Minuten am Nachmittag) mit reziproken Rollen teil. Das Lesetraining fokussierte auf Tandemlesen und den Erwerb und die Einübung von Lesestrategien (Wörtererklären, Zusammenfassen, Vorhersagen).

Die hier vorgestellten Analysen verwenden Daten von 168 Kindern, davon 138 Türkisch-Deutsch bilingual. 86 dieser Kinder nahmen an der Leseintervention teil (LG1: 32, davon 16 bilingual; LG2: 25; LG3: 25). 62 Kinder waren weiblich und das durchschnittliche Alter betrug 9,6 Jahre. Das LV wurde vor und nach der Intervention mit dem ELFE 1-6 (Lenard & Schneider, 2006;  $\alpha = .92$ ) erhoben.

#### Ergebnisse

Die gerichteten Hypothesen wurden durch Varianzanalysen mit Messwiederholung überprüft. Hypothese 1 wurde bestätigt: Kinder die am Lesetraining teilgenommen haben, verbesserten ihr Leseverständnis stärker als Kinder die am Rechentraining teilgenommen hatten ( $F(1,159)=3.22, p=.035$ ). Die Hypothesen 2 ( $F(1,37)=0.58, p=.23$ ) und 3 ( $F(1,47)=1.12, p=.15$ ) werden zurückgewiesen: Es gab keine signifikanten Unterschiede im Leseverständnisszuwachs zwischen den drei Lese-Interventionsgruppen. Die mittlere Post-Prä-Differenz betrug in der LG1 9.93 (SD=6.81), in der LG2 12.00 (SD=8.99) und in der LG3 9.32 (SD=8.72) Punkte im Leseverständnistest.

#### Diskussion

Trotz der relativ kurzen Dauer und des außerunterrichtlichen Settings konnten die Kinder vom Lesetraining profitieren. Die deskriptiven Daten deuten darauf hin, dass sich Kinder in allen drei Gruppen im Leseverständnis verbesserten – PL scheint somit für die deutlich benachteiligte Gruppe der Kinder mit türkischem Migrationshintergrund eine effektive Methode zur Leseförderung zu sein. Darüber hinaus scheint es für bilinguale Kinder keinen wesentlichen Unterschied zu machen, ob sie mit

einem mono- oder bilingualen Partner lernen und ob Sie nur Deutsch oder Türkisch und Deutsch sprechen. Das Lernen in Tandems mit gleichem Sprachhintergrund bei Nutzung nur des Deutschen als Arbeitssprache scheint den deskriptiven Daten nach vielversprechend zu sein. Die Ergebnisse werden sowohl im Hinblick auf die grundlagentheoretischen wie auch anwendungsbezogenen Implikationen diskutiert.

### **Fremdeinschätzungen als Maße zur Erfassung sprachlicher Fähigkeiten in Peernetzwerken?**

**Nora Dünkel<sup>1</sup>, Michel Knigge<sup>2</sup>, Jürgen Wilbert<sup>2</sup>**

<sup>1</sup>Universität Hamburg, Deutschland, <sup>2</sup>Universität Potsdam, Deutschland

Beziehungen zwischen Gleichaltrigen eröffnen vielfältige Erfahrungs- und Bildungsgelegenheiten. Je nach sozialer Einbettung bieten sich unterschiedliche Zugänge zu kulturellem und sozialem Kapital und damit unterschiedliche Chancen für Bildungserfolg (vgl. Bourdieu 1983; Coleman 1988). Innerhalb von Lernräumen ist Sprache zugleich Medium zur Aneignung und Wiedergabe von Wissen und stellt daher einen wesentlichen Faktor für Bildungserfolg dar (Gogolin 2009; Esser 2006). Neben dem Deutschen als Verkehrssprache spielen auch die in den Familien gelernten Herkunftssprachen bei der Interaktion unter Gleichaltrigen eine lebhafte Rolle.

Trotz der hohen Bedeutung von Beziehungen unter Jugendlichen, liegen zu ihrem Einfluss auf die Entwicklung sprachlicher Fähigkeiten nur wenige Forschungsergebnisse vor. Egozentrierte Netzwerke stellen eine erprobte und ökonomische Variante dar, um Personen aus dem sozialen Umfeld unabhängig von deren Studienteilnahme zu erfassen. Eine Befragungsperson (Ego) kann dabei Angaben zum Peerumfeld machen und z.B. Einschätzungen über die sprachlichen Fähigkeiten von Peers vornehmen. Die Qualität dieser Angaben hängt jedoch wesentlich von der Qualität der Fremdeinschätzungen ab. Forschungen, die sich mit Faktoren der Urteilsgenauigkeit von Fremdeinschätzungen befassen, kommen zu teilweise konträren Ergebnissen. Zum einen scheint die Qualität von Einschätzungen mit steigendem Bekanntheitsgrad zuzunehmen (vgl. Borkeau & Liebler 1993; Marsh & Craven 1991). Zum anderen sprechen Forschungsarbeiten gegen einen positiven Effekt von Freundschaft. So können sowohl Fremde als auch Freunde sich gleichermaßen gut einschätzen, sofern es sich um Eigenschaften handelt, die gut beobachtbar sind (vgl. Vazire 2010). Andererseits konnten leistungsbezogene Vorurteile auch unter befreundeten Schüler(inne)n nachgewiesen werden (vgl. Zander et al. 2014). Neben der Beziehung zwischen urteilender und beurteilter Person, scheint auch die Kompetenz der einschätzenden Person für die Urteilsgenauigkeit eine Rolle zu spielen (vgl. Siebert 2006).

Der Beitrag fokussiert die Frage nach der Genauigkeit von Fremdeinschätzungen, indem Ergebnisse einer Validierungsstudie zu Fremdeinschätzungen sprachlicher Fähigkeiten im Deutschen sowie den Herkunftssprachen Russisch und Türkisch von Peers vorgestellt werden. Grundlage bildete eine Teilstichprobe der Daten des Forschungsprojekts "Mehrsprachigkeitsentwicklung im Zeitverlauf" (MEZ). In einer Art soziometrischem Experiment wählten Schüler(innen) mit russischem (N=289), türkischem (N=440) monolingual deutschem (N=545) Sprachhintergrund bis zu drei Studienteilnehmer(innen) aus, die sie hinsichtlich ihrer sprachlichen Fähigkeiten einschätzten. Mit Hilfe von R wurden in Mehrebenenmodellen die Einschätzungen als Prädiktoren für das Leseverständnis (LGVT 5-12 nach Schneider et al 2007) der eingeschätzten Personen verwendet, wobei die einschätzende Person (Ego) auf Level 2 und die beurteilten Personen auf Ebene 1 verortet wurden. Dieses Verfahren ermöglichte es, Merkmale der einschätzenden Person zu kontrollieren und das Vorhandensein von Cross-Level-Interaktionen angemessen zu prüfen. Ziel war es, neben der Validität der Einschätzungen zu bestimmen, inwieweit unter Berücksichtigung individueller sowie relationaler Merkmale von Ego und der eingeschätzten Person die Güte der Fremdeinschätzungen als Maße für sprachliche Fähigkeiten erhöht werden können.

Hinsichtlich der Fremdeinschätzungen kann gezeigt werden, dass diese sowohl für das Deutsche als auch die Herkunftssprachen Türkisch und Russisch in einem positiven Zusammenhang mit dem Leseverständnis der eingeschätzten Personen steht und in allen Modellen auch unter Kontrolle individueller und relationaler Merkmale einen signifikanten Beitrag zur Aufklärung des Lesetestergebnisses leistet. Die Ergebnisse deuten ferner an, dass Schüler(innen), sofern sie mit der eingeschätzten Person nicht befreundet sind, tendenziell leistungstärkere Personen zur Einschätzung auswählen und leistungstärkere Personen ebenso Personen mit höherer Testleistung einschätzen. Während für die Qualität der Einschätzung des Leseverständnisses im Deutschen eine Freundschaft zwischen der Befragungsperson und der eingeschätzten Person von Bedeutung zu sein scheint, spielt im Türkischen die Lesekompetenz der einschätzenden Person eine wichtige Rolle für die Vorhersage der Lesestellung der Peers. Insgesamt sprechen die Ergebnisse dafür, dass Fremdeinschätzungen genutzt werden können, um die sprachlichen Ressourcen im Peerumfeld sowohl für das Deutsche als auch für eine geteilte Herkunftssprache zumindest ordinal zu bewerten.

### **„Mehrsprachiges reziprokes Lesen“ – Evaluation der Wirksamkeit und empirische Prüfung eines sprachenübergreifenden Zugriffs auf das Leseverstehen**

**Christoph Gantefort, Ina-Maria Maahs**

Universität Köln, Deutschland

In einem Kölner Verbund von Grundschulen, die im Rahmen der Bildungsinitiative BiSS zusammenarbeiten, wird mit dem ‚Mehrsprachigen reziproken Lesen‘ (MRL) derzeit ein Sprachförderkonzept für die Primarstufe entwickelt, erprobt und evaluiert (vgl. Gantefort & Sánchez-Oroquieta 2015). Das Konzept beruht auf den Prinzipien des Translanguaging-Ansatzes (vgl. Garcia 2009), indem wir davon ausgehen, dass mehrsprachig aufwachsende Schülerinnen und Schüler sich Sachverhalte und hierarchiehohe sprachliche Fähigkeiten (wie etwa kognitive und metakognitive Lesestrategien) umso besser aneignen können, je mehr sie dazu ihr gesamtsprachliches Repertoire, also Mittel aus Erst- und Zweitsprache, einbringen können. MRL wird in insgesamt vier Kölner Grundschulen umgesetzt, die alle nach dem Prinzip der koordinierten Alphabetisierung (KOALA) und des Koordinierten Lernens (KOLE) arbeiten: Die Kinder werden nicht nur in Deutsch, sondern auch in ihrer Herkunftssprache Türkisch alphabetisiert und werden im Sachunterricht unter Einbezug des Türkischen unterrichtet (vgl. Bezirksregierung Köln 2014). Damit steht diesen Schülerinnen und Schülern im Vergleich mit einer einsprachigen deutschen Lernumgebung prinzipiell ein breiteres Repertoire an sprachlichen Mitteln zur Verfügung, um sich die Inhalte aus dem Sachunterricht und damit zusammenhängende kommunikative Verfahrensweisen anzueignen. An dieser Stelle setzt die entwickelte Intervention an: Das vielfach bewährte Konzept des reziproken Lesens (vgl. Palincsar & Brown 1984) wurde insofern modifiziert, als die sprachlichen Mittel, die Schülerinnen und Schüler in den Peer-Interaktionen in der kommunikativen Rahmung des reziproken Lesens verwenden, nicht von außen festgelegt sind. Sie werden stattdessen aktiv aufgefordert, im Sinne einer Mikro-Alternierung der Sprachen (vgl. Garcia & Sylvan 2011) nicht nur das Deutsche oder das Türkische, sondern auch innerhalb einzelner Sprachhandlungen beide Sprachen integriert zu gebrauchen, um schwierige Textstellen zu klären, Fragen zum Inhalt des Textes zu stellen, den Inhalt zusammenzufassen und Überlegungen dazu anzustellen, wie der Text weitergehen könnte. Während es in pragmatischer Perspektive im Rahmen der Peer-Interaktionen der Schülerinnen und Schüler mit vergleichbarer

Mehrsprachigkeitskonstellation angemessen ist, die gemeinsam verfügbaren sprachlichen Mittel voll auszuschöpfen, ist es mit zunehmender Reichweite einer Kommunikationssituation erforderlich, sich auf die sprachlichen Mittel einer Sprache zu fokussieren. Eine solche Kontextverschiebung findet auch im Rahmen des MRL statt, indem die Schülerinnen und Schüler schriftsprachliche Produkte erstellen, die entweder im Rahmen des Regelunterrichts in Deutsch oder im Rahmen des herkunftssprachlichen Unterrichts in Türkisch angefertigt und vorgestellt werden. Im Zuge dieses höheren Anforderungsniveaus sollen die Lehrkräfte geeignete sprachliche Hilfen zur Verfügung stellen.

Unser Beitrag adressiert auf der Grundlage der in der wissenschaftlichen Begleitung der Verbundarbeit gewonnenen Daten neben der Frage der Wirksamkeit die Hypothese, das Leseverstehen eine Fähigkeit darstellt, auf die unabhängig von der Einzelsprache, in der Texte präsentiert werden, zugegriffen werden kann (vgl. Francis 2000). Vom ersten bis zum aktuell dritten Schuljahr wurden die Leseflüssigkeit und das Leseverstehen in jeweils Deutsch und Türkisch, die sozioökonomischen Kontextdaten, der familiäre Sprachgebrauch sowie die kognitiven Grundfähigkeiten der Schülerinnen und Schüler bei einer Stichprobe von insgesamt 224 Schülerinnen und Schülern erhoben. In einem quasiexperimentellen Untersuchungs-Kontrollgruppendesign konnten unter Anwendung eines Matchingverfahrens Treatmenteffekte mittlerer Stärke für das Leseverstehen im Deutschen nachgewiesen werden. Regressionsanalysen zeigen darüber hinaus, dass das Leseverstehen im Türkischen durch die Leseflüssigkeit im Türkischen und das Leseverstehen im Deutschen, nicht jedoch von der Leseflüssigkeit im Deutschen erklärt wird.

### **Kulturelle Sozialisierung durch Peers und schulische und psychologische Anpassung – Die medierende Rolle von kultureller Identität bei Jugendlichen mit und ohne Migrationshintergrund**

**Jana Vietze<sup>1</sup>, Linda Juang<sup>2</sup>, Maja Schachner<sup>1</sup>**

<sup>1</sup>Universität Potsdam, Deutschland, <sup>2</sup>Universität Potsdam

In der Jugend wird die Sozialisierung durch Peers immer wichtiger (Brown & Larson, 2009). Zum einen können Peers als Bezugspunkt für schulische Anpassung, wie Schulleistung und Motivation, als auch für psychologische Anpassung, wie Lebenszufriedenheit, dienen (Lynch, Lerner, & Leventhal, 2013; Rogers, Niwa, & Way, 2017). Zum anderen können, besonders für ethnische Minderheiten, Peers auch zum Erhalt und Erlernen von kulturellen Praktiken, Traditionen und Geschichte der ethnischen Gruppe (ethnische Sozialisierung) als auch der nationalen Mehrheitsgesellschaft (nationale Sozialisierung) beitragen (Wang, Benner, & Kim, 2015). Ethnische bzw. nationale Sozialisierung stärken die Identifizierung mit der ethnischen bzw. nationalen Gruppe und sowohl die ethnische als auch die nationale Identität sind wiederum positiv mit psychologischer und soziokultureller Anpassung verknüpft (Berry, Phinney, Sam, & Vedder, 2006). Jedoch fehlen weitestgehend empirische Befunde, wie sich der Zusammenhang von Sozialisierung, Identität und Anpassung bei Jugendlichen ohne Migrationshintergrund im Gegensatz zu Jugendlichen mit Migrationshintergrund gestaltet. In dieser Studie untersuchten wir daher (a) den Zusammenhang von ethnischer bzw. nationaler Sozialisierung durch Peers mit psychologischer und schulischer Anpassung bei Jugendlichen mit und ohne Migrationshintergrund, (b) die medierende Funktion von nationaler Identität (mit und ohne Migrationshintergrund) bzw. ethnischer Identität (mit Migrationshintergrund), und (c) dem unterschiedlichen Einfluss von Sozialisierung, wenn sie mit Peers der eigenen, ethnischen Minderheit (co-ethnische Peers), mit Peers anderer ethnischer Minderheiten (anders-ethnische Peers) oder mit Peers der Mehrheitsgesellschaft (nationale Peers) erlebt wird.

#### **Methoden**

Die Analysen basierten auf 1331 Sekundarschüler\_innen (MAlter= 14,69; 48,3% weiblich) aus kulturell vielfältigen 9. Klassen in Deutschland. Von diesen hatten 1057 Schüler\_innen einen Migrationshintergrund (mind. ein Eltern oder Großelternanteil nicht in Deutschland geboren). In einem Fragebogen wurden die Schüler\_innen gebeten, jeweils eine(n) Freund\_in zu identifizieren, der/die der eigenen Herkunftskultur bzw. einer anderen Kultur angehörte. Anschließend wurden die Teilnehmer\_innen gebeten, Fragen zu ethnischen und nationalen Sozialisierungspraktiken beider Freunde zu beantworten (z.B. „Mein Freund/meine Freundin ermutigt mich, stolz auf Deutschland zu sein“). Sie wurden außerdem gebeten, Fragen zu ihrer ethnischen und nationalen Identität zu beantworten (z.B. „Ich bin froh, Teil meiner Herkunftskultur zu sein“), sowie Fragen zu Lebenszufriedenheit (Ponizovsky, Dimitrova, Schachner, & Van de Schoot, 2013) und subjektiven schulischen Werten, inklusive intrinsischer Motivation, Nützlichkeit und Zielerreichung (SESSW; Steinmayr & Spinath, 2010). Die Hypothesen wurden mithilfe von Strukturgleichungsmodellen in Mplus überprüft.

#### **Ergebnisse & Diskussion**

Gemäß der Erwartungen hingen sowohl die ethnische als auch die nationale Identität positiv mit subjektiven schulischen Werten (SESSW) und Lebenszufriedenheit von Jugendlichen mit und ohne Migrationshintergrund zusammen. Gemäß der Hypothesen, gab es einen indirekten positiven Zusammenhang zwischen ethnischer bzw. nationaler Sozialisierung und SESSW, mediert durch die ethnische Identität bei Jugendlichen mit Migrationshintergrund und durch die nationale Identität bei Jugendlichen ohne Migrationshintergrund. Allerdings war dieser Zusammenhang für beide Gruppen nur signifikant, wenn die Sozialisierung durch einen Peer der eigenen ethnischen Gruppe (co-ethnischer Peer) erfolgte. Der positive Zusammenhang zwischen Sozialisierung und Lebenszufriedenheit bestand nur für Jugendliche mit Migrationshintergrund, und dort nur über die ethnische Sozialisierung durch einen co-ethnischen Peer und ethnische Identität.

Unsere Ergebnisse lassen vermuten, dass besonders in kulturell vielfältigen Umgebungen, wie kulturell heterogenen Schulen, Peers und deren kulturelle Sozialisierung eine Ressource für kulturelle Identität, psychologische und schulische Anpassung darstellen können. Peers der eigenen ethnischen Gruppe schienen hier sowohl für Jugendliche mit als auch ohne Migrationshintergrund die wichtigsten Sozialisierungsagenten zu sein, da sie als Orientierung für die eigene Kultur und Sprache dienen können und einen Austausch über eigene kulturelle Erfahrungen ermöglichen (Yip, 2005). In zukünftigen Studien gilt es daher herauszufinden, wie sich kulturelle Sozialisierungsprozesse zwischen co- und anders-ethnischen Peers inhaltlich und formal unterscheiden, und wie weitere Risikofaktoren, wie z.B. wahrgenommene Diskriminierung oder die kulturelle Zusammensetzung der Schulklasse, den Zusammenhang von Peersozialisation, schulischer und psychologischer Anpassung bei Jugendlichen mit und ohne Migrationshintergrund beeinflussen.

## E17

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H114

### Diagnostische Kompetenzen – von der Urteilsgenauigkeit zu Urteilsprozessen

Chair(s): **Timo Leuders** (Pädagogische Hochschule Freiburg, Deutschland)

DiskutantIn(nen): **Jörg Wittwer** (Albert-Ludwigs-Universität Freiburg)

Dem diagnostischen Handeln von Lehrkräften wird eine Schlüsselfunktion für die Unterrichtsqualität und die Kompetenzentwicklung von Schülerinnen und Schülern zugesprochen (Artelt & Gräsel, 2009; Weinert, 1998). Diagnostische Fähigkeiten werden daher seit vielen Jahren untersucht, vor allem im Rahmen des so genannten Urteilsgenauigkeitsparadigmas (Schrader & Helmke, 1988). Für die empirisch festgestellte Multidimensionalität des so operationalisierten Konstruktes (Spinath, 2005) finden sich allerdings bislang kaum befriedigende theoretische Erklärungen (Südkamp, Kaiser & Möller, 2012). Als Erklärungsansätze sind insbesondere die Spezifität der Diagnosesituationen, die Eigenschaften der diagnostizierenden Lehrkräfte und die Dynamik der Urteilsprozesse von Interesse (Schrader, 2011). Diese sich hiermit abzeichnende komplexe und situative Sicht auf diagnostisches Urteilen, legt es nahe, für weitere Untersuchungen auf den Kompetenzbegriff zurückzugreifen (Weinert, 2001).

Ein deskriptives Modell zur Beschreibung von Kompetenzen, insbesondere von Lehrkräften, wird von Blömeke, Gustafson und Shavelson (2015) vorgeschlagen. Blömeke et al. (2015) beschreiben als Rahmen für die Kompetenzforschung ein Modell der „Kompetenz als Kontinuum“, welches eine Verbindung zwischen latenten Personeneigenschaften, situationsspezifischen Fähigkeiten und Performanz als Komponenten herstellt. Neuere Untersuchungen (z. B. Heinrichs, 2015; Gruppe NeDiKo, Herppich et al. 2017; Leuders, Leuders & Philipp, i.Dr.) betrachten insbesondere diagnostische Kompetenzen in einem solchen Rahmen und unterscheiden dabei (1) diagnoserelevante Dispositionen (z.B. diagnostisches Wissen, diagnostische Einstellungen), (2) diagnostische Urteilsprozesse (z. B. Wahrnehmen, Interpretieren, Entscheiden) und (3) diagnostisches Handeln.

Im Symposium werden aktuelle Studien vorgestellt und diskutiert, die neue Erkenntnisse zum Konstrukt der diagnostischen Kompetenz erwarten lassen. Der erste Beitrag greift Daten der COACTIV-Studie auf und zeigt, dass die betrachteten (Teil-)Konstrukte diagnostischer Kompetenz nur schwach und unsystematisch zusammenhängen. Zudem werden überraschende Befunde zur (negativen) prädiktiven Validität bestimmter diagnostischer Fähigkeiten für die Mathematikleistung diskutiert. Im zweiten Beitrag wird anhand zweier exemplarischer Studien analysiert, wie Beziehungen zwischen Aspekten des NeDiKo-Modells untersucht und die Trainierbarkeit diagnostischer Kompetenz geprüft werden kann. Im dritten Beitrag wird das integrative Dia-Kom-Modell zur Prozessebene diagnostischer Kompetenz vorgestellt und anhand einer Interventionsstudie gezeigt, wie bisherige theoretische Annahmen integriert untersucht werden können.

Nach einem einführenden Überblick über den Diskussionsstand zu diagnostischen Kompetenzen (10 min) folgen drei Vorträge (jeweils 25 min, inkl. Diskussion), die verschiedene neuere Perspektiven anreißen. Die abschließende übergreifende Diskussion (10 min Diskutant, 10 min Plenum) soll eine Einschätzung der Desiderate und Entwicklungen im Forschungsgebiet ermöglichen.

#### Beiträge des Symposiums

### Facetten diagnostischer Fähigkeiten in der COACTIV-Studie: Neue Analysen zu Zusammenhängen und zur prädiktiven Validität für Schülerleistungen

**Karin Binder, Stefan Kraus, Sven Hilbert**  
Universität Regensburg,

#### Theoretischer Hintergrund

In der COACTIV-Studie wurden neben dem Fachwissen und fachdidaktischen Wissen auch verschiedene Facetten diagnostischer Fähigkeiten von Mathematiklehrkräften fokussiert (Kunter et al., 2013). Professionelle Kompetenz von Lehrkräften wird in dieser Studie allgemein auf der Basis von vier übergeordneten Kompetenzaspekten (Professionswissen, professionelle Überzeugungen, motivationale Orientierungen und selbstregulative Fähigkeiten) modelliert. Diagnostische Fähigkeiten werden dabei dem Professionswissen und dort konkret dem Überschneidungsbereich von fachdidaktischem und pädagogischem Wissen zugeordnet. Im vorliegenden Beitrag stellen wir konzeptuelle und empirische Ansätze zur Untersuchung „diagnostischer Kompetenz“ in der COACTIV-Studie vor (Binder et al., im Druck; Anders et al., 2010 bzw. Brunner et al., 2013).

#### Fragestellung

Im vorliegenden Beitrag werden vor allem drei Fragestellungen fokussiert: 1. Zusammenhänge verschiedener Facetten diagnostischer Kompetenz, 2. Schulformunterschiede, und 3. die prädiktive Validität der Facetten diagnostischer Kompetenz für Unterrichtsqualität und den Leistungszuwachs von Schülerinnen und Schülern.

#### Methoden

Es werden acht verschiedene Operationalisierungen diagnostischer Fähigkeiten betrachtet (die anhand der PISA-Ergebnisse der Schüler der COACTIV-Lehrer bestimmt werden können): 1. Einschätzung des Leistungsniveaus der Gesamtklasse bezogen auf das mittlere Leistungsniveau einer PISA-Klasse, 2. Einschätzung der Leistungsstreuung, 3. Einschätzung des Anteils der Schüler, die im unteren Leistungsdrittel liegen, 4. Einschätzung des Anteils der Schüler, die im oberen Leistungsdrittel liegen, 5. Einschätzung der Leistungsbereitschaft der Klasse, 6. Einschätzung des Anteils der Schüler, die vier spezielle PISA-Aufgaben richtig lösen. Zwei weitere untersuchte diagnostische Fähigkeiten von Lehrkräften beziehen sich nicht auf die Gesamtklasse, sondern auf individuelle Schüler: 7. Die Lehrkräfte sollten die Performanz (falsch/richtig) von sieben zufällig ausgewählten Schülern bei je zwei PISA-Aufgaben einschätzen, 8. Bildung einer Rangfolge aus diesen sieben Schülern bezüglich ihres Abschneidens im PISA-Gesamttest (Binder et al., im Druck; Anders et al., 2010; Brunner et al., 2013).

Neben diesen acht Aspekten diagnostischer Fähigkeiten untersuchen wir Zusammenhänge zu drei weiteren bei COACTIV erhobenen Konstrukten, die theoretisch eng mit diagnostischen Fähigkeiten verknüpft sein sollten, die in bisherigen Analysen aber noch nicht unter diesem Blickwinkel betrachtet wurden (Binder et al., im Druck): 9) Die Fähigkeit des schnellen Beurteilens von Schülerantworten im Hinblick auf deren fachliche Richtigkeit (Reaktionszeittest; siehe Krauss & Brunner, 2011). 10) Das fachdidaktische Wissen von Mathematiklehrkräften (PCK; Krauss et al., 2013). 11) Diagnostische Fähigkeiten in Bezug auf soziale Aspekte.

## Ergebnisse

Die bivariaten Korrelationen (auf Basis von  $N = 255$  Mathematiklehrkräften verschiedener Sekundarschulformen, die 2003 eine PISA-Klasse unterrichteten) zwischen diesen elf Fähigkeiten deuten auf unsystematische und lediglich mäßige Zusammenhänge hin und bestätigen somit bisherige Befunde zur Heterogenität diagnostischer Fähigkeiten. Unsere neu durchgeführten Analysen zu Schulformunterschieden (Binder et al., im Druck) zeigten allerdings (erwartungskonform), dass gymnasiale Lehrkräfte bei einem Großteil der erfassten diagnostischen Fähigkeiten signifikant besser als nicht-gymnasiale Lehrkräfte abschnitten. Einzig bei der diagnostischen Fähigkeit bezüglich sozialer Belange waren die nicht-gymnasialen Lehrkräfte den gymnasialen überlegen, was ebenfalls konform zu bisherigen COACTIV-Ergebnissen ist.

Ferner werden Strukturgleichungsmodelle vorgestellt, in denen die prädiktive Validität diagnostischer Fähigkeiten für die Mathematikleistung von Schülern untersucht wird („Black-Box“ Modelle sowie „Mediationsmodelle“ über Aspekte der Unterrichtsqualität). Überraschenderweise scheint die Fähigkeit, das Leistungsniveau der Klasse adäquat einschätzen zu können (Fähigkeit 1) sogar zu verhindern, dass kognitiv fordernde Mathematikaufgaben im Unterricht eingesetzt werden. Lediglich Fähigkeit 6 (siehe oben) hat einen nachweislichen Einfluss auf den Leistungsfortschritt der Schüler in Mathematik.

## Diskussion

Zusammengefasst bestätigen unsere Analysen die Verwendung des Begriffs „diagnostische Fähigkeiten“ (statt diagnostischer Kompetenz) von Brunner et al. (2013), da wir lediglich kleine und unsystematische Zusammenhänge zwischen den betrachteten (Teil-)Konstrukten vorfinden können. Unsere Ergebnisse stehen damit im Einklang mit Spinath (2005) und Schrader (1989), die ebenfalls keine oder nur schwache Zusammenhänge zwischen verschiedenen Indikatoren diagnostischer Kompetenz gefunden haben. Der provozierende Befund, dass mangelnde Einschätzung der „wahren“ kognitiven Leistungsfähigkeit einer Klasse sogar zu kognitiv aktivierenderem Unterricht („Mut zu anspruchsvolleren Aufgaben aufgrund einer Fehleinschätzung“) und somit zu besseren Schülerleistungen führen kann, soll ebenfalls diskutiert werden.

## **Diagnostische Kompetenz von Lehrkräften: Das NeDiKo-Modell, illustriert durch empirische Studien**

**Stephanie Herppich, Inga Glogger-Frey**  
Albert-Ludwigs-Universität Freiburg,

### Theoretischer Hintergrund

Diagnostische Kompetenz ist grundlegend für verschiedenstes professionelles Handeln von Lehrkräften. Das Modell des Netzwerks Diagnostische Kompetenz (NeDiKo) konzeptualisiert sie dabei als erlernbare und messbare kognitive Voraussetzung für die Beschreibung, Erklärung, Vorhersage und Förderung diagnostischen Handelns (Herppich, Praetorius et al., 2017). Anders als in früheren Modellen werden daher nicht nur einzelne Aspekte von Kompetenz und Handeln betrachtet (z.B. Diagnoseakkuratheit, Schrader, 1989), sondern Wissen, Überzeugungen und Motivation sowie Diagnoseprozesse, Diagnosehandlungen und deren Ergebnisse in unterschiedlichen Situationen, in Beziehung zueinander gesetzt.

Zur Entwicklung des Modells wurde pädagogische Diagnostik definiert als „Prozess des Einschätzens von Lernenden im Hinblick auf lernrelevante Merkmale mit dem Ziel, pädagogische Entscheidungen vorzubereiten“ (Herppich, Praetorius et al., 2017, S. 3; vgl. Leutner, 2010). Diagnostische Kompetenz wurde definiert als „kognitive Leistungsdisposition . . . , die es Lehrkräften, relativ konsistent und relativ stabil pädagogische–diagnostische ermöglicht Anforderungen in verschiedenen pädagogischen Handlungssituationen zu meistern“ (Herppich, Praetorius et al., 2017, S. 7; vgl. Klieme & Leutner, 2006). Im hieraus entwickelten Modell führt eine Lehrperson auf Basis ihrer kognitiven Leistungsdisposition, deren Wirksamwerden durch sonstige Leistungsdispositionen gefiltert (z.B. moderiert) wird, in verschiedenen pädagogischen Handlungssituationen pädagogische Diagnostik durch, an deren Ende eine Diagnose steht. Komponenten der kognitiven Leistungsdisposition können Aspekte pädagogischen, fachdidaktischen und Fachwissens sein (z.B. Wissen über lernrelevante Merkmale). Sonstige Leistungsdispositionen sind beispielsweise Motivation. Das Wirksamwerden der kognitiven Leistungsdisposition ist weiterhin abhängig von den Anforderungen der Situationen, die sich beispielsweise nach Zweck, Planbarkeit und Konsequenz für Lernende unterscheiden. Auf Basis der Diagnose trifft die Lehrperson eine pädagogische Entscheidung. Diese ist das Ziel der Diagnostik, jedoch nicht Teil von ihr. Diagnostik kann in Abhängigkeit von Moderatoren kontrolliert und kognitiv aufwändig oder automatisiert und mit geringerem kognitiven Aufwand ablaufen (Böhmer, Gräsel, Krolak-Schwerdt, Hörstermann, & Glock, 2015). Diagnostische Handlungen und Diagnosen sind als Performanz messbar, also quantifizierbar.

### Fragestellung

Mit zwei exemplarisch ausgewählten Studien wird der Frage nachgegangen, wie Aspekte des Modells und ihre Zusammenhänge empirisch untersucht werden können.

### Methode

In Studie 1 (Herppich, Südkamp et al., 2017) wurden konzeptuelles Wissen über pädagogische Diagnostik (Komponente der kognitiven Leistungsdisposition) und der Erfolg bei der Auswahl fiktiver Schultests zum Treffen einer Förderentscheidung bei 70 Lehramtsstudierenden und 28 Referendar\_innen in Beziehung zueinander gesetzt. Somit wurde die Auswahl von Diagnosemethoden als Teil des diagnostischen Prozesses untersucht.

In Studie 2 (Glogger-Frey, Deutscher, & Renkl, 2017) wurden konzeptuelles Wissen über Lernstrategien und Erkennen von Lernstrategien in authentischen Schülerprodukten vor und nach einem kurzen Training zu Wissen und Diagnostik bei 47 Lehramtsstudierenden in Beziehung zueinander gesetzt. Somit wurde der Zusammenhang zwischen einer Komponente der kognitiven Leistungsdisposition und der Diagnose untersucht. Zudem wurde der Zuwachs in der Diagnoseleistung (Lernstrategien erkennen) erfasst.

### Ergebnisse

Studie 1 zeigte, dass Studierende sowie Referendar\_innen etwa sechs von elf Wissenstestsitems richtig beantworteten und etwa zwei von drei Tests richtig auswählten. Referendarinnen wählten signifikant ( $d=0.88$ ) mehr Tests richtig aus. Wissen und Testauswahl hingen signifikant für die Studierenden ( $r=.31$ ) aber nicht für die Referendar\_innen ( $r=.06$ ) zusammen. Pädagogisch-diagnostisches Wissen scheint bei zunehmender Praxis kein Prädiktor für die Auswahl diagnostischer Methoden mehr zu sein.

In Studie 2 sagten unter Kontrolle der Lehrerfahrung konzeptuelles Wissen ( $b^*=.32$ ), aber auch zunächst als misskonzepthaft zu sehendes Wissen ( $b^*=.28$ ) signifikant die Diagnoseleistung vorher (Gesamtmodell  $R^2=.25$ ). Der Leistungszuwachs war mittel bis hoch ( $d=0.77$ ). Studie 2 belegt die Trainierbarkeit von Komponenten diagnostischer Kompetenz.

### Diskussion

Beide Studien zeigen, wie Aspekte des Modells, hier potenzielle Komponenten der kognitiven Leistungsdisposition (Wissen) und Schritte im diagnostischen Prozess, mit korrelativen und Interventions-Designs zueinander in Beziehung gesetzt werden können, um das Modell zu validieren. Zukünftige Forschung kann durch ähnliche Herangehensweisen weitere Beziehungen prüfen und so das Modell nutzen, um diagnostisches Handeln umfassend zu beschreiben, zu erklären, vorherzusagen und zu fördern.

## **Analysen von diagnostischen Urteilsprozessen nach dem integrativen DiaKom-Modell**

**Katharina Loibl<sup>1</sup>, Timo Leuders<sup>1</sup>, Tobias Dörfler<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Freiburg, <sup>2</sup>Pädagogische Hochschule Heidelberg

### Theoretischer Hintergrund

Bestehende Forschungsansätze nutzen ganz verschiedene theoretische Modelle zur Erklärung und Untersuchung diagnostischer Kompetenz (vgl. Südkamp & Praetorius, 2017). Es besteht daher Bedarf nach einer theoretischen Rahmung, die es ermöglicht, künftige Forschung konvergenter zu gestalten und bestehende Befunde zu integrieren. Gleichzeitig fokussiert die bestehende Forschung insbesondere auf die Urteilsgenauigkeit (Hoge & Coladarci, 1989; Südkamp, Kaiser & Möller, 2012), wohingegen bislang eine direkte Verbindung zu den diagnostischen Urteilsprozessen fehlt. Im Rahmen des Promotionskollegs „DiaKom“ (Diagnostische Kompetenzen von Lehrkräften) wurde daher ein integratives Modell entwickelt, das die Prozessebene fokussiert. Dabei verbindet das DiaKom-Modell Komponenten aus folgenden Ansätzen: das Linsenmodell (Brunswik, 1956), welches die Urteilsgenauigkeit mit der Signalvalidität (Wie relevant ist ein Hinweisreiz für die Diagnose?) und der Signalnutzung (Welche Hinweisreize werden beachtet?) in Verbindung bringt; das Kompetenzmodells nach Blömeke, Gustafsson und Shavelson (2015), welches auf der Prozessebene Wahrnehmung, Interpretation und Entscheidung unterscheidet, jedoch keine Hinweise enthält, wie diese Teilprozesse (empirisch) zu trennen sind; und Dual Process-Theorien (z.B. Croskerry, 2009), die zwischen heuristischen und analytischen Urteilsprozessen unterscheiden. So hängen die Art und Menge an Hinweisreizen, die beobachtet und in die Urteilsbildung einbezogen werden (d. h. die Signalnutzung), mit der Art der Diagnoseprozesse zusammen: Heuristische Prozesse werden ausgelöst durch eine Mustererkennung, die auf wenigen Hinweisreizen basiert. Analytische Prozesse zeichnen sich dadurch aus, dass eine Vielzahl von Hinweisreizen beachtet wird. Dabei besteht die Annahme, dass heuristische Urteilsprozesse schneller vollzogen werden und fehleranfälliger sind. Dahingegen sollten analytische Urteilsprozesse, die unter Berücksichtigung fachdidaktischen Wissens erfolgen, bei Urteilen zu Schülerlösungen die Fehlkonzepte beinhalten zu einer höheren Urteilsgenauigkeit führen.

### Fragestellungen

Um exemplarisch aufzuzeigen, wie die im DiaKom-Modell enthaltenen Annahmen zu diagnostischen Urteilsprozessen überprüft werden können, wird eine Interventionsstudie vorgestellt, die folgende Forschungsfragen untersucht: 1) Lassen sich analytische vs. heuristische Prozesse gezielt anregen? 2) Zeigen sich diese Prozessunterschiede sowohl in der Bearbeitungszeit als auch in der Anzahl beachteter Hinweisreize ?

### Methode

Teilnehmer waren 35 Mathematiklehramtsstudierende, die randomisiert zwei Gruppen zugewiesen wurden. Studierende der Intuitionsbedingung erhielten einen Text zur Bedeutung von Intuition beim schnellen diagnostischen Handeln im Lehrkontext. Studierende der PCK-Bedingung (PCK=pädagogical content knowledge) erhielten einen fachdidaktischen Text zu typischen Fehlkonzepten beim Bruchvergleich. Anschließend arbeiteten beide Gruppen in einer computergestützten Lernumgebung, in der sie vier SchülerInnen auf Basis der Lösungen zu zehn Bruchvergleichsaufgaben beurteilen sollten. Das diagnostische Urteil bestand in der Vorhersage, wie die SchülerInnen zwei weitere Aufgaben lösen würden. Dabei war jeder Fall als Repräsentant für einen typischen Schülerfehler gewählt, der bei den Bruchvergleichsaufgaben teils zu richtigen, teils zu falschen Antworten führt. Zur Verfügung stehende Hinweisreize waren Geschlecht, Größe der Ziffern, Anzahl der Fehler, subjektive Schwierigkeit der Aufgaben, Art der Brüche (echte vs. unechte Brüche), Vergleich der Zähler und Vergleich der Nenner. Nur die beiden letztgenannten Hinweisreize stehen mit den typischen Schülerfehlern in Beziehung und waren für die Diagnosegüte relevant. Die Lernumgebung erfasste die Verweildauer auf jeder Seite. Im Anschluss wurden die Studierenden zu den beachteten Hinweisreizen befragt.

### Ergebnisse

Die Gruppen unterschieden sich signifikant in der Urteilsgenauigkeit ( $t[33]=2.837$ ,  $p=.008$ ,  $d=0.979$ ) zugunsten der PCK-Bedingung. Entsprechend unserer Hypothesen hatten die Studierenden der PCK-Bedingung signifikant längere Bearbeitungszeiten als die Studierenden der Intuitionsbedingung ( $t[33]=2.762$ ,  $p=.009$ ,  $d=0.953$ ) und beachteten signifikant mehr Hinweisreize ( $t[15.82]=2.202$ ,  $p=.043$ ,  $d=0.884$ ).

### Diskussion

Zusammengenommen sprechen diese Ergebnisse dafür, dass durch die Manipulation in der PCK-Bedingung ein analytischer Diagnoseprozess angeregt wurde. Dieser zeigte sich in der Bearbeitungszeit (vgl. Dual Process-Theorien) und der Signalnutzung (vgl. Linsenmodell). Im Kontext von Schülerlösungen mit Fehlkonzepten und der Verfügbarkeit fachdidaktischen Wissens führte dieser analytische Prozess zu einer höheren Urteilsgenauigkeit. Die Studie zeigt exemplarisch auf, wie sich verschiedene theoretische Annahmen im Rahmen des integrativen DiaKom-Modells verknüpfen lassen. Im Vortrag werden die Grenzen der Studie kritisch diskutiert und es wird aufgezeigt, wie das DiaKom-Modell weitere Forschung informieren kann.

## E18

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H102

### Vergleichbarkeit von Bildungsabschlüssen durch verschiedene Standardisierungsstrategien bei Abschlussprüfungen – Empirische Befunde aus Deutschland und der Schweiz

*Chair(s):* **Pauline Schröter** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB)), **Lars Hoffmann** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB)), **Petra Stanat** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB))

*DiskutantIn(nen):* **Marko Neumann** (Deutsches Institut für Internationale Pädagogische Forschung (DIPF))

Abschlussprüfungen erheben den Anspruch, die von Prüflingen erzielten Leistungen so zu erfassen, dass die resultierenden Ergebnisse vergleichbar sind. Dabei kommen unterschiedliche Strategien zur Anwendung, die in ihrem Standardisierungsgrad variieren (Kühn, 2012). Eine Strategie besteht in der Durchführung zentraler Prüfungen, bei denen alle Prüflinge dieselben Aufgaben bearbeiten. Andere Strategien zielen darauf ab, die Vergleichbarkeit von Anforderungen durch die gemeinsame Entwicklung von Prüfungsaufgaben zu gewährleisten. Bei beiden Ansätzen erfolgt die Entwicklung von Prüfungsaufgaben in einem komplexen Spannungsfeld, das durch administrative Vorgaben einerseits und die Zusammenarbeit von Akteuren aus unterschiedlichen Bereichen (z.B. Bildungsverwaltung, Fachdidaktik und schulische Praxis) andererseits gekennzeichnet ist. Die Qualität der dabei entstehenden Aufgaben sowie der Entwicklungsprozesse selbst waren im deutschsprachigen Raum bislang jedoch kaum Gegenstand empirischer Forschungsarbeiten (Maag Merki, 2016), obwohl sie für die angestrebte Vergleichbarkeit von Bildungsabschlüssen von zentraler Bedeutung sind.

Im Symposium werden Studien präsentiert, die diese unterschiedlichen Ansätze untersucht haben. Dabei werden verschiedene Aspekte der Vergleichbarkeit in den Blick genommen, die von Prozessen der Aufgabenentwicklung über Anforderungen von Aufgabenstellungen bis hin zu Vorgaben für die Korrektur und Bewertung von Schülerantworten reichen. Alle drei Beiträge beschreiben Verfahren zur Sicherung von Qualitätsstandards bei der Entwicklung, Nutzung und Bewertung von Prüfungsaufgaben.

Im ersten Beitrag werden von Svenja Kühn die Ergebnisse aus zwei Studien vorgestellt, die die Objektivität von Korrektur- und Bewertungsvorgaben (KBV) aus zentralen Prüfungsverfahren zum Erwerb des MSA untersuchen. Während die erste Studie darauf abzielt, die KBV der Bundesländer hinsichtlich ihres Potenzials für die objektive Beurteilung von Schülerleistungen zu typisieren, untersucht die zweite Studie, wie Lehrkräfte die KBV nutzen.

Im zweiten Beitrag werden von Lars Hoffmann und Pauline Schröter erste Ergebnisse der Begleitforschung zum Projekt „Gemeinsame Abituraufgabenpools der Länder“ vorgestellt. Dieses Projekt zielt darauf ab, den Bundesländern Aufgaben für den Einsatz in der Abiturprüfung zur Verfügung zu stellen, die den Anforderungen der Bildungsstandards für die Allgemeine Hochschulreife entsprechen. Nach einer Darstellung des Vorgehens, das dabei zur Anwendung kommt, wird anhand von Befunden zu unterschiedlichen Evaluationsbereichen dargestellt, wie sich die Aufgaben im ersten Prüfungsjahr 2017 bewährt haben.

Im dritten Beitrag gehen die Autorinnen Monika Holmeier, Carmen Hirt und Katharina Maag Merki der Frage nach, inwieweit unterschiedliche Verfahren des „Gemeinsamen Prüfens“, bei dem mehrere Lehrkräfte Abiturprüfungsaufgaben gemeinsam entwickeln, in der schulischen Praxis implementiert werden können. Dabei werden Stärken und Schwächen verschiedener Verfahren aufgezeigt und im Hinblick auf die Einführung standardisierter Prüfungsformen reflektiert.

Im Anschluss diskutiert Marko Neumann die Beiträge unter anderem mit Bezug auf die Frage, welche Standardisierungsstrategien dazu geeignet sind, um das Ziel der Vergleichbarkeit von Bildungsabschlüssen zu gewährleisten und gleichzeitig eine gewisse Flexibilität bei der inhaltlichen Gestaltung von Abschlussprüfungen zu erhalten.

*Beiträge des Symposiums*

### Mehr Vergleichbarkeit durch zentrale Abschlussprüfungen? Möglichkeiten, Herausforderungen und Grenzen einheitlicher Korrektur- und Bewertungsvorgaben

**Svenja Kühn**

Universität Koblenz-Landau

Theoretischer Hintergrund

Zentrale Prüfungen haben das Ziel, einheitliche Leistungsanforderungen zu sichern und Schulabschlüsse vergleichbarer zu machen. Landesweit einheitlichen Korrektur- und Bewertungsvorgaben (KBV) wird dabei das Potenzial zugesprochen, die Objektivität der Leistungsbeurteilung sowie der Notenvergabe steigern und so die Vergleichbarkeit von Abschlüssen erhöhen zu können (vgl. z.B. Klein et al., 2014). Die Qualität der KBV ist demnach entscheidend für eine objektive, transparente und vergleichbare Leistungsbeurteilung (vgl. Köster, 2010). Gleichwohl ist im Rekurs auf schulische Governance-Ansätze (z.B. Altrichter, Brüsemeister & Wissinger, 2007; Altrichter & Maag Merki, 2016) davon auszugehen, dass die Nutzung der KBV nicht nur durch deren konkrete Ausgestaltung bedingt wird, sondern durch strukturelle (z.B. Ausgestaltung des Korrekturverfahrens, bspw. keine, schulinterne oder -externe Zweitkorrektur, vgl. Kühn, 2013) und organisational-situative Handlungsbedingungen (z.B. einzelschulische Bewertungsstandards, vgl. Terhart, 2014, S. 890ff.) sowie vielfältige „Wahrnehmungs-, Verarbeitungs- und Übersetzungsleistungen“ (Altrichter & Maag Merki, 2016, S. 5) der beurteilenden Lehrkräfte, die Fend (2006) als Rekontextualisierungsprozesse bezeichnet, multifaktoriell beeinflusst wird. Im Forschungsfeld ‚zentrale Abschlussprüfungen‘ sind diese Aspekte aber bislang kaum untersucht worden.

Fragestellung

Im Beitrag werden die Ergebnisse aus zwei Studien vorgestellt, die sich auf zentrale Prüfungen zum Erwerb des Mittleren Schulabschlusses beziehen. Im ersten Schritt wurden die KBV der Prüfungsfächer Deutsch und Mathematik aus allen deutschen Ländern kategoriengeleitet hinsichtlich verschiedener Aspekte von Objektivität analysiert, wobei folgende Forschungsfragen handlungsleitend waren:

- Welche Merkmale charakterisieren die KBV? Inwieweit lassen sich Ländertypen identifizieren, deren KBV ein vergleichbares Potenzial für die objektive Beurteilung von Schülerleistungen aufweisen?

Im zweiten Schritt wurden Lehrkräfte zur Wahrnehmung und Nutzung der KBV im Kontext zentraler Prüfungen befragt, wobei folgende Forschungsfragen handlungsleitend waren:

- In welchem Ausmaß nutzen Lehrkräfte die KBV und welche Faktoren beeinflussen die Nutzung?



## Methode

Zunächst wurden die KBV aus allen 15 Ländern mit zentralen Abschlussprüfungen über einen Zeitraum von 5 Jahren (2011-2015) auf Basis von zwei fachspezifischen, deduktiv-induktiv entwickelten Kategoriensystemen analysiert. Insgesamt wurden im Rahmen des hier dokumentierten Projekts N=140 KBV mit N=4.863 einzelnen Lösungsdarstellungen zu den dazugehörigen Prüfungsaufgaben untersucht. Die Qualität der Kategoriensysteme wurde über einen mehrstufigen Prozess der Erprobung, Diskussion, Ausdifferenzierung und Reduktion sichergestellt. Auf die Deskription von Gestaltungsmerkmalen folgte eine Typenbildung in Anlehnung an Kuckartz (2012), um Ländertypen zu identifizieren, deren KBV ein vergleichbares Potenzial für die objektive Beurteilung von Schülerleistungen zugesprochen werden kann.

Auf Basis der empirisch ermittelten Typologie wurden anschließend in drei bzw. vier ausgewählten Ländern (s.u.) leitfadengestützte Interviews mit Lehrkräften der Fächer Deutsch und Mathematik durchgeführt (N=21), transkribiert und inhaltsanalytisch ausgewertet.

## Ergebnisse

In der Gesamtschau zeigt sich ein breites Spektrum an Gestaltungsvarianten der KBV, insbesondere hinsichtlich des Abstraktions- und Detailliertheitsgrad der Beurteilungskriterien, das sich von konkreten Musterlösungen mit detailreicher Ausdifferenzierung in verschiedene Teilaspekte bzw. Handlungsschritte bis hin zu abstrakten Kriterien (z.B. „Zusammenfassung des Hauptteils“) erstreckt. Die ermittelten Typen lassen sich auf einem Kontinuum anordnen: Auf der einen Seite gibt es einen Typ, dem auf Grund seines analytischen Merkmalsprofils ein großes Potenzial für die objektive Beurteilung von Schülerleistungen und eine transparente Notenvergabe zugesprochen werden kann. Auf der anderen Seite gibt es in beiden Fächern auch einen Typ, der im Rekurs auf theoretische Überlegungen zur Objektivität kaum das Potenzial aufweist, Objektivität zu befördern. Für beide Fächer konnten drei Ländertypen herausgearbeitet werden: in jeweils einem dieser Länder wurden pro Fach je drei Lehrkräfte zur Wahrnehmung und Nutzung der KBV befragt. Die Befunde der explorativen Interviewstudie deuten auf eine differenzielle Wahrnehmung der KBV – auch innerhalb desselben Bundeslandes – hin. Hinsichtlich der Nutzung stellen die Einstellung zu zentralen Prüfungen und den KBV sowie die Berufserfahrung zentrale Einflussfaktoren dar. Von besonderer Relevanz sind zudem die vielfältigen Abstimmungsprozesse zwischen den an der Leistungsbeurteilung beteiligten Lehrkräften, die bundeslandspezifische und einzelschulische Besonderheiten aufweisen.

## **Gemeinsame Abituraufgabenpools der Länder – Evaluation des erstmaligen Einsatzes von länderübergreifenden Prüfungsaufgaben im Abitur**

**Lars Hoffmann, Pauline Schröter**

Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

### Theoretischer Hintergrund

Im Jahr 2012 verabschiedete die Kultusministerkonferenz (KMK) für die Fächer Deutsch, Englisch, Französisch und Mathematik die Bildungsstandards für die Allgemeine Hochschulreife (Pant, & Stanat, 2013). Diese spezifizieren zum einen, welche Kenntnisse, Fähigkeiten und Fertigkeiten Schülerinnen und Schüler bis zum Ende der Sekundarstufe II entwickelt haben sollen und beinhalten zum anderen Vorgaben zur Gestaltung der Abiturprüfungen, die für alle Länder verbindlich sind (Stanat, Becker-Mrotzek, Blum, & Tesch, 2016). Die Bildungsstandards für die Allgemeine Hochschulreife lösten die bis dahin geltenden Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA) ab. Neben der Förderung eines stärker kompetenzorientierten Unterrichts dient die Einführung von Bildungsstandards in der Sekundarstufe II insbesondere dazu, die Anforderungen für den Unterricht und die Abschlussprüfung in der gymnasialen Oberstufe transparent zu machen, diese bundesweit anzugleichen und damit die Durchlässigkeit vom sekundären ins tertiäre Bildungssystem zu sichern.

Eine zentrale Maßnahme zur Implementation der Bildungsstandards für die Allgemeine Hochschulreife bildet der Aufbau von gemeinsamen Abituraufgabenpools, der vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) koordiniert und in Zusammenarbeit mit den Bundesländern gestaltet wird. Diese Aufgabenpools zielen darauf ab, die Vergleichbarkeit von Anforderungen zwischen den Ländern erhöhen, ohne die Abiturprüfungen bundesweit vollständig zu standardisieren.

Auf der Grundlage der Bildungsstandards für die Allgemeine Hochschulreife und weiterer Vereinbarungen für die ländergemeinsame Entwicklung von Aufgaben veröffentlichte das IQB im Jahr 2015 zunächst eine Sammlung von Abiturprüfungsaufgaben, die der Orientierung in Bezug auf die Gestaltung und die zu erwartenden Anforderungen von Abituraufgaben dient. Im Prüfungsjahr 2017 wurden erstmalig Aufgaben aus den Abituraufgabenpools in den Abiturprüfungen der Länder eingesetzt.

### Fragestellung

Um die Qualität der gemeinsam entwickelten Aufgaben empirisch zu untersuchen, zu sichern und weiter zu optimieren, wird im Projekt wissenschaftliche Begleitforschung durchgeführt. In diesem Rahmen erfolgte im Prüfungsjahr 2017 parallel zum erstmaligen Einsatz von Poolaufgaben eine bundesweite Evaluation der Abiturprüfungsaufgaben. In diesem Beitrag soll zunächst dargestellt werden, wie bei der Aufgabenentwicklung vorgegangen wird, um zu gewährleisten, dass die Anforderungen den Bildungsstandards entsprechen. Vor diesem Hintergrund werden anschließend zentrale Befunde der Evaluation dieser Aufgaben zu präsentieren. Der Schwerpunkt liegt dabei auf den Evaluationsfragen,

- wie häufig die Aufgaben aus den Pools von den Schülerinnen und Schülern ausgewählt wurden,
- wie sich die von den Schülerinnen und Schülern erzielten Ergebnisse bei den Aufgaben aus den Pools von den landeseigenen Aufgaben unterscheiden,
- wie trennscharf die Aufgaben aus den Pools sind,
- wie die Lehrkräfte den Anspruch der Aufgaben aus den Pools und die Nützlichkeit ihrer Erwartungshorizonte einschätzen.

### Methode

Bei der formativ angelegten Evaluation wurden in den Ländern Daten von 20 allgemeinbildenden Schulen mit gymnasialer Oberstufe erfasst. Die Informationen zu den verschiedenen Evaluationsbereichen wurden dabei anonymisiert über ein elektronisches Eingabeinstrument erhoben. Die Ergebnisse zu diesen Evaluationsbereichen werden getrennt nach den Fächern Deutsch, Englisch, Französisch und Mathematik vorrangig deskriptiv berichtet. Für die Beschreibung der Schwierigkeit und kriterialen Validität der Aufgaben werden inferenzstatistische Analysen mittels gemischter Modelle durchgeführt, um schulspezifische Effekte zu berücksichtigen.

### Ergebnisse

Da die Dateneingabe noch nicht abgeschlossen ist, werden erste Zwischenergebnisse der Evaluation erst im vierten Quartal 2017 vorliegen. Aus diesem Grund können in diesem Abstracts noch keine Befunde berichtet werden. Die im Symposium dargestellten Ergebnisse zu den oben genannten Fragestellungen werden erste Anhaltspunkte dazu geben, inwieweit durch die Entwicklung gemeinsamer Aufgabenpools das Ziel erreichbar ist, eine stärkere Vergleichbarkeit von Anforderungen zu gewährleisten.

### **„Gemeinsames Prüfen“ in Gymnasien der Schweiz**

**Monika Holmeier<sup>1</sup>, Carmen Hirt<sup>2</sup>, Katharina Maag Merki<sup>2</sup>**  
<sup>1</sup>Pädagogische Hochschule FHNW, <sup>2</sup>Universität Zürich

#### Theoretischer Hintergrund

Während in Deutschland in fast allen Bundesländern das Zentralabitur eingeführt worden ist, um einheitliche Anforderungen und Beurteilungen in den Abiturprüfungen zu erreichen, wurde in der Schweiz das „Gemeinsame Prüfen“ implementiert. Dabei entwerfen zwei oder mehr Lehrpersonen im gymnasialen Unterricht und teilweise am Ende des Gymnasiums gemeinsame Prüfungen und/oder korrigieren und bewerten diese nach gemeinsam festgelegten Kriterien. Ziele des ‚Gemeinsamen Prüfens‘ sind eine stärkere Verbindlichkeit gemeinsamer Anforderungen in den Prüfungen, eine bessere Vergleichbarkeit der Noten und eine Professionalisierung der Lehrpersonen.

Bei der Umsetzung des ‚Gemeinsamen Prüfens‘ zeigt sich eine große Variabilität in den Schulen (Holmeier, Maag Merki & Hirt, 2017). So unterscheiden sich die implementierten Verfahren in Abhängigkeit des Faches, des Einführungsmodus (top-down vs. bottom-up), der Art (Abschlussprüfung, Jahrgangsprüfung) und Funktion (Selektions- oder Diagnosefunktion) der gemeinsamen Prüfungen, aber auch hinsichtlich der eingebundenen Akteure (einzelne Lehrpersonen, gesamte Fachschaften, mehrere Kantone). Die unterschiedliche Implementation kann dabei auf Rekontextualisierungsprozesse (Fend, 2008) in den Schulen zurückgeführt werden, wie sie auch bei der Einführung weiterer standardisierter Prüfungsformen nachgewiesen werden konnten (Bryk et al., 2010; Maag Merki, 2016a; Stringfield et al., 2008). So ist die Realisierung des ‚Gemeinsamen Prüfens‘ zwar von formalen Vorgaben und Rahmenbedingungen beeinflusst, die konkrete Umsetzung ist aber Ergebnis eines komplexen ‚Übersetzungsprozesses‘ der schulinternen Akteure.

Empirische Untersuchungen zu den Wirkungen standardisierter Prüfungsformen liegen international und national vor (u.a. Maag Merki, 2016b, Klein et al., 2014). Studien, die einen vertieften Blick auf die Implementation und die Prozesse der Durchführung neuer Prüfverfahren werfen, sind hingegen rar. Noch seltener lassen sich verfahrenvergleichende Analysen finden, die zu einem spezifischen Verständnis der Gemeinsamkeiten und Unterschiede verschiedener Verfahren beitragen und die es ermöglichen, verfahrensspezifische Stärken und Schwächen herauszuarbeiten.

#### Fragestellung

Dieser Beitrag setzt an der oben aufgeführten Forschungslücke an und untersucht folgende Fragestellungen:

- Inwiefern weisen die unterschiedlichen Verfahren des ‚Gemeinsamen Prüfens‘ Potenziale und Herausforderungen für eine nachhaltige Implementation in der schulischen Praxis auf?
- Welche Spannungsfelder lassen sich aus den Ergebnissen für die Implementation neuer Prüfungsverfahren ableiten?

Angenommen wird, dass die nachhaltige und effektive Implementation von einem komplexen Zusammenspiel interner (z.B. Support, Freiwilligkeit) und externer Faktoren (z.B. bottom-up, kein Kontrollinstrument) beeinflusst wird (Valli et al., 2008) und dass die jeweiligen Verfahren ihre je eigenen Potenziale und Herausforderungen aufweisen.

#### Methode

Datengrundlage sind 12 Fokusgruppengespräche in 4 Gymnasien der Deutschschweiz mit je 5-6 Lehrpersonen. Die Daten wurden mittels strukturierender Inhaltsanalyse (Mayring, 2008) fall-, verfahrens- und fachspezifisch sowie fall-, verfahrens- und fachvergleichend ausgewertet. Dabei geben einheitliche, übergreifende Befunde generelle Trends bei der Implementation des ‚Gemeinsamen Prüfens‘ wieder. Unterschiede deuten hingegen darauf hin, dass die Implementation vom schul- oder fachspezifischen Kontext oder vom Verfahren selbst abhängig ist. Über die verfahrenvergleichenden Analysen wurden die spezifischen Stärken und Schwächen der verschiedenen Verfahren ausgewiesen, aus denen wiederum Spannungsfelder herausgearbeitet wurden, die für die Gestaltung des ‚Gemeinsamen Prüfens‘ in der Praxis bedeutsam scheinen.

#### Ergebnisse

In den 4 untersuchten Gymnasien konnten insgesamt 37 unterschiedliche Verfahren des ‚Gemeinsamen Prüfens‘ identifiziert werden. Dies bestätigt die Annahme, dass die mit der Einführung verbundenen Vorgaben durch die beteiligten Akteure rekontextualisiert und an den schulischen aber auch unterrichtlichen Kontext angepasst werden. Zudem bestätigen die verfahrens- und fallvergleichenden Analysen, dass die Umsetzung von einem komplexen Zusammenspiel interner und externer Faktoren beeinflusst wird. So erweisen sich insbesondere die Art der Einführung (top-down vs. bottom-up; freiwillig vs. Pflicht), die Zusammenarbeit im Team, aber auch verfahrensinterne Faktoren wie bspw. die Transparenz der Ziele und die Qualität der gemeinsamen Prüfungen als bedeutsam. An der Konferenz werden weitere Faktoren sowie die sich daraus ergebenden Spannungsfelder bei der Implementation gemeinsamer Prüfungen präsentiert. Diese ermöglichen einen vertieften Einblick in die Implementation des ‚Gemeinsamen Prüfens‘ und geben Hinweise auf zu reflektierende Entscheidungen und Überlegungen bei der Einführung standardisierter Prüfungsformen.

## E19

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H001

### Individuelle, systemische und kompositionelle Effekte auf (erfolgreiche) Übergänge

Chair(s): **Nele Usslepp** (Universität Tübingen, Deutschland)

DiskutantIn(nen): **Markus Neuenschwander** (Pädagogische Hochschule der Fachhochschule Nordwestschweiz)

Verschiedene aufeinanderfolgende institutionalisierte Übergänge charakterisieren das Bildungssystem in Deutschland und der Schweiz. Da Übergangsentscheidungen weitreichende Konsequenzen für Bildungsbiographien, die spätere soziale Klasse und das Einkommen haben (z.B., Breen & Jonsson, 2005; Müller & Shavit, 1998; Tamborini, Kim, & Sakamoto, 2015), beschäftigte sich bereits eine Vielzahl an Studien mit diesen Gelenkstellen des Bildungssystems (z.B., Baumert, Stanat, & Watermann, 2006; Maaz, Hausen, McElvany, & Baumert, 2006; Trautwein & Baeriswyl, 2007). Dabei wurde gut belegt, dass an diesen Stellen Bildungsungleichheiten entstehen oder verstärkt werden können (z.B., Maaz, Trautwein, Lüdtke, & Baumert, 2008; Trautwein, Nagy, & Maaz, 2011). Der Fokus bisheriger Forschung lag jedoch überwiegend auf dem ersten Übergang von der Primarstufe in die weiterführende Schule (z.B., Ditton & Krüsken, 2006; Jonkmann, Maaz, McElvany, & Baumert, 2010; Paulus & Blossfeld, 2007). Dementsprechend fanden spätere Übergänge, Bildungsverläufe sowie Erfolgsindikatoren von Übergängen kaum Beachtung. Somit besteht großer Bedarf diesbezügliches Wissen über individuelle, systemische und kompositionelle Faktoren auf (erfolgreiche) Übergänge zu generieren. Dieses Wissen ist notwendig, um abschätzen zu können wo Möglichkeiten und Grenzen der Einflussnahme auf Bildungsentscheidungen und -verläufe liegen und inwiefern den angesprochenen Bildungsungleichheiten entgegengewirkt werden kann. Vor diesem Hintergrund thematisiert das vorliegende Symposium relevante Aspekte der Übergangsforschung, indem interdisziplinäre Perspektiven (Soziologie, Erziehungswissenschaft und Psychologie) und methodisch unterschiedliche Herangehensweise verwendet werden.

Beitrag 1 analysiert verschiedene Determinanten aus der Grundschulzeit auf das Nachholen des Abiturs und das vorzeitige Verlassen des Gymnasiums. Es wird gezeigt, dass Eltern und Kinder ihre Aspirationen verwirklichen sobald sie die Möglichkeit haben eine freie Bildungsentscheidung zu treffen und ansonsten die Aspirationen über Jahre stabil gehalten werden. Der zweite Beitrag betrachtet ebenfalls verschiedene Bildungsverläufe und analysiert dabei die Berufsmatura als Öffnungsoption des Bildungswesens. Dabei werden differenzielle Systemeffekte unterschiedlicher Personengruppen identifiziert. Beispielsweise scheinen Jugendliche mit Migrationshintergrund stärker von der Öffnung der Berufsmaturität zu profitieren. Beitrag 3 beschäftigt sich mit institutionellen und kompositionellen Effekten der Schulzugehörigkeit beim Übergang nach der Pflichtschulzeit. Es wird unter anderem aufgezeigt, dass je nach besuchter Schulform eine unterschiedlich hohe Übertrittswahrscheinlichkeit für die postobligatorische Ausbildung einhergeht. Abschließend analysiert Beitrag 4 inwieweit Persönlichkeitseigenschaften – neben dem sozialen Hintergrund und kognitiven Fähigkeiten – den erfolgreichen Übergang in die Ausbildung vorhersagen, wobei sich die Gewissenhaftigkeit als einflussstarkes Merkmal zeigte. Diskutiert werden die Einzelbeiträge von Prof. Dr. Markus Neuenschwander.

Mit dem übergreifenden Blick auf verschiedene Bildungsetappen werden in diesem Symposium Übergänge und Verläufe von der Primarstufe bis hin zur Tertiärstufe in den Blick genommen. Diese Ansicht und ihre Erkenntnisse liefern an den wichtigen Gelenkstellen des Bildungssystems Hinweise für professionelles Handeln auf individueller und institutioneller Ebene.

*Beiträge des Symposiums*

### Übergänge nach der Sekundarstufe I: „Korrektur“ von einst getroffenen Bildungsentscheidungen oder auch von Bildungsaspirationen?

**Nele Usslepp<sup>1</sup>, Franz Baeriswyl<sup>2</sup>, Nicolas Hübner<sup>1</sup>, Benjamin Nagengast<sup>1</sup>, Ulrich Trautwein<sup>1</sup>**

<sup>1</sup>Universität Tübingen, <sup>2</sup>Universität Freiburg/Schweiz

Durch die vertikale Öffnung des Schulsystems (bezeichnet die Anschlussfähigkeit zweier Bildungsgänge) sollen einst getroffene Bildungsentscheidungen „korrigiert“ werden können (Deutscher Bildungsrat, 1970), die beispielsweise aufgrund von inkorrekt diagnostizierten Leistungen getroffen wurden. Im deutschsprachigen Teil des Kantons Freiburg ermöglicht die vertikale Öffnung gleichzeitig die erste ‚freie‘ Bildungsentscheidung von Eltern und Schüler/-innen, da das Zuweisungsverfahren nach der Grundschule (fast) ausschließlich auf Basis der Lehrerempfehlung und eines standardisierten Leistungstest getroffen wird (Baeriswyl, Wandeler, Trautwein, & Oswald, 2006). Somit ist es möglich, dass Eltern und Kinder, die beispielsweise nicht die erhoffte Zuweisung erhielten, bereits seit dem Zeitpunkt des Wechsels in die Sekundarstufe den Plan verfolgen, die nächste Übertrittsschwelle zu nutzen, um ihre Aspirationen (Wunsch einen bestimmten Bildungsgang bzw.–abschluss zu erreichen) zu verwirklichen. Hinweise für diese Annahme lieferten bereits Kleine, Paulus und Blossfeld (2009), die aufzeigen konnten, dass sich Eltern teilweise nicht von ihren Bildungsvorstellungen abbringen lassen. Darüber hinaus stellen Aspirationen neben der schulischen Leistung einen besonders starken Prädiktor für Bildungslaufbahnentscheidungen dar (vgl. Becker, 2000; Kleine et al., 2009; Meulemann, 1985). Bisherige Forschungsarbeiten untersuchen die prognostische Kraft von Aspirationen bei Bildungsverlaufsentscheidungen jedoch gar nicht oder nur unzureichend.

Die vorliegende Studie untersucht (1), ob das vorzeitige Verlassen des Gymnasiums bzw. das Nachholen des Abiturs als eine „Korrektur“ von einst getroffenen Bildungsentscheidungen oder auch von Bildungsaspirationen zu verstehen ist. Hierbei finden Eltern- und Schüleraspirationen getrennt Berücksichtigung, da jene voneinander abweichen können (z.B. Wohlking & Ditton, 2012). Darüber hinaus soll analysiert werden, (2) wie gut die Elemente der Zuweisung nach der Grundschule (Lehrerempfehlung und standardisierter Leistungstest) diese „Korrektur“-Möglichkeit prognostizieren.

Die Datengrundlage der Analysen bildete die Vollerhebung zweier Übertrittsjahrgänge aus Deutschfreiburg (z.B. Baeriswyl et al., 2006). Zum ersten Messzeitpunkt, bei dem Lehrkräfte, Eltern und Schüler/-innen schriftlich befragt wurden, befanden sich die Schüler/-innen am Ende der Primarstufe. Dieselben Schüler/-innen wurden ein zweites Mal am Ende der Sekundarstufe I und ein drittes Mal zum Zeitpunkt der Sekundarstufe II befragt. Zur Testung unserer Forschungsfragen wurden Angaben von N= 1,219 Personen berücksichtigt und eine Reihe von logistischen Regressionsmodelle in Mplus 7.4 (Muthén & Muthén, 1998-2012) berechnet. Die hierarchische Struktur der Daten sowie fehlende Werte wurden bei den Analysen durch Verwendung von robusten Schätzverfahren sowie Full-Information-Maximum-Likelihood-(FIML-)Verfahren berücksichtigt.

Bereits die deskriptiven Befunde verdeutlichten, dass sich Schüler/-innen, die später „korrigieren“, in der Grundschule von jenen, die nicht „korrigieren“, hinsichtlich ihrer Aspirationen unterscheiden. Gleiches Muster zeigte sich auch für die Elternaspirationen. Dementsprechend brachten die Regressionsmodelle den Einfluss der Aspirationen bei allen „Korrektur“-Möglichkeiten hervor: Die Aufwärts-„Korrektur“ (das Nachholen des Abiturs) wurde sowohl durch die Schüler- (OR = 1.80; KI = [1.20, 2.70]) als auch die Elternaspirationen (OR = 1.91; KI = [1.34, 2.73]), die am Ende der Grundschule gemessen wurden,

vorhergesagt – unter Kontrolle des sozialen Hintergrunds, den Schulnoten und den Zuweisungselementen (Pseudo-R2 = .49). Die Abwärts-„Korrektur“ (Vorzeitige Verlassen des Gymnasiums) wurde dagegen ausschließlich von der Schüleraspiration beeinflusst (OR = 2.00; KI = [.13, 1.25]) – unter Kontrolle aller oben genannten Prädiktoren (Pseudo-R2 = .38). Darüber hinaus hatten die Zuweisungselemente nach der Grundschule (z.B. Leistungstest) keine prognostische Kraft und auch die Schulnoten am Ende der Sekundarstufe I beeinflussten lediglich, ob jemand das Abitur nachholt (OR = 3.54; KI = [.87, 1.66]).

Für das Deutschfreiburger Schulsystem lässt sich festhalten, dass zwar der Übergang nach der Grundschule sehr strikt geregelt ist, um soziale Disparitäten zu verhindern (Baeriswyl et al., 2006). Schüler/-innen und ihre Eltern scheinen jedoch bereits in der Grundschule ihren weiteren Bildungsweg zu planen und so auf die Verwirklichung ihrer Aspirationen zu warten bzw. diese umsetzen sobald sie die Möglichkeit dazu haben. Somit lassen sich Hinweise finden, dass der Übergang nach der Sekundarstufe I zwar eine "Korrektur" von einstigen Bildungsentscheidungen ist - aber nicht von Bildungsaspirationen.

## **Öffnungsoption Berufsmatura – Wer profitiert? Ein Vergleich der Durchlässigkeit verschiedener Bildungswege zur Tertiärbildung**

**Dominique Oesch, Stefan Sacchi, Rolf Becker**  
Universität Bern

Durchlässige Übergänge zwischen Bildungsstufen erhöhen die Bildungsmobilität und erleichtern Anpassungen von individuellen Bildungslaufbahnen – auch zu späteren Zeitpunkten im Bildungsverlauf. Dabei ist mit dem Prinzip der Durchlässigkeit auch die Hoffnung verbunden, dass Jugendliche, die auf den klassischen Bildungswegen eher benachteiligt sind, einen alternativen Bildungsweg vorfinden, mit dem sie ihr Potential bestmöglich ausschöpfen können (Oesch, 2017).

Die Berufsmaturität gilt in der Schweiz seit den 1990er Jahren als eines der wichtigsten Instrumente zur Erhöhung der Durchlässigkeit zwischen der Sekundarstufe II und der Tertiärstufe, indem sie im jeweiligen Berufsfeld einen prüfungsfreien Zugang zu entsprechenden Fachhochschulen und über die „Passerelle“ den universitären Hochschulzugang ermöglicht. Da schätzungsweise lediglich 7 Prozent eines Altersjahrgangs auf diesem Weg eine Ausbildung auf Tertiärstufe A abschliessen, scheint sich die politische Hoffnung auf erhöhte Durchlässigkeit zu erfüllen, wenn auch vielleicht nicht in erhofftem Ausmass (Meyer, 2016). Zudem wirkt die soziale Herkunft auch auf die Chance, eine Berufsmaturität zu erwerben: Personen aus sozial benachteiligten Schichten erlangen ceteris paribus zwar eher eine Berufsmaturität als eine gymnasiale Matura (Schumann, 2011), nehmen jedoch seltener eine Tertiär-Ausbildung auf als GymnasiastInnen (Kost, Lüthi & Fischer, 2017). Weiter zeigen sich beim Weg über die Berufsmaturität zur Tertiärstufe A sowohl Effekte des Geschlechts zugunsten der Männer (Kost et al., 2017; Jäpel, 2017), als auch Effekte des Migrationshintergrunds zugunsten Einheimischer (Murdoch et al., 2017).

Bisherige Arbeiten im Bereich nachobligatorischer Übergänge in der Schweiz fokussieren vorrangig punktuelle Bildungsentscheidungen bezüglich der Berufsmatura (bspw. Jäpel, 2017, Glauser, 2015) resp. des Zugangs zum Tertiärbereich (Schumann, 2011; Schmid & Gonon, 2011). Ungeklärt bleibt indessen, in welchem Ausmass kumulierte Übergangsentscheidungen über verschiedene Bildungsetappen auf den Tertiärstufenzugang und -abschluss wirken, da entsprechende Analysen bislang fehlen. So stellen sich die leitenden Forschungsfragen nach der genutzten Durchlässigkeit verschiedener Bildungswege bis Ende des dritten Lebensjahrzehnts, und welche Personengruppen insbesondere von der Öffnung durch die Berufsmaturität profitieren.

Dieser Frage soll auf Basis der Daten der Längsschnittstudie TREE (Transitionen von der Erstausbildung ins Erwerbsleben) nachgegangen werden, da diese für die Schweiz repräsentative Analysen von nachobligatorischen Bildungsverläufen auf Basis von Individualdaten über einen Beobachtungszeitraum von 14 Jahren erlauben. Die Stichprobe umfasst 5'565 Jugendliche, für welche vollständige Angaben zu Herkunfts-Merkmalen wie der sozialen Herkunft und dem Migrationsstatus sowie dem Geschlecht vorliegen.

Die Analysestrategie baut auf einer empirisch gestützten Typologie der wichtigsten Bildungswege zu tertiären Bildungsabschlüssen auf. Für jeden dieser Bildungswege wird eine Reihe von aufeinander aufbauenden binärologistischen Regressionsanalysen geschätzt, mit denen die Übertrittswahrscheinlichkeiten zwischen den massgeblichen Zwischenetappen ermittelt werden, die auf dem Weg zum Tertiärabschluss zu durchlaufen sind. Gestützt auf die so ermittelten Übertrittswahrscheinlichkeiten kann ein kumulatives Mass für die Durchlässigkeit des betreffenden Bildungswegs für Jugendliche ermittelt werden, die von Anfang an unterschiedliche und je nachdem nachteilige Voraussetzungen bzgl. sozialer Herkunft, Migrationshintergrund oder Geschlecht mitbringen. Damit kann die relative Durchlässigkeit verschiedener institutionalisierter Bildungswege für unterschiedliche Personengruppen verglichen werden – so z. B. Wege, die über die gymnasiale bzw. die Berufsmaturität zu einem Fachhochschul- oder zu einem Universitätsabschluss führen.

Erste vorläufige Analysen des Hochschulzugangs weisen auf differenzielle Effekte unterschiedlicher Personengruppen hin. Der Bildungsweg über die Berufsbildung mit Berufsmatur zur Hochschule ist für Männer ceteris paribus vergleichbar durchlässig wie derjenige ohne Berufsmatur. Für Frauen hingegen ist die Durchlässigkeit bei einem Bildungspfad mit Maturitätszeugnis höher als bei einem ohne Maturität. Jugendliche der zweiten Einwanderergeneration scheinen von der Öffnung der Berufsmaturität zur Hochschule stärker zu profitieren als Einheimische. Zudem zeigt sich der Bildungsweg über die Berufsmaturität auch für die sozioökonomisch Schwächsten (unterstes SES-Quartil) als wesentlich offener als über die gymnasiale Matur. Differenziertere Analysen mit Fokus auf Eintritt und Abschluss weiterer Angebote auf Tertiärstufe folgen.

## **Schultypen als differenzielle Entwicklungsmilieus beim Übergang nach dem Ende der Pflichtschulzeit?**

**Katja Scharenberg**

Pädagogische Hochschule Freiburg

Ein Charakteristikum des schweizerischen Schulsystems ist die externe Differenzierung in verschiedene Bildungsgänge mit unterschiedlichen Anforderungsniveaus. Für andere Schulsysteme (z.B. Deutschland) wurde gezeigt, dass eine solche externe Leistungsdifferenzierung einerseits zu sozialer Stratifizierung bzw. Segregation (Baumert et al., 2003, 2006) führen kann. Andererseits können auch „differenzielle Lern- und Entwicklungsmilieus“ (Baumert & Schümer, 2001, S. 462) den Schülerinnen und Schülern je nach besuchtem Schultyp unterschiedliche Entwicklungschancen eröffnen. Diese können dazu führen, dass die individuelle Entwicklung von Jugendlichen ungleich verläuft, erschwert oder sogar verhindert wird (Baumert et al., 2006).

Auswirkungen differenzieller Lern- und Entwicklungsmilieus wurden für verschiedene Altersgruppen und Kompetenzdomänen nachgewiesen (z.B. Baumert et al., 2006, 2009, 2010; Becker et al., 2006, 2012; Gröhlich et al., 2010a, 2010b; Nikolova, 2011): So deutlichten Baumert et al. (2006) auf der Grundlage von PISA 2000 die Auswirkungen der Schulzusammensetzung und Schulformzugehörigkeit, die im deutschen Schulsystem zu deutlichen Schereneffekten bei der Leistungsentwicklung bis zum Ende der Sekundarstufe I führt. Für das schweizerische Sekundarschulsystem zeigten Neumann et al. (2007) für die Kantone

Wallis und Fribourg, dass die Leistungsentwicklung im Fach Französisch als Fremdsprache in Schultypen mit erweiterten Anforderungen signifikant höher ausfällt als in Schultypen mit Grundansprüchen. Es ließen sich zudem Leistungsvorteile aufgrund einer leistungsstärkeren und sozial privilegiierteren Schulkomposition nachweisen. Diese waren allerdings mit institutionellen Effekten des Schultyps konfundiert.

Basierend auf diesen Überlegungen geht dieser Beitrag folgenden Forschungsfragen nach: 1.) Lassen sich institutionelle Effekte der Schultypzugehörigkeit für den Übertritt in die postobligatorische Ausbildung nach dem Ende der Sekundarstufe I nachweisen? 2.) Inwiefern sind schultypspezifische Übergangsraten auch auf die jeweilige Schülerkomposition zurückzuführen?

Datengrundlage ist die Studie Transitionen von der Erstausbildung ins Erwerbsleben (TREE, Kohorte 1). TREE ist in der Schweiz die erste repräsentative Längsschnitterhebung zum Übergang Jugendlicher von der Schule ins Erwerbs- und Erwachsenenleben. Im Zentrum der Untersuchung stehen die Ausbildungs- und Erwerbsverläufe von Jugendlichen nach dem Ende der obligatorischen Schulzeit. Die TREE-Stichprobe der ersten Kohorte umfasst mehr als 6.000 Jugendliche, die an PISA 2000 teilnahmen, im selben Jahr aus der obligatorischen Schulpflicht entlassen wurden und seitdem über nunmehr neun Messzeitpunkte anhand von standardisierten Erhebungsinstrumenten schriftlich und telefonisch zu ihren Ausbildungs- und Erwerbsverläufen sowie weiteren damit in Zusammenhang stehenden Kontextmerkmalen befragt wurden. Die Stichprobe ist national und sprachregional repräsentativ für rund 80.000 Schulabgängerinnen und Schulabgänger aus dem Jahr 2000. Die Analysen fokussieren die Übergangswahrscheinlichkeiten beim Übertritt in die postobligatorische Ausbildung (n=4.192 Schülerinnen und Schüler in 217 Schulen).

Methodisch wird ein logistisches Mehrebenenmodell (HLM 7.01; Raudenbush et al., 2013) spezifiziert, das auf Individualebene das Geschlecht (Referenz: Mädchen), die Sprache im Haushalt (Referenz: keine offizielle Landessprache), den Sozialstatus (HISEI; Ganzeboom et al., 1992) und die PISA-Lesekompetenz (Adams & Wu, 2002) als Merkmale der differenziellen Eingangsselektivität umfasst. Kontinuierliche Merkmale wurden auf Individualebene z-standardisiert (M=0; SD=1). Auf Aggregatebene wurden die durchschnittliche leistungsbezogene, soziale und kulturelle Zusammensetzung der Schülerschaft sowie die Schultypzugehörigkeit als Institutionsmerkmale kontrolliert (anhand der Schulzugehörigkeit über den Mittelwert aggregiert).

Die Ergebnisse zeigen, dass 29,8 % der Schülerinnen und Schüler aus Schultypen mit Grundanforderungen der direkte Eintritt in eine zertifizierende Ausbildung auf der Sekundarstufe II nicht gelingt, während dieser Anteil bei jenen aus Schultypen mit erweiterten Anforderungen mit 13,2 % signifikant geringer ausfällt ( $\chi^2=126,16$ ,  $p<.001$ ). Zudem bestehen deutliche Unterschiede in der sozialen ( $d=1,63$ ), kulturellen ( $d=1,02$ ) und leistungsbezogenen Schulkomposition ( $d=2,31$ ) zugunsten der Schultypen mit erweiterten Anforderungen (jeweils  $p<.001$ ). Mehrebenenanalytisch lassen sich unter Kontrolle individueller Merkmale zunächst signifikant höhere Übertrittswahrscheinlichkeiten an Schultypen mit erweiterten Anforderungen beobachten (OR=1,44,  $p<.05$ ). Diese Übergangsvorteile sind jedoch auf eine günstigere soziale Zusammensetzung der Schülerschaft an den Schultypen mit erweiterten Anforderungen zurückzuführen (OR= 2,13,  $p<.01$ ).

Die Befunde erweitern somit den bisherigen Forschungsstand zu differenziellen Entwicklungsmilieus um bildungsbiografische Kriteriumsvariablen und unterstreichen die Bedeutung kompositioneller und institutioneller Schulmerkmale für Entwicklungsverläufe bis ins junge Erwachsenenalter.

## **Erfolgreich in die Berufsausbildung: Die Rolle der Persönlichkeit für den Übergang von der Schule in die Ausbildung**

**Désirée Nießen, Clemens Lechner, Daniel Danner**

GESIS – Leibniz-Institut für Sozialwissenschaften

Persönlichkeitseigenschaften wie die Big Five (Extraversion, Verträglichkeit, Gewissenhaftigkeit, Emotionale Stabilität und Offenheit; Costa & McCrae, 1992) beeinflussen wichtige Indikatoren des Lebenserfolgs wie z. B. Gesundheit (Bogg & Roberts, 2004) und Lebenszufriedenheit (Soto, 2015) sowie Ausbildungs- und Berufserfolg, auch über den sozioökonomischen Status (SES) und über kognitive Fähigkeiten hinaus (Heckman & Kautz, 2012; Roberts et al., 2007; Spengler et al., 2015). In der Forschung zu Bildungsübergängen werden Persönlichkeitseigenschaften als Determinanten von Bildungsentscheidungen und Erfolg jedoch nur selten berücksichtigt (Shanahan et al., 2014). Dabei sind zwei mögliche Rollen von Persönlichkeitseigenschaften auf Bildungsübergänge denkbar: eine additive (Bildungserfolg als Resultat voneinander unabhängiger Einflussfaktoren) und eine interaktive (Zusammenspiel von Persönlichkeitseigenschaften und SES oder kognitiven Fähigkeiten in Hinblick auf Bildungserfolg). Während empirische Studien zu additiven Effekten schon eine inkrementelle Validität von Persönlichkeitseigenschaften hinsichtlich Lernmotivation, Kompetenzerwerb und Bildungserfolg nachweisen konnten (Chamorro-Przemuzic & Furnham, 2003; Poropat, 2009; Spengler, Lütke, Martin & Brunner, 2013), befassten sich wenige Studien mit möglichen interaktiven Effekten (Damian et al., 2015; Rammstedt, Danner & Martin, 2016). Darüber hinaus nehmen bisherige Befunde nur auf einen engen Kreis an globalen Erfolgsindikatoren (erreichter Notendurchschnitt, höchster Bildungsabschluss) Bezug.

Die vorliegende Studie beabsichtigte, diese Lücke zu schließen und sowohl additive als auch interaktive Effekte von Persönlichkeitseigenschaften auf ein breiteres Spektrum von Erfolgsindikatoren zu erforschen. Das Ziel war es zu überprüfen, ob Persönlichkeitseigenschaften auch nach Berücksichtigung des unabhängigen Einflusses von SES und kognitiven Fähigkeiten noch interindividuelle Unterschiede im Erfolg von Bildungsübergängen von der Schule in die Ausbildung aufklären können.

Die Fragestellung wurde mit Hilfe der Startkohorte 4 des National Educational Panel Survey (NEPS) untersucht, einer Längsschnittstudie, die den Bildungsweg von Personen nachverfolgt, die zu Beginn der ersten Erhebungswelle im Jahr 2010 die neunte Klasse besuchten. In die vorliegende Untersuchung wurden nur Abgänger der Haupt- und Realschule miteinbezogen (N = 4.314), die im Anschluss an ihren Schulabschluss nicht in die Sekundarstufe II wechselten.

Die Messung der Big-Five-Persönlichkeitsdimensionen (BF) erfolgte mit dem BFI-10 (Rammstedt & John, 2007). Als Indikatoren für den Übergangserfolg dienten objektive Indikatoren (Erlangen eines Ausbildungsplatzes, Ausbildungsabbruch, Abschluss und Note der Ausbildung) sowie subjektive Indikatoren (Gesundheit und Zufriedenheit mit der Ausbildung). Zentrale soziodemografische Faktoren (SES der Eltern, Migrationshintergrund, Geschlecht) und kognitive Fähigkeiten fungierten als Kovariaten. Zur Analyse wurden lineare Regressionen verwendet.

Die Analysen zeigen deutliche additive Effekte, der zufolge Persönlichkeitseigenschaften den Erfolg in Bildungsübergängen vorherzusagen vermögen, und zwar über soziodemografische Faktoren und kognitive Fähigkeiten hinaus. Die stärksten positiven signifikanten Effekte auf den Übergangserfolg zeitigte hierbei Gewissenhaftigkeit bei der Vorhersage folgender Erfolgsindikatoren: Ausbildungsplatz erreicht (inkrementelles  $R^2_{BF} = .006$ ), Ausbildung begonnen (inkrementelles  $R^2_{BF} =$

.005), Anzahl an Ausbildungsplatzzusagen (inkrementelles  $R^2BF = .004$ ), Ausbildungsnote (inkrementelles  $R^2BF = .025$ ), Gesundheitszustand (inkrementelles  $R^2BF = .026$ ) und Zufriedenheit mit der Ausbildung (inkrementelles  $R^2BF = .012$ ). Bei vereinzelt Erfolgsindikatoren waren zudem Extraversion (Ausbildungsplatz erreicht, Ausbildungsnote) bzw. Emotionale Stabilität (Gesundheitszustand) prädiktiv.

Auch zusätzlich aufgenommene Interaktionsterme zwischen den Persönlichkeitsdimensionen und Kovariaten spielten bei der Vorhersage diverser Kriterien eine Rolle. So zeigte sich z. B. bei der Ausbildungsplatzerreichung ein statistisch signifikanter Interaktionseffekt zwischen Extraversion und Migrationshintergrund: Ein hohes Maß an Extraversion führte bei Personen, die Deutsch als Muttersprache hatten, eher zu einer Ausbildungsplatzzusage, als bei Personen, deren Muttersprache nicht Deutsch war. Ein signifikanter positiver Zusammenhang zwischen Ausbildungsplatz erreicht und Extraversion blieb als Haupteffekt bestehen.

Zusammenfassend lässt sich sagen, dass es sich bei der vorliegenden Studie um eine der ersten handelt, die den Zusammenhang zwischen Persönlichkeit und Bildungsübergangserfolg untersuchte, und die Ergebnisse das Vorhandensein eines ebensolchen im Sinne eines additiven und eines interaktiven Modells demonstrieren. Als theoretische Implikation ist die bislang unterschätzte Rolle von Persönlichkeit im Vergleich zu SES und kognitiven Fähigkeiten zu nennen und als praktische Gewissenhaftigkeit als mögliches Ziel für Interventionen hervorzuheben.

## E20

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Kollegienhaus H120

### Internalisierendes Verhalten im schulischen Peerkontext

Chair(s): **Carmen Zurbrigen** (Universität Bielefeld, Deutschland), **Christoph Michael Müller** (Universität Fribourg, Schweiz)

DiskutantIn(nen): **Tina Hascher** (Universität Bern, Schweiz)

Internalisierendes Problemverhalten von Kindern und Jugendlichen umfasst 'nach innen' gerichtete Schwierigkeiten, wie Ängstlichkeit, Depressivität und sozialen Rückzug (Abela & Hankin, 2008). Trotz seiner ungünstigen Wirkung auf schulisches und soziales Lernen wird internalisierendes Verhalten in der Schule bislang wenig beachtet (Bilz, 2008; Lane, Oakes, Menzies & Germer, 2014). Als ein zentraler Einflussfaktor für solches Verhalten gelten die Peers (Giletta, Burk, Scholte, Engels & Prinstein, 2013). Der Einfluss der Gleichaltrigen auf internalisierendes Verhalten kann sich beispielsweise durch gegenseitige Verstärkung im Rahmen exzessiver Diskussionen eigener Probleme untereinander vollziehen (sog. Co-rumination; Rose, 2002).

Solche Prozesse wurden bisher in Bezug auf freizeitbezogene Freundschaftsdyaden und Cliques untersucht (z.B. Brechwald & Prinstein, 2011). Unklar ist hingegen, inwiefern Peereinfluss auf internalisierendes Verhalten auch von Merkmalen der Klassenkameradinnen und Klassenkameraden abhängt. Der Einfluss der schulischen Peers dürfte dabei auch je nach individuellen Merkmalen der beeinflussten Schülerinnen und Schüler variieren. Neben Peersozialisationswirkungen können sich internalisierende Verhaltensweisen zudem als Reaktion auf negative Erlebnisse unter den Peers herausbilden (Zwierzynska, Wolke & Lereya, 2013). In diesem Zusammenhang ist mehr Wissen dazu nötig, welche Peerkonstellationen im schulischen Kontext für die Entwicklung internalisierenden Verhaltens von Schülerinnen und Schülern bedeutsam sind. Das geplante Symposium vereint zu der skizzierten Thematik drei Studien aus Deutschland und der Schweiz.

Der erste Beitrag untersucht Sozialisationswirkungen auf internalisierendes Verhalten unter Klassenkameradinnen und Klassenkameraden der Sekundarstufe I. Die Studie zeigt, dass jugendliche Mädchen vor allem durch die Ausprägung internalisierenden Verhaltens unter ihren Klassenkameradinnen beeinflusst werden. Internalisierendes Verhalten von Jungen wurde hingegen weder durch die männlichen noch durch die weiblichen Peers beeinflusst.

Im zweiten Beitrag wird der Zusammenhang von sozialem Status in der Klasse und internalisierendem Verhalten anhand der Kontrastierung einer herkömmlichen und einer neu entwickelten Statusgruppenmethode untersucht. Die Ergebnisse deuten darauf hin, dass unbeachtete Schülerinnen und Schüler mit einer erhöhten Wahrscheinlichkeit internalisierende Verhaltensauffälligkeiten zeigen, wobei sich die Effektstärken je nach Analyseverfahren unterscheiden.

Der dritte Beitrag geht der Frage nach, wie Schülerinnen und Schüler mit internalisierendem Verhalten verschiedene Sozialformen (Partner- und Gruppenarbeit vs. Einzelarbeit) im Unterricht erleben. Des Weiteren wird überprüft, ob die Klassenkomposition bezüglich Problemverhalten und prosozialem Verhalten den Zusammenhang zwischen dem emotionalen Erleben und den Interaktionen mit den Peers moderiert.

Gesamthaft betrachtet liefern die Beiträge dieses Symposiums neue Erkenntnisse zu einer im Schulkontext bisher wenig berücksichtigten Problematik. Die Befunde der Studien werden von einer DiskutantIn zusammenfassend in methodischer und inhaltlicher Hinsicht diskutiert, um anschließend weiterführende Forschungsfragen und praktische Implikationen aufzuzeigen.

#### Beiträge des Symposiums

### Peereinfluss auf internalisierendes Verhalten - Unterschiede zwischen Jungen und Mädchen

**Sybille Arm, Verena Hofmann, Christoph Michael Müller**  
Universität Fribourg, Schweiz

Forschungsergebnisse zeigen, dass die Ausprägung an internalisierendem Verhalten unter den Freund(inn)en einen Einfluss auf die individuelle Entwicklung von internalisierenden Verhaltensproblemen hat (Conway, Rancourt, Adelman, Burk, & Prinstein, 2011). Als ein zentraler Peereinflussmechanismus gilt dabei die sogenannte „Co-rumination“, ein kommunikativer Prozess, der durch das gemeinsame exzessive und nicht lösungsorientierte Diskutieren eigener Probleme unter den Peers gekennzeichnet ist (Rose, 2002). Da sich die frühere Forschung auf Freundschaftsbeziehungen beschränkt hat, ist offen, ob auch die Charakteristika der „unfreiwilligen“ Peergruppe der Klassenkamerad(inn)en internalisierendes Verhalten beeinflussen. Darüber hinaus könnte sowohl das Geschlecht der beeinflussten als auch der beeinflussenden Schüler(innen) bedeutsam sein, da internalisierendes Verhalten oft als geschlechterkonformes Verhalten für Mädchen wahrgenommen wird (Broderick & Korteland, 2002; Rose & Rudolph, 2006).

In der vorliegenden Studie wurde erwartet, dass eine höhere Ausprägung internalisierenden Verhaltens unter allen Klassenkamerad(inn)en zu einer Zunahme individuellen internalisierenden Verhaltens beiträgt (Hyp. 1). Da internalisierende Verhaltensweisen in der Adoleszenz oft als geschlechterkonform für Mädchen betrachtet werden, wurde zudem vermutet, dass Mädchen in diesem Bereich stärker von den Klassenkamerad(inn)en beeinflusst werden als Jungen (Hyp. 2). In Bezug auf den Effekt des Geschlechts der beeinflussenden Schüler(innen) wurde angenommen, dass Mädchen und Jungen stärker durch Klassenkamerad(inn)en des gleichen Geschlechts beeinflusst werden (Hyp. 3 und Hyp. 4).

Die Hypothesen wurden anhand von Daten aus vier Messzeitpunkten der longitudinalen „Freiburger Studie zum Peereinfluss in Schulen“ überprüft, wobei sich zwei Erhebungen auf das siebte Schuljahr (T1, T2) und je eine auf das achte (T3) und neunte Schuljahr (T4) bezogen. Insgesamt nahmen 864 Schüler(innen) aus 56 Schulklassen und 8 Sekundarschulen teil (Alter bei T1 M = 13.12 Jahre; SDAAlter = .48; 52% Jungen). Die Schüler(innen) verblieben während der gesamten Sekundarstufe I in ihren nach Bildungsgängen gegliederten Klassen. Zur Erhebung internalisierenden Verhaltens bearbeiteten die Teilnehmenden zu allen vier Messzeitpunkten anonym eine standardisierte Selbstauskunftsskala zu ängstlichem und depressivem Verhalten mit 10 Items (Skala Ängstlichkeit/Depression des Screenings Psychischer Störungen im Jugendalter; Hampel & Petermann, 2005). Die Ausprägung internalisierenden Verhaltens unter den Klassenkamerad(inn)en wurde durch Klassenmittelwerte der genannten Skala repräsentiert (sog. deskriptive Gruppennorm; Henry et al., 2000). Angesichts der geschachtelten Datenstruktur (Messzeitpunkte innerhalb von Individuen; Individuen innerhalb von Schulklassen) wurden mit der Software MLwiN 2.36 Mehrebenenmodelle mit drei Ebenen berechnet. Um reziproke Kausalitäten zu vermeiden, wurden die Variablen zeitlich versetzt gemessen (sog. lagged models), d.h. die Klassenmittelwerte von T1 bis T3 sagten jeweils die Individualwerte von T2 bis T4 voraus (Singer & Willett, 2003). Die Faktoren Geschlecht und Bildungsgang wurden in allen Modellen statistisch kontrolliert.

Modell 1 zeigte, dass das mittlere Niveau an internalisierendem Verhalten in der Klasse kein signifikanter Prädiktor der individuellen Ausprägung zukünftigen internalisierenden Verhaltens war (Hyp. 1). Allerdings fand sich in Modell 2 ein signifikanter Interaktionseffekt zwischen dem Klassenmittelwert internalisierenden Verhaltens und dem Geschlecht auf das zukünftige individuelle Verhalten. Dieser zeigte, dass Mädchen stärker von dem Verhalten in ihrer Klasse beeinflusst wurden als Jungen (Hyp. 2). Zur Prüfung der weiteren Hypothesen wurde der Datensatz nach dem Faktor Geschlecht der Jugendlichen aufgeteilt. Die folgenden Analysen zeigten, dass der Verhaltensmittelwert der Klassenkameradinnen (nicht aber jener der Klassenkameraden) das individuelle internalisierende Verhalten der Mädchen signifikant vorhersagte (Hyp. 3). Das internalisierende Verhalten von Jungen wurde weder vom internalisierenden Verhalten der männlichen noch von jenem der weiblichen Mitschüler(innen) vorhergesagt (Hyp. 4).

Die dargestellten Ergebnisse erweitern bisherige Befunde zu Geschlechterunterschieden bei Peereinflussprozessen. Beispielsweise ergänzen sie den Befund früherer Untersuchungen, dass Jungen im Bereich der Delinquenz beeinflussbarer sind als Mädchen (Müller, Hofmann, & Arm, 2016; Steinberg & Monahan, 2007). Da Delinquenz von Jugendlichen typischerweise als ein maskulines und internalisierendes Verhalten als ein feminines Verhalten betrachtet wird, kann die Geschlechterkonformität spezifischer Verhaltensweisen eine wichtige Erklärung für eine unterschiedliche Peereinflussbarkeit von Jugendlichen sein.

## **Zeigen unbeachtete Kinder tatsächlich ein höheres Maß an internalisierenden Verhaltensweisen?**

**Pawel R. Kulawiak, Jürgen Wilbert**  
Universität Potsdam, Deutschland

### Theorie

Vielfach wurde argumentiert, dass internalisierende Verhaltensweisen (soziale Zurückgezogenheit, Ängstlichkeit, Depressivität) mit einer ungünstigen sozialen Integration zusammenhängen (Strauss, 1988). Dies ist intuitiv nachvollziehbar, da davon ausgegangen werden kann, dass Kinder aufgrund ihrer internalisierenden Verhaltensweisen von ihren Peers unbeachtet bleiben. Die empirische Evidenz zum Zusammenhang zwischen internalisierenden Verhaltensweisen und dem soziometrischen Status „unbeachtet“ ist jedoch widersprüchlich (Rubin et al., 1989). Die Befunde deuten einerseits darauf hin, dass soziometrisch unbeachtete Kinder, im Vergleich zu ihren Peers, ein höheres Maß an internalisierenden Verhaltensweisen aufweisen (Crews et al., 2007; Harrist et al., 1997). Andererseits berichten Untersuchungen keinen oder sogar einen schwachen negativen Zusammenhang (Boivin et al., 1994; Cassidy & Asher, 1992). Potenzielle Gründe für die widersprüchlichen Evidenzen könnten methodische Limitationen in der Bestimmung der soziometrischen Statusgruppen (Rubin et al., 1989) sowie die datenanalytische Vernachlässigung der genesteten Datenstruktur (Kinder in Peergruppen/Schulklassen) (Snijders, 2016) sein.

### Ziel

Der vorliegende Beitrag verdeutlicht anhand empirischer Daten die bisherigen Limitationen in der Bestimmung der soziometrischen Statusgruppen. Dabei wird eine klassische mit einer neu entwickelten Statusgruppenmethode kontrastiert. Die Methoden werden in Bezug auf die Analyse des Zusammenhangs zwischen internalisierenden Verhaltensweisen und soziometrischem Status miteinander verglichen. Beide Methoden werden jeweils mit und ohne Berücksichtigung der genesteten Datenstruktur angewendet. Insgesamt soll die Studie zeigen, ob die methodischen Neuerungen einen Beitrag zum besseren Verständnis der bisher widersprüchlichen Befunde leisten können.

### Methode

Die Querschnittsdaten entstammen einer Untersuchung an Brandenburger Grundschulen (22 Klassen, Klassenstufen 3-6, N=316 Kinder). Internalisierendes Verhalten wurde mit der Lehrerversion des SDQ gemessen (Subskala: internalisierendes Verhalten, Goodman et al., 2010). Die Soziometrie erfolgte anhand von Peernominationen. Die Statusgruppenzuordnung erfolgte anhand der soziometrischen Klassennorm (unbeachtete Kinder sind innerhalb ihrer Klassen unterdurchschnittlich beliebt/unbeliebt) gemäß der Methode nach Coie und Dodge (1983). Alternativ dazu beschreibt die neue Methode jedes Kind im Sinne einer kontinuierlichen Anpassung an soziometrische Kriterienpunkte (Extrempunkte). Die Anpassung der Kinder an die Extrempunkte bietet somit eine kriteriumsorientierte Interpretation (der unbeachtete Extrempunkt ist definiert durch nicht vorhandene Beliebtheit/Unbeliebtheit). Die Analysen ohne Berücksichtigung der genesteten Datenstruktur wurden mit OLS-Regressionsanalysen durchgeführt, wohingegen die genestete Datenstruktur durch Multilevel-Regressionsanalysen berücksichtigt wurde (internalisierendes Verhalten regressiert auf den soziometrischen Status).

### Ergebnisse

Mit beiden Statusgruppenmethoden zeigte sich bei Verwendung von Multilevelanalysen ein positiver Zusammenhang zwischen dem soziometrischen Status „unbeachtet“ und internalisierendem Verhalten. Der Effekt des soziometrischen Status war größer für den kriteriumsorientierten ( $b=9.05$ , 95%-CI[-2.47;19.78]) als für den klassennormorientierten soziometrischen Status ( $b=1.22$ , 95%-CI[-1.53;3.64]). Die Effektvariation zwischen den Klassen war beträchtlich (in einigen Klassen zeigten sich negative Zusammenhänge) und größer für den kriteriumsorientierten (random-slope-SD=20.60) als für den klassennormorientierten soziometrischen Status (random-slope-SD=2.68). Bezüglich der OLS-Regressionsanalysen (ohne Berücksichtigung der genesteten Datenstruktur) zeigte sich für den klassennormorientierten soziometrischen Status ( $b=2.07$ , 95%-CI[-0.06; 4.20]) ein höherer Effekt auf internalisierendes Verhalten als in der Multilevelanalyse und es kam zu einer Effekturnkehrung für den kriteriumsorientierten soziometrischen Status ( $b=-0.80$ , 95%-CI[-4.17;2.57]).

### Diskussion

Grundsätzlich bestärkt die Mehrheit der Analysen die Annahme, dass unbeachtete Kinder ein höheres Niveau an internalisierenden Verhaltensweisen aufweisen. Allerdings unterschieden sich die Effektstärken deutlich, je nachdem ob eine klassennorm- oder kriteriumsorientierte Analyse eingesetzt wurde. Auch die Wahl der einer Multilevel- oder OLS-Regression beeinflusste die Effekte maßgeblich bis hin zu einer Effekturnkehrung. Die teilweise höheren Effektstärken bei der OLS-Regressionsanalyse hängen vermutlich mit der fehlenden Berücksichtigung der Effektvariationen zwischen Klassen zusammen und sind daher als überschätzt zu beurteilen. Effekverzerrungen bei OLS-Regressionsanalysen (Huang, 2017) könnten auch zur Erklärung widersprüchlicher Evidenzen beitragen. Die in vielen Studien verwendete Statusgruppenmethode nach Coie und Dodge (1983) forciert mittels Intra-Klassen-Standardisierung ebenfalls die Probleme einer Missachtung der genesteten Datenstruktur (Elimination der soziometrischen Inter-Klassen-Unterschiede). Nebst diesen methodischen Erklärungen muss auch die Vielfalt der unterschiedlichen Operationalisierungen internalisierenden Verhaltens als eine Quelle für widersprüchliche Evidenzen im Blick behalten werden (Rubin et al., 1989).



## **Emotionales Erleben von Schülerinnen und Schülern mit internalisierendem Verhalten bei Peerinteraktionen im Unterricht**

**Carmen Zurbriggen<sup>1</sup>, Martin Venetz<sup>2</sup>**

<sup>1</sup>Universität Bielefeld, Deutschland, <sup>2</sup>Interkantonale Hochschule für Heilpädagogik Zürich, Schweiz

Internalisierende Verhaltensweisen beeinträchtigen das schulische Lernen von Jugendlichen und deren soziale Beziehungen zu ihren Klassenkameraden (Olympia, Heathfield, Jenson, Majszak, Ramos-Matias & Thacker, 2013). Im schulischen Alltag werden subklinische Ausprägungen internalisierender Verhaltensweisen wie emotionale Probleme oder Peer-Probleme jedoch aufgrund der nach innen gerichteten Charakteristik solcher Symptome oft übersehen (Lane, Oakes, Menzies, & Germer, 2014). Darüber hinaus gibt es wenig Forschung über mögliche Auswirkungen des Peer-Umfeldes auf das emotionale Erleben der Schülerinnen und Schüler im Allgemeinen (vgl. Dishion & Tipsord, 2011) und des schulischen Peer-Umfeldes im Besonderen.

Ziel dieser Studie war, erstens das emotionale Erleben von Schülerinnen und Schülern mit internalisierendem Verhalten bei Interaktionen mit ihren Peers direkt im Unterrichtsaltag zu untersuchen. Zweitens wurde überprüft, ob die Klassenkomposition bezüglich Problemverhalten und prosozialem Verhalten den Zusammenhang zwischen dem emotionalen Erleben und Interaktionen mit Peers moderiert.

An der Studie nahmen 719 Jugendliche (MAlter = 12,2; SDAlter = 0,8; 49% Mädchen) aus 40 Klassen der 6. Klassenstufe teil. Mit Hilfe der Experience Sampling Method (ESM) berichteten sie ihr momentanes affektives Erleben und die aktuell ausgeübte Aktivität an M = 12,3 zufällig ausgewählten Zeitpunkten während einer Schulwoche – mit einem Total von 8870 Protokollen. Das aktuelle affektive Erleben wurde mit den beiden Skalen Positive Aktivierung (PA) und Negative Aktivierung (NA) erhoben. Die je vier Items umfassenden Skalen PA und NA wurden von Schallberger (2005) entwickelt und basieren auf dem Circumplex-Modell affektiver Zustände von Watson und Tellegen (1985). Verhaltensprobleme und das prosoziale Verhalten der Jugendlichen wurden mit der Lehrerversion des Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001) erhoben.

Zunächst wurde eine latente Profilanalyse auf Basis der SDQ-Daten durchgeführt, um innerhalb der Schülerschaft Subgruppen mit Verhaltensproblemen zu identifizieren. Anschließend wurden mittels Mehrebenenregressionsanalysen die Struktur affektiven Erlebens (basierend auf den ESM-Daten) sowie Effekte von Peer-Interaktionen und der Klassenkomposition bezüglich Problemverhalten und prosozialem Verhalten auf das affektive Erleben der Schülerschaft als Funktion der Gruppenzugehörigkeit untersucht.

Im Rahmen der latenten Profilanalysen wurden drei Subgruppen identifiziert: Schülerinnen und Schüler mit unproblematischem Verhalten (74%), Schülerinnen und Schüler mit externalisierenden (14%) sowie solche mit internalisierenden Verhaltensweisen (12%). Die Struktur affektiven Erlebens erwies sich für Schülerinnen und Schüler mit internalisierendem und unproblematischem Verhalten vergleichbar. Demgegenüber zeigten Schülerinnen und Schüler mit externalisierendem Verhalten eine höhere intraindividuelle Variabilität, zudem erlebten sie generell weniger PA und mehr NA. Allgemein betrachtet hatte das gemeinsame Arbeiten mit Peers im Unterricht im Vergleich zu Einzelarbeit einen positiven Effekt auf das affektive Erleben der Schülerinnen und Schüler (PA+, NA-). Der positive Effekt war bei Schülerinnen und Schülern mit internalisierendem Verhalten (PA = .21; NA = -.24) und externalisierendem Verhalten (PA = .37; NA = -.23) ausgeprägter ( $p < .05$ ) als bei ihren Klassenkameraden mit unproblematischem Verhalten. Hinsichtlich der Kompositionseffekte im Unterricht konnten nur Tendenzen ( $p < .10$ ) des prosozialen Verhaltens von Klassenkameraden auf PA während der Partner- oder Gruppenarbeit festgestellt werden. Für Schülerinnen und Schüler mit internalisierendem Verhalten war der positive Effekt des gemeinsamen Arbeitens mit Peers in Klassen mit weniger prosozialem Verhalten etwas stärker ausgeprägt als in vergleichsweise prosozialen Klassen.

Die Ergebnisse dieser Studie deuten darauf hin, dass das gemeinsame Arbeiten mit Peers von Schülerinnen und Schülern mit internalisierendem Verhalten mit einem positiven momentanen affektiven Zustand assoziiert wird. Die Bedeutung von Peer-Interaktionen im Unterricht für Schülerinnen und Schüler mit internalisierendem Verhalten wird diskutiert und weiterführende Forschungsfragen werden aufgezeigt.

## E21

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Bernoullianum 30, H148

### Umgang mit professionellen Herausforderungen im Lehrerberuf und dessen Auswirkungen auf Verbleib und Wohlbefinden im Beruf

Chair(s): **Catherine Eve Bauer** (Pädagogische Hochschule Bern, Schweiz)

DiskutantIn(nen): **Johannes Mayr** (Alpen-Adria-Universität Klagenfurt)

Viel wurde in den letzten Jahrzehnten geforscht und geschrieben zu Belastung und Beanspruchung von Lehrpersonen (z.B. Klusmann, 2011; Schaarschmidt, 2002; van Dick & Stegmann, 2013) bis hin zum Burnout (Poschkamp, 2008). In den letzten Jahren rückte neben diesen eher defizitorientierten Ansätzen auch die Untersuchung positiver Arbeitsempfindungen wie Wohlbefinden, Sinnerleben und Glück zunehmend in den Fokus (Day & Qing, 2009; Gehrman, 2013). Auch die berufsbio-graphische Entwicklung und die individuellen Berufsverläufe von Lehrpersonen wurden im Kontext der Frage nach der Bewältigung beruflicher Herausforderungen vermehrt zum Gegenstand empirischer Untersuchungen (Keller-Schneider & Hericks, 2014; Herzog, 2014). Diese Perspektivenerweiterung ist nicht zuletzt dem Einfluss des breit rezipierten Modells professioneller Lehrkompetenzen von Baumert & Kunter (2006) geschuldet, das neben Wissenskomponenten auch motivationale Orientierungen, Werthaltungen, Ziele und Selbstregulationsaspekte von Lehrpersonen als in gleichem Masse relevante Dimensionen professionellen Handelns voraussetzt. Die Forschung zeigt, dass berufliche Herausforderungen und Beanspruchungen keineswegs mit Belastung gleichzusetzen sind, sondern durchaus mit einer hohen Berufszufriedenheit einhergehen können (Gehrman, 2013), sofern die notwendigen Bedingungen für eine erfolgreiche Bewältigung in Form von günstigen Arbeitsbedingungen und persönlichen Handlungsvoraussetzungen gegeben sind (Rudow, 1994). In einer positiven Rückkoppelungsschleife trägt die erfolgreiche Bewältigung beruflicher Herausforderungen zum weiteren Kompetenzaufbau bei (Keller-Schneider & Hericks, 2014) und führt zu grösserer Berufszufriedenheit und Berufstreue (Herzog, Herzog, Brunner, & Müller, 2007; Rudow, 1994, 1999). Auch positive Wirkungen des beruflichen Wohlbefindens auf die Unterrichtsqualität konnten nachgewiesen werden (Frenzel et al., 2016).

Doch welches sind die Bedingungen, die zur erfolgreichen Bewältigung von Herausforderungen beitragen und das subjektive Wohlbefinden und Sinnerleben im Beruf fördern? Im Symposium wird diese Frage aus verschiedenen Perspektiven untersucht und diskutiert:

1. Kathi V. Wachnowski (M.Ed.), Ruhr-Universität Bochum

Lehrer'glück' als Ergebnis professionellen Handelns von Lehrkräften. Präsentiert werden Auszüge aus der Studie „Glück im Lehrerberuf“ zur Entstehung und Wirkung subjektiven Wohlbefindens von Lehrerinnen und Lehrern. Der Fokus der Untersuchung mit der Day Reconstruction Method liegt auf folgenden Fragen: Wie hängen die Tätigkeiten von Lehrpersonen mit ihrem subjektiven Wohlbefinden zusammen? Wie werden diese Zusammenhänge von Persönlichkeit, berufsbezogenen und soziodemographischen Aspekten beeinflusst?

2. Dr. Sören Umlauf, Universität Halle

Arbeitsorientierungen, Arbeitszufriedenheit und subjektives Wohlbefinden von Lehrpersonen. Für die arbeitsbezogene Sinnerfüllung spielen Arbeitsorientierungen eine zentrale Rolle; insbesondere die Calling-Orientierung kann als Ressource für Lehrpersonen verstanden werden. Präsentiert werden die Ergebnisse einer Fragebogenstudie zu den Zusammenhängen zwischen Calling-Orientierung, Variablen des subjektiven Wohlbefindens und der Arbeitszufriedenheit.

3. Dr. Larissa Trösch & Dr. Catherine Bauer, Pädagogische Hochschule Bern

Berufliche Herausforderungen, Bewältigung und Berufsverbleib. Präsentiert werden Auszüge aus der Studie „Berufsleute als Lehrpersonen“, in der Lehrkräfte auf dem zweiten Bildungsweg zu beruflichen Herausforderungen der ersten Berufsjahre, subjektivem Wohlbefinden und Berufsverbleib befragt wurden. Präsentiert werden die Ergebnisse einer inhaltsanalytischen Auswertung von Leitfadenterviews mit noch aktiven und wieder ausgeschiedenen Lehrpersonen auf dem zweiten Bildungsweg.

Es werden qualitative und quantitative Befunde zur gemeinsamen Fragestellung präsentiert und im Hinblick auf die zugrundeliegenden theoretischen Bezüge wie auch auf ihre Bedeutung für die Lehrerinnen- und Lehrerbildung sowie die individuelle Laufbahngestaltung diskutiert.

#### Beiträge des Symposiums

### Lehrer'glück' als Ergebnis professionellen Handelns von Lehrkräften. Auszüge aus der Studie „Glück im Lehrerberuf“ zur Entstehung und Wirkung subjektiven Wohlbefindens von Lehrerinnen und Lehrern

**Kathi V. Wachnowski**

Ruhr-Universität Bochum

#### Theoretischer Hintergrund

Lehrerinnen und Lehrer (LuL) sind in ihrem beruflichen Alltag mit vielfältigen Herausforderungen, unklaren Situationen und unterschiedlichsten Aufgaben konfrontiert (z.B. Fend, 1998; Helsper, 2002; Rothland & Terhart, 2007). Die von ihnen zu bewältigenden Tätigkeiten gehen dabei weit über rein unterrichtsbezogene Aktivitäten hinaus. Bisherige Studien, in denen sich Informationen zu Tätigkeiten von LuL finden lassen, stammen zumeist aus dem Diskurs zu Lehrbelastung und der (übermäßigen) Arbeitszeit von LuL (Mummert & Partner, 1999; Schaarschmidt, 2002) und sind vielfach programmatisch geprägt (z.B. die von der GEW in Auftrag gegebene niedersächsische Arbeitszeitstudie; Mußmann, Riethmüller & Hardwig, 2016). Häufig beziehen sie ihre Daten zudem aus Single-Time-Questionnaires und fokussieren die Art und Dauer von Tätigkeiten. Erhebungen, die LuL über eine gewisse Zeit begleiten und damit zur Minimierung von (zeitlichen) Verzerrungseffekten (z.B. retrospective reporting biases; Hektner et al., 2007; Kahnemann et al., 2004) beitragen sowie Untersuchungen, die weitere Tätigkeitsaspekte wie beispielsweise den Ort und das soziale Setting berücksichtigen, stellen die Ausnahme dar bzw. sind nicht vorhanden. Darüber hinaus wird in den Studien vorwiegend ein defizitorientierter Blick auf die Bedeutung der Arbeitstätigkeit als Gefahr für die eigene Gesundheit eingenommen. Insgesamt gibt der aktuelle Stand der Forschung damit nur begrenzt Einblick in die Komplexität professionellen Handelns von Lehrkräften im beruflichen Alltag.

#### Fragestellung

Im vorliegenden Beitrag werden ausgewählte Ergebnisse der Studie „Glück im Lehrerberuf“ vorgestellt. Die Studie nähert sich dem Feld aus einer ergänzenden Perspektive, in dem sie das Zusammenspiel von Tätigkeiten / Selbstorganisation mit dem subjektiven Wohlbefinden (SWB) von LuL in den Fokus des Erkenntnisinteresses stellt. SWB wird dabei in Anlehnung an Diener (1984) definiert als Zusammenspiel von affektiven Komponenten (positiver und negativer Affekt) und kognitiven Komponenten (allgemeine Lebenszufriedenheit, domänenspezifische Zufriedenheiten, professionsspezifische Zufriedenheiten). Zentrale Fragestellungen der Studie sind: Was tun LuL und wie hängt die mit ihrem SWB zusammen? Wie werden diese Zusammenhänge von Persönlichkeit, berufsbezogenen und soziodemographischen Aspekten beeinflusst?

#### Methodik

Zur Klärung der Fragestellungen wurde ein mehrstufiges Erhebungsdesign genutzt: In der Eingangsbefragung wurden ausführliche Informationen zum SWB und zur Persönlichkeit (Big Five) sowie soziodemographische und berufsbezogene (formale Charakteristika der LuL und Schulen, motivationale Aspekte der Arbeitsgestaltung) Daten erhoben. Es folgte eine einwöchige Phase, in der die LuL täglich einen Tagebucheintrag zur Erfassung von Tätigkeiten (Was? Wann? Wo? Mit wem?) und SWB im Tagesverlauf verfassten (Day Reconstruction Method). Optional und ergänzend konnten die Teilnehmenden sich dafür entscheiden, täglich vier Momentaufnahmen zu beantworten, in denen sie ebenfalls zu ihren Tätigkeiten, dem SWB und dem Flow-Erleben befragt wurden (Experience Sampling Method). Zur Kontrolle der instrumentellen Effekte schloss sich an den letzten Tage-bucheintrag eine erneute Erfassung der SWB-Batterie aus der Eingangsbefragung an. LuL hatten die Möglichkeit, online oder über die für die Studie entwickelte App teilzunehmen. Die Studie richtete sich an Lehrkräfte in Deutschland und Österreich (N=877). Studienteilnehmerinnen und -teilnehmer wurden über unterschiedliche Kanäle (Anschreiben an eine repräsentative Auswahl von Schulen, Werbung über Verteiler der großen Gewerkschaften, in Lehrerforen und -zeitschriften) akquiriert. Die erreichte Stichprobe wurde einem Vergleich mit statistischen Daten zur Grundgesamtheit unterzogen.

#### Ergebnisse und Diskussion

Im Rahmen des Vortrags wird unter Nutzung der erhobenen Daten das Bedingungsmodell des SWB von LuL vorgestellt. Besondere Berücksichtigung findet dabei die differenzierte Betrachtung von Tätigkeiten und der Wirkung der individuellen Arbeitsgestaltung / Selbstorganisation (Wann und wo mache ich was mit wem?) auf das SWB. Die vorgestellten Ergebnisse werden in Hinblick auf das Verständnis von Selbstregulation als Teilfacette der professionellen Kompetenz von Lehrkräften diskutiert (Baumert & Kunter, 2006). Zudem werden Überlegungen zu praktischen Implikationen der Befunde für die Aus-, Fort- und Weiterbildung von Lehrerinnen und Lehrern getätigt.

### **Arbeitsorientierungen, Arbeitszufriedenheit und subjektives Wohlbefinden von Lehrpersonen**

#### Sören Umlauf

Martin-Luther-Universität Halle-Wittenberg

Theoretische Ansätze, die das Sinnerleben im Beruf thematisieren, stimmen im Grundsatz darin überein, dass das Erleben der Arbeitstätigkeit eine zentrale Rolle spielt. Hackman und Oldham (1976) erklären dies mit objektiven Charakteristika von Jobs, wie Anforderungsvielfalt, Ganzheitlichkeit und Bedeutung von Arbeitsaufgaben. Wrzesniewski, Dutton und Debebe (2003) nehmen an, dass jeder Job sinnhaft erlebt werden kann, wenn im all-täglichen Arbeitsgeschehen erfahrbar wird, dass die Umwelt dem Inhalt der Arbeit, den damit verbundenen Rollen und dem individuellen Selbst Wert und Bedeutung zuschreibt. Letzterer Ansatz legt damit nahe, dass arbeits-bezogene Sinnerfüllung vor allem von interindividuellen Unterschieden und kontextuellen Faktoren abhängt. Im Hinblick auf interindividuelle Unterschiede könnten Arbeitsorientierungen eine zentrale Rolle spielen. Unter diesem Begriff werden drei relativ stabile Einstellungen zur eigenen Berufstätigkeit verstanden (Job, Career, Calling), welche sich durch verschiedene Zielschwerpunkte auszeichnen (Wrzesniewski, McCauley, Rozin, & Schwartz, 1997). Während bei der Job-Orientierung der materielle Nutzen und bei der Career-Orientierung der berufliche Aufstieg im Vordergrund stehen, leitet sich die Calling-Orientierung von der Idee der „Berufung“ ab, was im Kern meint, dass den jeweiligen Arbeitsaufgaben an sich Wert und Bedeutung beigemessen werden und der eigene Beruf eine zentrale Stellung im Leben einnimmt (Wrzesniewski, 2003). In diesem Beitrag wird vorgeschlagen, die Calling-Orientierung als Ressource für Lehrpersonen aufzufassen. Dafür spricht vor allem der Befund, dass Calling mit höherer Arbeitszufriedenheit assoziiert ist, wobei Arbeit als Quelle von Zufriedenheit und Erfüllung außerdem stärker gewichtet wird als andere Lebensdomänen (Wrzesniewski et al., 1997; Wrzesniewski, 2003). Folgerichtig findet sich auch eine positive Beziehung zur Lebenszufriedenheit, was die Vermutung nahe legt, dass die Calling-Orientierung arbeitsbezogene Sinnerfüllung erklärt und damit zu subjektivem Wohlbefinden beiträgt (Wrzesniewski et al., 1997). Für diese Erklärung spricht, dass Calling die Suche nach subjektiv sinnstiftenden Tätigkeiten vorhersagt (längsschnittlich), also mit proaktiver Sinnsuche im Beruf einherzugehen scheint (Wrzesniewski & Dutton, 2001).

Mit der vorliegenden Studie sollten die drei Arbeitsorientierungen anhand eines neu erstellten, deutschsprachigen Fragebogens bei berufstätigen Lehrpersonen untersucht werden. Dazu wurden Lehrpersonen mehrerer Schulen in Sachsen-Anhalt einmalig schriftlich befragt (N = 102; weiblich: 75; MAlter = 47.7 Jahre). Eine erste Fragestellung war, wie die drei Arbeitsorientierungen bei aktiven Lehrpersonen durchschnittlich ausgeprägt sind. Die zweite Fragestellung war, ob die bekannten, günstigen Zusammenhänge zwischen der Calling-Orientierung und der Arbeitszufriedenheit sowie dem subjektiven Wohlbefinden repliziert (Lebenszufriedenheit) und erweitert (positive und negative Aktivierung) werden können. Aufgrund der Hypothese arbeitsbezogener Sinnerfüllung, wurde hierbei außerdem der subjektive Lebenssinn einbezogen (Schnell & Becker, 2007). Die dritte Fragestellung war, ob die Zusammenhänge zwischen der Calling-Orientierung und den Variablen des subjektiven Wohlbefindens durch die Arbeitszufriedenheit vermittelt werden. Dabei wurde zwischen tätigkeitsbezogener und arbeitsplatzbezogener Arbeitszufriedenheit unterschieden und angenommen, dass insbesondere die tätigkeitsbezogene eine zentrale Rolle spielt.

Es zeigte sich, dass die Calling-Orientierung substantielle, positive Zusammenhänge zu Arbeitszufriedenheit und den positiven Variablen des subjektiven Wohlbefindens aufweist (Lebenszufriedenheit, positive Aktivierung, Sinnerfüllung), aber nur geringfügig oder gar nicht mit den negativen (negative Aktivierung, Sinnkrise) korreliert. Mediationsanalysen stützen die Hypothese einer zentralen Rolle tätigkeitsbezogener Arbeitszufriedenheit bei Kontrolle der Schulform, des Alters und des Geschlechts. Die Ergebnisse werden kritisch diskutiert und auf ihre Implikationen untersucht. Hintergrund ist die Überlegung, dass sowohl Arbeitskontexte als auch Arbeitsori-entierungen von Lehrpersonen im Hinblick auf die Erfüllung des Sinnbedürfnisses entwickelbar sein könnten. Für beide werden Maßnahmen auf institutioneller und individueller Ebene erwogen, wobei auf Job Crafting (Wrzesniewski & Dutton, 2001) sowie das zugrunde liegende Model of Interpersonal Sensemaking (Wrzesniewski, Dutton & Debebe, 2003) zurückgegriffen wird.

## **Herausforderungen, Bewältigungsaspekte und Berufsverbleib von Lehrpersonen auf dem zweiten Bildungsweg**

**Larissa Maria Trösch, Catherine Eve Bauer**  
Pädagogische Hochschule Bern

### Theoretischer Hintergrund und Fragestellung

Der Berufsverbleib von Lehrpersonen ist seit Jahrzehnten ein Dauerbrennerthema in Öffentlichkeit und Wissenschaft. Die vielfältigen Herausforderungen des Lehrberufs und die dadurch gegebenenfalls entstehende hohe Belastung werden unter anderem als Grund für den Berufsausstieg angesehen (Herzog et al., 2007). Eine erfolgreiche Bewältigung beruflicher Beanspruchungssituationen (Lazarus & Folkman, 1984) führt zu einer Stärkung und Erweiterung der vorhandenen Kompetenzen (Keller-Schneider & Hericks, 2014) und wirkt sich positiv auf das berufliche Wohlbefinden und den Berufsverbleib aus, während fehlende Ressourcen und Bewältigungsprobleme mit Gefühlen der subjektiven Belastung und Überlastung verbunden sind (Herzog, Herzog, Brunner, & Müller, 2007; Rudow, 1999). Der erfolgreiche Umgang mit diesen Herausforderungen setzt jedoch breite Kompetenzen (Baumert & Kunter, 2006) und Ressourcen voraus, die nicht nur im Rahmen der institutionellen Lehrerinnen- und Lehrerbildung erworben werden, sondern sich über die gesamte Bildungs- und Berufslaufbahn akkumulieren (Cramer, 2012).

In diesem Kontext ist die professionelle Entwicklung von Lehrkräften auf dem zweiten Bildungsweg von Bedeutung. Um den wiederkehrenden Lehrermangel abzufedern, werden in der Lehrerinnen- und Lehrerbildung seit Jahrzehnten Programme ins Leben gerufen, die Personen mit Berufserfahrungen anziehen sollen (Criblez, 2017). Obwohl sie somit eine wichtige Zielgruppe für den Lehrberuf darstellen, sind die weitere professionelle Entwicklung dieser Lehrkräfte und die Bedingungen des weiteren Berufsverbleibs oder -ausstiegs nur sehr beschränkt untersucht. Der folgende Beitrag befasst sich mit der Bewältigung der beruflichen Herausforderung von Lehrpersonen auf dem zweiten Bildungsweg und wie diese mit dem Berufsverbleib oder -ausstieg zusammenhängt. Von besonderem Interesse dabei ist, ob sich Lehrpersonen auf dem zweiten Bildungsweg, die den Lehrberuf wieder verlassen haben, von im Lehrberuf Verbliebenen in Bezug auf die wahrgenommenen Herausforderungen und deren Bewältigung unterscheiden.

### Methodik

Als Datenbasis dienten problemzentrierte Interviews mit 23 ausgebildeten Lehrpersonen verschiedener Stufen der obligatorischen Schule. Alle Befragten hatten sieben bis zehn Jahre vor dem Befragungszeitpunkt die Pädagogische Hochschule absolviert und vor der Lehrerbildung einen Vorberuf erlernt.

Die Interviews wurden leitfadengestützt durchgeführt. Die Auswertung erfolgte anhand der qualitativen Inhaltsanalyse nach Mayring (2003). Das Kategoriensystem wurde nach deduktiv-induktivem Vorgehen erstellt und orientiert sich am transaktionalen Stressmodell nach Lazarus und Folkman (1984).

### Ergebnisse

Die Zusammenarbeit in Kollegium und Schulleitung wurde von sämtlichen Befragten mit deutlichem Abstand als häufigste Herausforderung genannt, wobei diese Kategorie von den Aussteigerinnen und Aussteigern noch dominanter angesprochen wurde als von den Berufsverbleibenden. Bei den Aussteigern zeigten sich besonders häufig Unsicherheiten in Bezug auf die eigene Leistung als Herausforderungsthematik sowie überhöhte Erwartungen und Ansprüche an sich selbst. Der Wegbruch der gewohnten beruflichen Leitplanken (Zusammenarbeit im Team, regelmässiges Feedback, sichtbare Arbeitsergebnisse) war für beide Verbleibsgruppen eine zentrale Herausforderung. Hingegen berichteten Aussteiger und Aussteigerinnen häufiger über Schwierigkeiten sich zu regenerieren als Berufsverbleibende.

Die Verbleibenden zeigten in den Interviews ein grösseres Spektrum an Bewältigungsthemen und -strategien im Vergleich zu den Aussteigerinnen und Aussteigern. Während letztere als Bewältigungsstrategien vornehmlich das Inanspruchnehmen von Hilfe, Abwarten, Stellenwechsel sowie Widerstand nannten, setzten die Verbleiberinnen und Verbleiber vermehrt auf weitere Strategien wie Kommunikation und Veränderung der beruflichen Rahmenbedingungen oder der eigenen Arbeitsstrategien. Die Korrektur der eigenen Erwartungen an den Beruf wurde von beiden Gruppen als Bewältigungsthema mehrfach angesprochen, bei den Aussteigerinnen und Aussteigern anteilmässig etwas häufiger.

Die Resultate werden im Zusammenhang ihrer Relevanz für die Berufsvorbereitung und -einführung, das subjektive Wohlbefinden von Lehrpersonen auf dem zweiten Bildungsweg sowie ihre berufliche Entwicklung diskutiert.

## E22

Zeit: Freitag, 16.02.2018: 15:30 - 17:15 · Ort: Musikwissenschaftliches Seminar, Hörsaal 006

### Studien zum Zusammenhang von Lerngelegenheiten und Kompetenzentwicklungen in der universitären Lehrerbildung

Chair(s): **Jörg Doll** (Universität Hamburg, Deutschland), **Gabriele Kaiser** (Universität Hamburg)

DiskutantIn(nen): **Nils Buchholtz** (Universität Oslo)

Das Symposium vereint Beiträge von vier Hochschulen, in denen Lerngelegenheiten – opportunities to learn – als wesentliche Prozessmerkmale der Lehr-Lern-Angebote im Lehramtsstudium konzipiert werden. Lerngelegenheiten beziehen sich auf institutionelle Ausbildungsangebote, von denen angenommen wird, dass sie einen bedeutsamen Einfluss auf den Erwerb professioneller Wissens- und Kompetenzfacetten haben (Tatto et al., 2008). Die Erhebung spezifischer bildungswissenschaftlicher, fachdidaktischer und schulpraktischer Lerngelegenheiten im Lehramtsstudium durch Einschätzungen der Studierenden und die Analyse ihres Zusammenhangs zu Lernergebnissen im Studium ermöglichen Einblicke in die Effektivität von Curricula und Möglichkeiten der Qualitätssicherung (König et al., 2010).

Der Beitrag von Klemenz, Ligtvoet, König und Rothland untersucht den Zusammenhang zwischen bildungswissenschaftlichen Lerngelegenheiten und dem Zuwachs pädagogischen Wissens. Die Daten wurden zu zwei Zeitpunkten, einmal zum Studienbeginn und ein zweites Mal im Abstand von vier Semestern, erhoben. In einem Zwei-Ebenen-Design werden Einflüsse auf individueller und institutioneller Ebene modelliert. Der Beitrag von Kaiser, Jentsch, Doll und König dokumentiert die Nutzungsverläufe fachdidaktischer Lerngelegenheiten in den Fächern Deutsch, Englisch und Mathematik im Lehramtsstudium. Der Beitrag von Hany, Protzel und Bock geht der Frage nach, welche praktischen Lerngelegenheiten die Lehramtsstudierenden in einem Halbjahrespraktikum, als komplexem Schulpraktikum, nutzen und wie diese Nutzung mit den Kompetenzzuwächsen in Beziehung steht. Die kompetenzbezogenen Selbsteinschätzungsskalen werden nach dem Stufenmodell der Expertise von Dreyfus und Dreyfus (Dreyfus, 2004) konstruiert. Die Daten werden in einem Prä-Posttest-Design erhoben. Auch der Beitrag von Doll, Meyer, Kaiser und König widmet sich den lernbezogenen Aktivitäten in Schulpraktika. Basierend auf dem Modell lernprozessbezogener Aktivitäten von König et al. (2014) werden die Bachelor- und Masterpraktika an zwei Hochschulen daraufhin untersucht, ob sie die Aktivitäten „Theorien auf Situationen beziehen“ unterschiedlich stark ermöglichen.

*Beiträge des Symposiums*

### Zur Wirksamkeit von Lerngelegenheiten auf den Erwerb pädagogischen Wissens

**Stefan Klemenz<sup>1</sup>, Rudy Ligtvoet<sup>1</sup>, Johannes König<sup>1</sup>, Martin Rothland<sup>2</sup>**

<sup>1</sup>Universität Köln, <sup>2</sup>Universität Siegen

Die Lehrer\*innenausbildung ist seit Jahren Gegenstand zahlreicher Reformdiskussionen. Allerdings mangelt es noch immer an empirischen Studien, die differenzierten Einblick in die Wirkungsweise von bestehenden Curricula auf relevante Ergebnisvariablen angehender Lehrer\*innen geben (König et al., 2017). Modellhafte Vorstellungen zur empirischen Untersuchung der Bedingungen von professioneller Kompetenz im Sinne einer Wirksamkeit der Lehrer\*innenausbildung betonen die Bedeutung von Prozessmerkmalen zur Beschreibung der Lehr-Lern-Arrangements, die angehende Lehrkräfte unter der Zielperspektive durchlaufen, um berufsrelevantes Wissen und Können zu erwerben und aufzubauen (Blömeke, 2011). „Lerngelegenheiten“ – opportunities to learn (OTL) – können als wesentliche Prozessmerkmale betrachtet werden und beziehen sich auf institutionelle Ausbildungsangebote, von denen angenommen wird, dass sie einen bedeutsamen Effekt auf den Erwerb professioneller Kompetenzen haben (Tatto et al., 2008). Die Erhebung von Lerngelegenheiten und die differenzierte Analyse ihres Zusammenhangs zu Lernergebnissen wie dem professionellen Wissen angehender Lehrer\*innen, ermöglichen Einblicke in die Effektivität von Curricula und Möglichkeiten ihrer Qualitätssicherung (König, Blömeke & Kaiser, 2010).

Die Forschungslage zum Zusammenhang von Lerngelegenheiten und berufsrelevantem Wissen bei angehenden Lehrkräften erweist sich jedoch als relativ dünn (Floden, 2015). Daran anschließend fokussiert der Beitrag den Einfluss von Lerngelegenheiten als Prozessmerkmal auf den Erwerb und Aufbau von pädagogischem Wissen.

Fragestellung

1. Welche bildungswissenschaftlichen Lerngelegenheiten (Inhalte und schulpraktische Tätigkeiten) stellen bedeutsame Prädiktoren für die Veränderung (bzw. den Zuwachs) pädagogischen Wissens während der Bachelor-Lehramtsausbildung dar?
2. Welche Lerngelegenheiten werden dabei nicht nur auf der Individualebene (d.h. der Ebene von Lehramtsstudierenden), sondern auch auf der institutionellen Ebene signifikant?

Methode

Für die Bearbeitung der beiden Fragestellungen wird ein Teildatensatz der EMW-Studie („Entwicklung von berufsspezifischer Motivation und pädagogischem Wissen in der Lehrerausbildung“; König & Rothland, 2013) mit 1347 Lehramtsstudierenden aus 37 Studiengängen von 18 Universitäten bzw. Pädagogischen Hochschulen aus Deutschland und Österreich einbezogen. Sie wurden zu zwei Messzeitpunkten - bei Eintritt in ihr Studium (1. Semester, Herbst 2011) sowie zwei Jahre später (5. Semester, Herbst 2013) - zu ihrem pädagogischen Wissen getestet sowie zu ihren Lerngelegenheiten befragt. Das pädagogische Wissen wurde mit dem TEDS-M Testinstrument erfasst (König et al., 2011). Es kann über vier Subskalen für die Bereiche „Umgang mit Heterogenität“, „Strukturierung von Unterricht“, „Klassenführung/Motivierung“, „Leistungsbeurteilung“ reliabel berichtet und eine Veränderung über beide Messzeitpunkte modelliert werden. Die Erfassung der Lerngelegenheiten erfolgte zum zweiten Messzeitpunkt mit einem Instrument, das umfangreiche Inhaltsabfragen zu den vier inhaltlichen Testbereichen umfasst. Darüber hinaus wurden Skalen zur Erfassung schulpraktischer Tätigkeiten (Unterricht planen, Unterricht durchführen, Theorien auf Situationen beziehen, mit Situationen analytisch-reflektiert umgehen; König et al., 2014) verwendet.

Um den Einfluss von Merkmalen der Lerngelegenheiten auf Ebene der Studierenden (Individualebene) sowie auf Ebene der Studiengänge innerhalb der Universitäten (Institutionenebene) simultan zu prüfen, werden Mehrebenenanalysen durchgeführt.

Ergebnisse

Zwischen den Zeitpunkten zeigt sich in allen vier Subskalen pädagogischen Wissens ein statistisch signifikanter Wissenszuwachs. Ferner liegen die ICCs zum zweiten Messzeitpunkt bei .10 oder höher, weshalb eine differenzierte Untersuchung anhand von Mehrebenenanalysen empfehlenswert erscheint.

Ergebnisse zu Frage 1: Die Inhaltsskalen sowie die schulpraktischen Lerngelegenheiten erweisen sich in Veränderungsmodellen als statistisch signifikant. Ergebnisse zu Frage 2: Allein die Inhaltsskalen zu den schulpädagogisch geprägten Bereichen „Umgang mit Heterogenität“ und „Strukturierung von Unterricht“ erweisen sich sowohl auf Individual- als auch auf Institutionsebene als signifikante Prädiktoren zur Veränderung pädagogischen Wissens in den entsprechenden Bereichen Heterogenität bzw. Strukturierung. Dagegen erweisen sich die übrigen, eher pädagogisch-psychologisch geprägten Inhaltsskalen „Klassenführung/Motivierung“ und „Leistungsbeurteilung“ sowie die schulpraktischen Skalen allein auf Individualebene als bedeutsame Prädiktoren für pädagogisches Wissen.

Die Modellierung des Einflusses von Lerngelegenheiten auf den Zuwachs pädagogischen Wissens sowie die Implikationen für die Profilierung von bildungswissenschaftlichen Curricula werden zur Diskussion gestellt.

## **Zur Nutzung fachdidaktischer Lerngelegenheiten im Lehramtsstudium in den Fächern Deutsch, Englisch und Mathematik**

**Gabriele Kaiser<sup>1</sup>, Armin Jentsch<sup>1</sup>, Jörg Doll<sup>1</sup>, Johannes König<sup>2</sup>**

<sup>1</sup>Universität Hamburg, <sup>2</sup>Universität Köln

Das Wissen und die Kompetenzen, über die Lehrpersonen am Ende ihrer Ausbildung verfügen, sind eng verknüpft mit Art, Umfang und Qualität der Lerngelegenheiten in dieser Ausbildung. Lerngelegenheiten werden auf drei Ebenen als intendiertes, implementiertes und erreichtes Curriculum beschrieben (Kunina-Habenicht et al., 2013; McDonnell, 1995). Das intendierte Curriculum besteht aus dem institutionell vorgegebenen Bildungsplan, das implementierte Curriculum beschreibt, welche Lehr-Lern-Angebote die Lehrenden entwickeln und anbieten, welche Aktivitäten die Studierenden zeigen und welche Lernprozesse sie durchlaufen, und das erreichte Curriculum umfasst erworbenes Wissen, Kompetenzen und berufliche Selbstkonzeptfacetten. Vor diesem theoretischen Hintergrund werden Ergebnisse der ersten und zweiten Welle des ProfaLe-Panels an der Universität Hamburg zur Nutzung fachdidaktischer Lerngelegenheiten durch Lehramtsstudierende der Fächer Deutsch, Englisch und Mathematik im Studienverlauf berichtet. ProfaLe („Professionelles Lehrhandeln zur Förderung fachlichen Lernens“) ist Teil der „Qualitätsoffensive Lehrerbildung“ und das an der Universität Hamburg angesiedelte Projekt.

Untersucht wird, wie sich das Nutzungsverhalten für fachdidaktische Studieninhaltsbereiche im Studienverlauf entwickelt. Erfragt wurde die Nutzung der Studieninhalte bezogen auf die gesamte bisherige Ausbildung, sodass ein monotoner Nutzungsanstieg im Studienverlauf ein wahrscheinliches Nutzungsmuster beschreibt. Möglich ist aber auch ein Nutzungsverlauf mit einem Anstieg ausschließlich im Bachelorstudium ohne weitere Zunahme im Masterstudium. Eine zweite methodische Fragestellung untersucht, wie sich die Art der Erhebung des Nutzungsverhaltens auf die Nutzungsmuster auswirkt. Während das Nutzungsverhalten in Welle 1 dichotom erhoben wurde, wurde in Welle 2 eine 7-stufigen Skala der Nutzungsintensität eingesetzt.

Die deutschdidaktischen Studieninhaltsbereiche wurden mit insgesamt 31 Items erhoben, die von König, Bremerich-Vos et al. (2017) entwickelt worden sind: „Sprechen/Zuhören“, „Schreiben“, „Lesen“, „Literaturunterricht“, „Sprachreflexion“ und „Curriculare Aspekte“; sieben englischdidaktische Studieninhaltsbereiche mit 86 Items stammen aus dem PKE-Projekt (Professionelle Kompetenzen von angehenden Englischlehrkräften; König, Tachtsoglou et al., 2017): „Sprachdidaktik“, „Literaturdidaktik“, „Kulturdidaktik“, „Mediendidaktik“, „Methodik“, „Curriculum/Beurteilung“ und „Fachdidaktische Forschung“ und sechs mathematikdidaktische Studieninhaltsbereiche mit 61 Items wurden von Buchholz und Doll (2017) konstruiert: „Didaktische Grundlagen“, „Umgang mit Heterogenität“, „Moderne Medien“, „Methodik“, „Curriculare Aspekte“ und „Fachdidaktische Forschung“.

Das Projekt der Begleitforschung zu ProfaLe ist als Panelstudie mit drei Wellen in den Jahren 2016, 2017 und 2018 konzipiert. Die Datenerhebung erfolgt durch online-Befragungen am Ende des jeweiligen Sommersemesters. An jeder Welle nehmen fünf Studierendengruppen teil: Lehramtsstudierende im 2., 4. oder 6. Semester des Bachelorstudiums und im 2. und 4. Semester des Masterstudiums. An der Welle 1 im Sommer 2016 beteiligten sich Studierende der Fächer Deutsch (n = 263), Englisch (n = 227) und Mathematik (n = 187). Die Daten der Welle 2 im Sommer 2017 werden aktuell aufbereitet. Die Nutzungsdaten werden als Multikohorten-Sequenzdesign ausgewertet. Als charakteristisch für alle drei Fächer ergaben sich in der Welle 1 große signifikante Nutzungsanstiege im Bachelorstudium. Weitere signifikante Nutzungsanstiege im Masterstudium traten ausschließlich im Fach Englisch in der „Sprachdidaktik“ und der „Methodik“ auf.

## **Subjektiver Kompetenzzuwachs im Langzeitpraktikum als Folge spezifischer, allgemein und individuell genutzter Lerngelegenheiten**

**Ernst Hany, Madlen Protzel, Thomas Bock**

Universität Erfurt

Angesichts der bezweifelbaren Wirkung von Praxisphasen im Rahmen der Lehrer\*innenbildung tun jene Hochschulen, die ihre Praxiskonzepte modifizieren, gut daran, die Effekte dieser Veränderungen zu evaluieren. Auch wenn die Verwendung von Standards der Wirkungsforschung (z.B. eines randomisierten Kontrollgruppendesigns) bei Studien an realen Studierendengruppen nur eingeschränkt möglich ist (Scriven, 2009), gilt es, die theoretische und methodische Solidität der Erhebung zu gewährleisten. Zur Prüfung der Wirksamkeit des neu eingeführten Praxissemesters an der Universität Erfurt wurden deshalb gängige Kompetenz- bzw. Angebots-Nutzungsmodelle (Hascher & Kittinger, 2014) und empirisch bewährte Messkonzepte (Meier, 2015) herangezogen sowie vier Praktikumsgruppen, die im Halbjahreszyklus zwischen 2015 und 2017 nacheinander im Einsatz waren, zur empirischen Wirkungsprüfung in einem Prätest-Posttest-Design berücksichtigt.

Da es im Sinne der Evaluationsfunktion der Studie erforderlich war, Studierende der Grundschule und der Regelschule sowie alle im Praktikum berücksichtigten Unterrichtsfächer einzubeziehen, mussten schulart- und fachübergreifende Kompetenzerfassungen eingesetzt werden, für die jedoch keine objektiven Tests vorliegen. Deshalb wurden subjektive Kompetenzeinschätzungen verwendet, die wegen des möglichen Overconfidence-Effekts in Verruf geraten sind (Kunter & Klusmann, 2010). Allerdings zeigt eine Reihe von Studien, dass Selbsteinschätzungen von Kompetenzen bei ausreichender Erfahrung und enger Fassung des Kompetenzbereichs präzise sein können (Ackerman & Wolman, 2007; Hansson, 2001). Ferner konnte gezeigt werden, dass die Zusammenhänge mit der realen Leistung valider sind (Chamorro-Premuzic & Furnham, 2006) als beispielsweise die Zusammenhänge von professionellen Unterrichtsbeurteilungen bei Lehrkräften und den Lernfortschritten der Schulklassen (Kane & Staiger, 2012).

Der spezielle Untersuchungsansatz der durchgeführten Evaluationsstudie geht davon aus, dass der Umfang der praktischen Lerngelegenheiten eines zu einem erheblichen Teil selbstverantworteten Praxiseinsatzes an sich – nach dem opportunities to learn-Ansatz (Carroll, 1963) – zu einem (generellen) Kompetenzzuwachs führt, der bei den Erfurter Studierenden jedoch nicht erst mit dem Praxissemester beginnt, da diesem kürzere Praktika mit gestuften Ansprüchen vorangehen. Da die Studierenden

nicht alle Handlungsanforderungen in gleichem Umfang erleben dürften, wurden Unterschiede in den Lerngelegenheiten und den Kompetenzzuwächsen in Abhängigkeit vom Kompetenzfeld erwartet. Ferner sind aufgrund der individuellen Unterschiede der Studierenden hinsichtlich ihrer Studieninteressen und Lernmotivation (Richardson, Abraham & Bond, 2012), ihrer persönlichen Zielbildung im Zusammenhang mit dem Studium (Neave & Tramm, 2007) und ihrer Interaktion mit der Schule und den praktikumsbegleitenden Lehrkräften (Hoffman et al., 2015) wesentliche individuelle Unterschiede in den erlebten Lerngelegenheiten und den damit verbundenen Kompetenzveränderungen zu erwarten.

#### Fragestellung

Untersucht werden sollte, in welchem Umfang bereichsspezifische Lerngelegenheiten im Verlauf eines Halbjahrespraktikums berichtet werden. Dabei wurden theoretische und praktische Lerngelegenheiten unterschieden. Erwartet wurde ein erheblicher Anstieg an praktischen und kein nennenswerter Zuwachs an theoretischen Lerngelegenheiten, bezogen auf das gesamte Studium. Ferner sollte untersucht werden, inwieweit sich Zuwächse in den selbst erlebten Kompetenzen ergeben, inwieweit diese bereichsspezifisch ausfallen und wie sehr diese allgemein wie differentiell von den erlebten Lerngelegenheiten abhängen. Die Rolle der aktiven Nutzung möglicher Lerngelegenheiten und die Wechselwirkungen mit erlebter Selbstwirksamkeit wurden ergänzend untersucht.

#### Methode

In den längsschnittlich nutzbaren Datensatz mit jeweils zwei Messzeitpunkten konnten 423 Studierende aus zwei Studierendenjahrgängen einbezogen werden, von denen jeweils die Hälfte im Sommer- bzw. im Wintersemester das „Komplexe Schulpraktikum“ (Erfurter Bezeichnung des Praxissemesters) absolvierten. Für die Einschätzungen der Lerngelegenheiten und Kompetenzen wurden keine Likert-Skalen verwendet, sondern kompetenzbezogene Einordnungen nach dem Stufenmodell der Expertise von Dreyfus und Dreyfus (Dreyfus, 2004) konstruiert. Die betrachteten Kompetenzfacetten folgen der empirischen Strukturanalyse von Meier (2015), wobei schwer verständliche Items durch griffigere Formulierungen aus anderen Studien (z. B. Thiel & Blüthmann, 2009) ausgetauscht wurden.

#### Ergebnisse

Wie erwartet begannen die Studierenden das Praxissemester mit maßvollen und bereichsspezifisch stark variierenden Selbstzuschreibungen ihrer handlungsbezogenen Kompetenzen. Das Praktikum erbrachte durchweg erhebliche praktische Lerngelegenheiten und einen bereichsspezifisch differenzierten, dem Ausmaß an Lerngelegenheiten entsprechenden Kompetenzzuwachs. Daneben lassen sich erhebliche individuelle Unterschiede in der Nutzung der Lerngelegenheiten und den Kompetenzzuwächsen beobachten, die aber interindividuell wieder hoch korreliert sind.

### **Zur Nutzung schulpraktischer Lerngelegenheiten an zwei deutschen Hochschulen: die Anwendung von theoriebasiertem Wissen auf pädagogische Situationen in Praktika**

**Jörg Doll<sup>1</sup>, Dennis Meyer<sup>1</sup>, Gabriele Kaiser<sup>1</sup>, Johannes König<sup>2</sup>**

<sup>1</sup>Universität Hamburg, <sup>2</sup>Universität Köln

Berichtet werden Ergebnisse des Hamburger Projekts „ProfaLe“ und des Kölner Projekts ZuS, die im Rahmen der „Qualitätsoffensive Lehrerbildung“ kooperieren. Die Bachelor- und Masterpraktika an den Universitäten von Hamburg und Köln werden als komplexe Lerngelegenheiten im Lehramtsstudium (Hascher, 2012; König & Klemenz, 2015) vergleichend analysiert. Als theoretischer Rahmen dient die Unterscheidung von intendierten und implementierten Curricula (König et al., 2017; McDonnel, 1995). Die intendierten Curricula weisen zwar an beiden Standorten große Ähnlichkeiten auf, jedoch finden sich bedeutende Unterschiede in der zeitlichen Positionierung des Bachelorpraktikums im Studienverlauf und der Verweildauer an den Praktikumschulen im Masterpraktikum. Die implementierten Curricula werden gemäß der Angebots-Nutzungs-Konzeption (Kunina-Habenicht et al., 2013) auf der Angebotsseite durch die Merkmale der Kohärenz von hochschulbezogener und schulpraktischer Ausbildung und der Mentorenunterstützung und auf der Nutzungsseite durch das Modell lernprozessbezogener Tätigkeitsdimensionen von König, Tachtsoglou, Darge und Lünemann (2014) beschrieben.

Aus dem Unterschied in der zeitlichen Positionierung des Bachelorpraktikums in Köln im 2. Semester und in Hamburg im 5. Semester wird die Hypothese abgeleitet, dass die Studierenden an der Universität Hamburg in diesem Praktikum mehr Aktivitäten in der Dimension „Theorien auf Situationen beziehen“ ausüben sollten als die Studierenden an der Universität zu Köln. Im Masterpraktikum sollte der Unterschied in dieser Aktivitätsdimension zwischen beiden Hochschulen geringer ausfallen. Außerdem wird an beiden Hochschulen auf Grund der curricularen Vorgaben eine Zunahme in der Ausübung von lernprozessbezogenen Tätigkeiten in der Dimension „Theorien auf Situationen beziehen“ vom Bachelor- zum Masterpraktikum erwartet. Insgesamt 1036 Lehramtsstudierende der beiden Hochschulen schätzten in online-Umfragen die Angebotsmerkmale der schulpraktischen Lerngelegenheiten ein und gaben in der Tätigkeitsdimension „Theorien auf Situationen beziehen“ an, welche der 11 Tätigkeiten, die zur Operationalisierung dieser Dimension formuliert worden waren, sie ausgeübt hatten. In allen Merkmalen fanden sich große interindividuelle Unterschiede, die auf die Komplexität der schulpraktischen Lerngelegenheiten und die Bedeutung individueller Lernorientierungen bei der Nutzung dieser Lerngelegenheiten hinweisen (van Oosterheert, Vermunt & Denessen, 2002). Hypothesenkonform wurden sowohl in Hamburg häufiger „Theorien auf Situationen beziehen“ als in Köln als auch in den Masterpraktika beider Hochschulen häufiger als in den Bachelorpraktika. Die signifikante Interaktion zwischen dem Studienort und der Studienphase belegte die Verringerung des Unterschieds zwischen den Hochschulen im Masterpraktikum. Die Angebotsvariablen der Mentorenunterstützung und der Kohärenz erwiesen sich als signifikante Mediatoren, die aber nur einen Teil der Effekte der Studienphasen und Standorte aufklären konnten, die nach Ausparialisierung des Einflusses der Mediatoren weiterhin signifikant blieben.

## F01

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus S103

### **Leistungsbezogene und soziale Kompositionseffekte im Mathematikunterricht der Grundschule – Ein HLM-Ansatz zu Effekten der Schulzusammensetzung in TIMSS 2015**

**Mario Vennemann, Birgit Eickelmann**

Universität Paderborn, Deutschland; [mario.vennemann@upb.de](mailto:mario.vennemann@upb.de)

In der Diskussion um bildungsbezogene Benachteiligungen werden von der empirischen Bildungsforschung – neben Bildungsdisparitäten, die auf individuelle und familienbezogene Merkmale zurückzuführen sind (z.B. soziale Lage; vgl. Stubbe et al., 2016; OECD, 2016) – seit einigen Jahren zusätzlich klassen- und schulbezogene Kompositionseffekte fokussiert, die dann anzutreffen sind, wenn das Lernen in einer Klasse oder einer Schule durch die leistungsbezogene oder soziale Schülerzusammensetzung beeinflusst wird (Maaz et al., 2009), was zusätzliche Anforderungen an das professionelle Handeln der Lehrkräfte stellt. Während Kompositionseffekte für den Primarbereich in Deutschland bisher vielfach lediglich auf der Grundlage der Angaben von Schulleitungen (z.B. zum Einzugsgebiet der Schule) und mithilfe deskriptiver Statistiken dargestellt wurden, werden in diesem Beitrag konkrete Indikatoren der leistungsbezogenen und sozialen Komposition von Grundschulen (z.B. Anteil der Kinder mit Migrationshintergrund) unter Nutzung multivariater Analysemethoden zugrunde gelegt, um Effekte der Schülerzusammensetzung auch unter Kontrolle individueller Merkmale der Grundschülerinnen und Grundschüler zu untersuchen, was beispielsweise für Studien aus Europäischen Nachbarländern (vgl. Boonen et al., 2014; Opdenakker et al., 2002) oder aus dem US-amerikanischen Raum (Southworth, 2010) bereits seit einigen Jahren zum State-of-the-Art gehört.

Für Deutschland ist für die mathematischen Kompetenzen im Primarbereich in diesem Zusammenhang unbekannt, ob sich schulbezogene Kompositionseffekte vor dem Hintergrund dieser Operationalisierungen und insbesondere vor dem Hintergrund individueller Merkmale zeigen. Dieses Forschungsdesiderat wird mit dem vorliegenden Beitrag auf Grundlage von Sekundäranalysen des aktuellen Studienzyklus' der Trends in International Mathematics and Science Study (TIMSS 2015; Wendt et al., 2016) bearbeitet. Im Einzelnen werden die folgenden Forschungsfragestellungen fokussiert:

1. Inwieweit lässt sich die Varianz in den mathematischen Kompetenzen der Schülerinnen und Schüler in TIMSS 2015 durch individuelle und schulische Merkmale erklären?
2. Sind in Mathematik schulische Kompositionseffekte zu beobachten und wie sind diese im Vergleich zu den Effekten auf der Individualebene einzuordnen?

Mittels hierarchisch-linearer Modelle (HLM; Raudenbush & Bryk, 2002) wird auf der Grundlage der repräsentativen Schülerstichprobe zunächst untersucht, inwieweit die mathematischen Kompetenzen der Viertklässlerinnen und Viertklässler (N= 3948) durch Merkmale auf der Individual- und Schulebene zu erklären sind (One-Way ANOVA with Random Effects). Anschließend werden sowohl individuelle Merkmale der Schülerinnen und Schüler (Geschlecht, kulturelles Kapital, Migrationshintergrund und sozioökonomischer Status [SES]) einbezogen als auch die Kompositionsmerkmale der Schulen durch Aggregation der Daten der Individualebene modelliert. Dazu gehören einerseits die leistungsbezogenen Indikatoren der Schülerzusammensetzung (Leistungsniveau und -heterogenität der Schule in Mathematik) als auch die Indikatoren der sozialen Schülerzusammensetzung (Anteile der Jungen, der Schülerinnen und Schüler mit hohem kulturellem Kapital und SES und mit Migrationshintergrund). Die für den vorliegenden Beitrag genutzten hierarchisch-linearen Modelle werden mit Mplus7 (Muthén & Muthén, 2012) unter Berücksichtigung der komplexen TIMSS-2015-Stichprobe modelliert.

Die Analysen zeigen, dass 17.5 Prozent der Variation in Mathematik durch Unterschiede zwischen den Schulen erklärt werden kann ( $ICC = .175$ ) und Hinweise darauf zu finden sind, dass schulische Merkmale geeignet sind, die Unterschiede in den mathematischen Kompetenzen von Grundschulkindern in Deutschland zu erklären (Forschungsfrage 1). Weiterhin wird deutlich, dass unter Kontrolle individueller Merkmale sowohl ein signifikanter Zusammenhang mit dem allgemeinen mathematischen Leistungsniveau als auch mit der Leistungsheterogenität besteht (Forschungsfrage 2). Weiterhin sind signifikante Zusammenhänge mit dem Anteil der Schülerinnen und Schüler, die über mehr als 100 Bücher verfügen (kulturelles Kapital), mit dem Anteil der Schülerinnen und Schülern mit Migrationshintergrund (mindestens ein Elternteil im Ausland geboren) sowie mit dem Anteil an Schülerinnen und Schülern mit hohem SES zu beobachten. Im Vergleich der beiden Indikatorgruppen ist den individuellen Merkmalen jedoch größere Relevanz beizumessen.

Die Ergebnisse zeigen damit, dass für Grundschülerinnen und Grundschüler in Deutschland nicht nur leistungsbezogene, sondern auch sozialbezogene Kompositionseffekte bedeutsam sind. Für das professionelle Handeln ergeben sich somit Entwicklungsbedarfe speziell für Lehrpersonen, aber auch Herausforderungen für die Bildungsforschung ganz allgemein, die abschließend vor dem Hintergrund der individuellen und aggregierten Disparitäten diskutiert werden.

### **Bedingungen systematisch verzerrter Lehrpersonenerwartungen an Leistungen, Anstrengungsbereitschaft und Verhaltensauffälligkeiten von Schülerinnen und Schülern**

**Tamara Stotz, Edith Niederbacher, Markus P. Neuenschwander**

Pädagogische Hochschule FHNW; [markus.neuenschwander@fhnw.ch](mailto:markus.neuenschwander@fhnw.ch)

Zahlreiche internationale Studien bestätigten, dass Schülerinnen und Schüler mit Migrationshintergrund geringere schulische Leistungen erbringen als solche ohne Migrationshintergrund (Bonsen, Kummer & Bos, 2008; Konsortium PISA.ch, 2010; Schwippert, Hornberg, Freiberg & Stubbe, 2007). Es gibt Hinweise darauf, dass Lehrpersonen mit ihren Erwartungen die Leistungsentwicklung von Schülerinnen und Schülern beeinflussen (self-fulfilling prophecy, Jussim, Robustelli & Cain, 2009). In ihrem theoretischen Modell zeigten Jussim, Eccles und Madon (1996), dass Lehrpersonen Merkmale der Schülerinnen und Schüler (z. B. frühere schulische Leistungen, kognitive Fähigkeiten, familiäre Merkmale) wahrnehmen, entsprechende Erwartungen und Überzeugungen entwickeln und damit die Schulleistungen der Kinder beeinflussen. Diese Erwartungen sind allerdings je nach sozialer Herkunft, ethnischer Herkunft sowie aufgrund der Geschlechtszugehörigkeit der Schülerinnen und Schüler verzerrt (Lorenz, Gentrup, Kristen, Stanat & Kogan, 2016; Hannover, 2004). Für den Schweizer Kontext liegen diesbezüglich jedoch kaum empirische Befunde vor. Zudem gibt es kaum Studien, die untersuchten, ob Lehrpersonenerwartungen neben den Leistungen auch hinsichtlich der Anstrengungsbereitschaft und Verhaltensauffälligkeiten von Schülerinnen und Schülern systematisch verzerrt sind.

Im vorliegenden Beitrag wird daher der Frage nachgegangen, wie sehr Erwartungen von Lehrpersonen in der Schweiz bezüglich der Leistungen in Deutsch und Mathematik, der Anstrengungsbereitschaft in Deutsch und Mathematik sowie der



Verhaltensauffälligkeiten von Schülerinnen und Schülern hinsichtlich der sozialen Herkunft (sozioökonomischer Status), der kulturellen Herkunft (Nationalität) sowie der Geschlechtszugehörigkeit der Schülerinnen und Schüler systematisch verzerrt sind.

Diese Frage wurde anhand einer Stichprobe von 75 Primarlehrpersonen der 4.-6. Klasse aus 6 Deutschschweizer Kantonen und ihren Schülerinnen und Schülern (N=1122) untersucht. Die Schülerinnen und Schüler bearbeiteten im Herbstsemester 2016 einen Fragebogen und standardisierte Leistungstests in den Fächern Deutsch und Mathematik.

Die Ergebnisse von stufenweisen Mehrebenenregressionsanalysen wiesen insgesamt auf bedeutsame Verzerrungen in den Einschätzungen der Lehrpersonen aufgrund der sozialen und kulturellen Herkunft und des Geschlechts der Schülerinnen und Schüler hin. Hinsichtlich der Leistungserwartungen der Lehrpersonen im Fach Deutsch zeigten sich – auch nach Kontrolle der Deutschleistungen – negativ verzerrte Erwartungen (d.h. tiefere Erwartungen) gegenüber Kindern mit Migrationshintergrund, positiv verzerrte Erwartungen (d.h. höhere Erwartungen) gegenüber Kindern aus Familien mit höherem sozioökonomischen Status sowie positiv verzerrte Erwartungen zugunsten von Mädchen. Im Fach Mathematik fanden sich dieselben Effekte, mit Ausnahme von der Geschlechtszugehörigkeit der Schülerinnen und Schüler: Die Leistungserwartungen in Mathematik waren an Jungen höher als an Mädchen. Dieser Effekt war nach Kontrolle der Leistungen in Mathematik nicht mehr signifikant. Die höheren Leistungserwartungen an Knaben im Fach Mathematik sind demnach offenbar zutreffend.

Die Lehrpersonenerwartungen an die Anstrengungsbereitschaft der Schülerinnen und Schüler in den Fächern Deutsch und Mathematik waren ebenfalls systematisch verzerrt, auch nach Kontrolle der selbstberichteten fachspezifischen Anstrengungsbereitschaft der Kinder. Es fanden sich negativ verzerrte Lehrpersonenerwartungen an die Anstrengungsbereitschaft der Kinder in beiden Fächern gegenüber Kindern mit Migrationshintergrund, positiv verzerrte Erwartungen gegenüber Kindern mit höherem sozioökonomischen Status sowie positiv verzerrte Erwartungen gegenüber Mädchen.

Hinsichtlich der Lehrpersonenerwartungen an die Verhaltensauffälligkeiten der Schülerinnen und Schüler fanden sich ebenfalls systematische Verzerrungen. Auch nach Kontrolle der selbstberichteten Verhaltensauffälligkeit der Kinder erwarteten Lehrpersonen mehr Verhaltensauffälligkeiten bei Kindern mit Migrationshintergrund und bei Knaben. Keinen Effekt gab es hinsichtlich des sozioökonomischen Status der Kinder.

Die Ergebnisse zeigen, dass die schülerbezogenen Erwartungen von Primarschullehrpersonen in der Schweiz aufgrund von Herkunftsmerkmalen sowie der Geschlechtszugehörigkeit der Schülerinnen und Schülern systematisch verzerrt sind. Dies gilt nicht nur bezüglich der Leistungen, sondern auch bezüglich der Anstrengungsbereitschaft und der Verhaltensauffälligkeiten der Kinder. Die Bedeutung dieser Ergebnisse für die Forschung zu Lehrpersonenerwartungen wird diskutiert.

## **Leistungsgerechte Benotung in der 8. Schulstufe? Eine Analyse auf Grundlage der österreichischen Bildungsstandardüberprüfung Mathematik 2012**

**Christine Schmid<sup>1</sup>, Michaela Katstaller<sup>2</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland; <sup>2</sup>School of Education, Universität Salzburg, Österreich; [christine.schmid@dipf.de](mailto:christine.schmid@dipf.de)

Die rechtlichen Bestimmungen in Österreich sehen eine kriteriale, an den Inhalten der Lehrpläne orientierte Benotungspraxis durch Lehrkräfte vor (BMUKK, 2007). Mit der Einführung von Bildungsstandards wurden die Inhalte in Form von zu erreichenden Kompetenzen formuliert. Die Bildungsstandardüberprüfung, die nach einer Baselineerhebung im Jahre 2009 für das Fach Mathematik im Jahr 2012 in der 8. Schulstufe erstmals als Vollerhebung durchgeführt wurde, bietet einen einheitlichen, d.h. Schulformen und Leistungsgruppen übergreifenden Referenzrahmen, der zur Beurteilung von Benotungspraxen herangezogen werden kann (Schreiner & Breit, 2012).

Zum Zeitpunkt der Bildungsstandardüberprüfung wurde das Fach Mathematik in den Schulen des Allgemeinen Pflichtschulzweiges (APS) nach Leistungsgruppen differenziert unterrichtet. Mit der 2012 erfolgten Umwandlung von Hauptschulen in Neue Mittelschulen wurde diese äußere Leistungsgruppen differenzierung zugunsten eines binnendifferenzierten Unterrichts im APS-Zweig aufgegeben.

Im Beitrag wird zum einen untersucht, inwieweit das Notenspektrum durch Lehrkräfte insgesamt sowie innerhalb der jeweiligen Schulzweige bzw. Leistungsgruppen ausgenutzt wird und welche Abweichungen sich im Benotungsverhalten bei gleichen Testleistungen über die Schulzweige bzw. Leistungsgruppen hinweg ergeben. Letztgenannte Analyse soll die Frage beantworten, ob es in einer bestimmten Leistungsgruppe „einfacher“ ist, eine gute Note zu erzielen (Eder & Dämon, 2010).

Zum zweiten werden systematische Verzerrungen der Notengebung in Abhängigkeit von Geschlecht, sozialer Schicht und dem Migrationshintergrund der Schüler/innen untersucht. Verschiedene Hypothesen hierzu lassen sich aus der Forschungsliteratur ableiten (z.B. Eder, 2003; Maaz, Baeriswyl & Trautwein, 2013) und entsprechende Ergebnisse für die österreichische Primarstufe sind bereits belegt (Schmid, Paasch & Katstaller, 2016).

Die Analysen wurden überwiegend in Form von logistischen Regressionen, zum Teil mehrebenenanalytisch, mit Mplus und unter Berücksichtigung der geschachtelten Datenstruktur durchgeführt. Zugrunde gelegt wurden jeweils 10 imputierte Datensätze (je N = 79 558). Einbezogen wurden das Testergebnis der Bildungsstandardüberprüfung Mathematik 8. Schulstufe sowie die von den Schüler/innen selbstberichteten Noten des Halbjahreszeugnisses. Die Angaben zu den Noten sowie auch Angaben zum sozialen Hintergrund der Schüler/innen entstammen den Schüler/innenfragebögen, die begleitend zur Testung der Schüler/innen eingesetzt wurden.

Die Ergebnisse dokumentieren, dass in der Leistungsgruppe 1 des APS-Zweiges bei gleichen Testleistungen signifikant bessere Noten erzielt wurden als im AHS-Zweig. Dasselbe gilt – allerdings in geringerem Ausmaß – für die Leistungsgruppen 2 und 3 im Vergleich zur Leistungsgruppe 1. Letztgenanntes Ergebnis wird mittels einer Mehrebenenanalyse dahingehend differenziert, als innerhalb von Klassen die Schüler/innen der Leistungsgruppe 2 und 3 bei gleichen Testleistungen eher schlechtere Noten erhielten. Hier zeichnet sich insgesamt eine Tendenz zur „Überbewertung“ von Schüler/innen der Leistungsgruppe 1 des APS-Zweiges ab.

In beiden Schulzweigen ergaben sich darüber hinaus systematische Verzerrungen. Mädchen erhielten bei gleichen Testleistungen bessere Noten als Jungen, dasselbe gilt für Schüler/innen aus Elternhäusern mit einem höheren sozioökonomischen Status sowie für Schüler/innen mit einem Migrationshintergrund. Alle drei Effekte hielten zudem auch der mehrebenenanalytischen Betrachtung stand. In der mehrebenenanalytischen Betrachtung wies außerdem der Klassenmittelwert der Testleistungen einen signifikant negativen Zusammenhang mit den Noten auf.

Der Befund bezüglich des Migrationshintergrundes weicht vom entsprechenden Befund für die österreichische Primarstufe ab, wo Schüler/innen mit Migrationshintergrund im Vergleich zu ihren Testleistungen innerhalb von Klassen schlechter bewertet

wurden. Der negative Effekt des Klassenmittelwerts der Testleistungen kann im Sinne eines Bezugsgruppeneffekts bzw. der Anwendung einer sozialen Bezugsnorm bei der Benotung interpretiert werden (Rheinberg & Fries, 2010).

Insgesamt werden die Ergebnisse im Hinblick auf ihre Bedeutung für Berechtigungen und Übergänge nach der 8. Schulstufe diskutiert. Außerdem werden mögliche Mechanismen angesprochen, die den systematischen Verzerrungen zwischen Testleistungen und Noten zugrunde liegen könnten (Lintorf, 2012). Die Studie kann als Baseline dienen für Analysen auf Grundlage der Bildungsstandardüberprüfung Mathematik 8. Schulstufe 2017, die nach der Umwandlung von Hauptschulen in Neue Mittelschulen erfolgt.

## **Wirkt sich die Ethnizität von Schülerinnen auf die Einschätzung des Lernzuwachses aus? Eine experimentelle Untersuchung zur Lernverlaufdiagnostik**

**Florian Klapproth**

Medical School Berlin, Deutschland; [florian.klapproth@medicalschoo-berlin.de](mailto:florian.klapproth@medicalschoo-berlin.de)

### Theoretischer Hintergrund

Auf Grundlage von Lernverlaufdiagnostik (Deno, 1985; Klauer, 2011) können Lehrkräfte Entscheidungen für oder gegen die Durchführung von Förderprogrammen treffen (Hosp, Hosp, & Howell, 2007) oder längerfristige Lernziele festlegen (Good & Shinn, 1990). Dabei ist die Qualität von Entscheidungen im Rahmen der Lernverlaufdiagnostik häufig Gegenstand empirischer Untersuchungen gewesen (z. B. Van Norman & Christ, 2016). Lehrkräfte neigen demnach dazu, das Ausmaß der Lernzuwächse von Schülerinnen und Schülern zu unterschätzen (Sanders, 1992) und können einen Lernzuwachs in stark schwankenden im Vergleich zu gering schwankenden Verläufen schlechter erkennen (Nelson, Van Norman, & Christ, 2017). Obwohl in zahlreichen Studien festgestellt wurde, dass auch die Ethnizität von Schülerinnen und Schülern die Beurteilung ihrer Leistungen beeinflusst (z. B. Glock, Krolak-Schwerdt, Klapproth, & Böhmer, 2013), wurde hingegen kaum untersucht, ob Beurteilungen von Lernverlaufskurven abhängig sind von der Ethnizität des Schülers bzw. der Schülerin sind.

### Fragestellung

Mit der vorliegenden experimentellen Studie wurden Beurteilungen von Schülerleistungen auf Grundlage unterschiedlicher Lernkurven geprüft. Ferner wurde geprüft, ob die Ethnizität von Schülerinnen und Schülern die Beurteilung des Ausmaßes des Lernzuwachses beeinflusst.

### Methode

99 Lehramtsstudierende (48 % weiblich) mit einem Durchschnittsalter von 20.5 Jahren (SD = 3.1) nahmen an der Untersuchung teil. Jede Versuchsperson erhielt 8 Vignetten, die jeweils von einer Schülerin die Leistung in einem Lesetest über 11 Wochen in Form einer Lernkurve abbildeten. Aufgabe der Versuchspersonen war einzuschätzen, ob die abgebildeten Lernkurven einen Lernzuwachs erkennen ließen oder nicht. Variiert wurden (1) die Steigung der Lernkurve (flach vs. steil), (2) das Ausmaß zufälliger Schwankungen innerhalb der Lernkurve (niedrig vs. hoch) und (3) die Ethnizität der Schülerinnen (deutsch vs. türkisch).

### Ergebnisse und Diskussion

Die Ergebnisse wurden mit Hilfe einer Mehrebenen-Regressionsanalyse für binäre Daten ausgewertet. Ebene 1 repräsentierte die wiederholten Messungen je Versuchsperson, Ebene 2 repräsentierte die Prädiktoren „Steigung der Lernkurve“, „Ausmaß zufälliger Schwankungen innerhalb der Lernkurve“ und „Ethnizität der Schülerinnen“. Geschätzt wurden sowohl ein Haupteffekt-Modell (Modell 1) als auch ein Modell mit Interaktionstermen (Modell 2). In Modell 1 waren alle Prädiktoren signifikant. Die Wahrscheinlichkeit der Identifikation von Lernzuwachs stieg, je steiler die Lernkurve und je geringer die Schwankungen waren und wenn die Schülerinnen deutsch statt türkisch waren. In Modell 2 war darüber hinaus die Ethnizität x Steigung-Interaktion signifikant. Diese bedeutet, dass die Unterschiede in der Wahrscheinlichkeit der Identifikation von Lernzuwachs zwischen deutschen und türkischen Schülern größer waren, wenn der Lernzuwachs hoch war. Die Ergebnisse werden im Hinblick auf ethnische Stereotype in der Urteilsbildung (Lee & Fiske, 2006) und ihre praktischen Implikationen diskutiert.

## Psychoedukation auf Klassenebene – eine Interventionsstudie zur Verringerung sozialer Distanz zu Kindern mit sonderpädagogischem Förderbedarf

**Anita Gerullis<sup>1</sup>, Simone Schulze<sup>2</sup>, Christian Huber<sup>3</sup>**

<sup>1</sup>Bergische Universität Wuppertal, Deutschland; <sup>2</sup>Bergische Universität Wuppertal, Deutschland; <sup>3</sup>Bergische Universität Wuppertal, Deutschland; [gerullis@uni-wuppertal.de](mailto:gerullis@uni-wuppertal.de), [schulze@uni-wuppertal.de](mailto:schulze@uni-wuppertal.de)

Für Schüler/innen mit einem sonderpädagogischem Förderbedarf im Bereich der emotional-sozialen Entwicklung besteht ein erhöhtes soziales Ausgrenzungsrisiko (Huber, 2008; Huber & Wilbert, 2012). Es fehlt jedoch an wirksamen Interventionsansätzen zur Förderung der sozialen Integration und zur Verringerung sozialer Distanz (Gerullis & Huber, 2017). Mit theoretischer Fundierung der Kontakthypothese (Allport, 1954) werden mögliche Interventionen zur Verbesserung sozialer Integration und dem Abbau von Vorurteilen genannt – eine dieser Möglichkeiten ist die Psychoedukation. Die Forschung im Rahmen der Kontakthypothese wurde bisher jedoch kaum auf Kinder mit Behinderungen und Förderbedarfe ausgeweitet (Gerullis & Huber, 2017). Diese Forschungslücke stellt ein großes Defizit in der Orientierung für professionelles Handeln von Lehrkräften im Förderbereich emotional-soziale Entwicklung dar.

Im Rahmen dieser Interventionsstudie wurde die aus klinisch-therapeutischen Settings bekannte Methode der Psychoedukation (Behrendt & Schaub, 2005) auf ihre Wirksamkeit zur Verringerung sozialer Distanz gegenüber Mitschüler/innen mit Beeinträchtigungen der emotional-sozialen Entwicklung, speziell Kindern mit Posttraumatischer Belastungsstörung (PTBS), untersucht. Psychoedukative Elemente wurden aus dem Bereich der Erwachsenentherapie kindgerecht adaptiert und in einer 4-stündigen Unterrichtsreihe zusammengeführt. Die Schüler/innen erhielten Informationen über das Krankheitsbild PTBS, eine Schulung im Benennen und inhaltlichen Trennen verschiedener Emotionen und erlernten im Rahmen von Rollenspielen Handlungsstrategien zum Umgang mit Kindern mit PTBS. Diese Intervention fand in N=6 Grundschulklassen statt. Weitere N=6 Klassen nahmen im Rahmen einer Kontrollgruppe an einer Wissensvermittlung anhand eines zu lesenden Textes mit Inhalten zur PTBS teil. In einem klassischen Prä-Post-Design wurden die soziale Distanz zu Kindern mit PTBS (SoDIS; Gerullis & Huber, 2016), die Einstellung der Teilnehmer/innen zu Kindern mit PTBS (CATCH-Skala; Schwab, 2015) das Wissen über das Krankheitsbild und das Klassenklima (FEESS, Rauer & Schuck, 2003) mittels paper-pencil Fragebögen erhoben.

Im Prä-Post-Vergleich zeigten sich keine signifikanten Effekte in der Veränderung der sozialen Distanz, im Klassenklima und in den Einstellungen zu Kindern mit PTBS in der Experimental und der Kontrollgruppe. Beide Gruppen zeigten jedoch einen Zuwachs an Wissen ( $t(90)=4.61, p=.000$ ). Der Vortrag stellt das Studiendesign und die Ergebnisse dar und diskutiert vor allem mit Fokus auf die nicht signifikanten Ergebnisse die Psychoedukation in ihrer Wirksamkeit als Interventionsmaßnahme im Rahmen inklusiver schulischer Settings und mögliche Verbesserungsmöglichkeiten der Intervention in zukünftigen Studien. Vor allem wird mit Blick auf die Kontakthypothese (Allport, 1954; Pettigrew & Tropp, 2011) diskutiert, welche Defizite in einer Interventionsmaßnahme ohne reale Kontaktherstellung bestehen und wie diese in Zukunft in ein Untersuchungsdesign aufgenommen werden könnten.

## Soziale Referenzierung und Einstellungen von Grundschulkindern gegenüber Peers mit sonderpädagogischem Förderbedarf

**Marwin Felix Löper, Frank Hellmich**

Universität Paderborn, Deutschland; [mloeper@mail.uni-paderborn.de](mailto:mloeper@mail.uni-paderborn.de)

Als eine wichtige Bedingung für die Gestaltung inklusiven Lernens im Grundschulunterricht wird die soziale Partizipation aller Kinder erachtet. Unter sozialer Partizipation wird mit dem Blick auf den inklusiven Grundschulunterricht die erfolgreiche Teilnahme aller Kinder unabhängig von ihren individuellen Lernvoraussetzungen verstanden. Als Bedingungen einer erfolgreichen Partizipation nennen Koster, Nakken, Pijl und Houten (2009) beispielsweise Freundschaften zwischen Schülerinnen und Schülern mit und ohne sonderpädagogischem/n Förderbedarf (SPF), Kontakte von und positive Interaktionen zwischen Schülerinnen und Schülern mit und ohne SPF, eine positive (Selbst-)Wahrnehmung der Kinder mit und ohne SPF im Klassenzimmer sowie positive Einstellungen von Kindern gegenüber Peers mit SPF. Nach Eagly und Chaiken (1993) werden dabei unter Einstellungen Bereitschaften von Individuen verstanden, auf Aspekte des persönlichen Umfelds durch Zuspruch oder Ablehnung wertend zu reagieren. Aus Untersuchungen geht hervor, dass Kinder zwar Peers mit Förderbedarf im Allgemeinen gegenüber positiv eingestellt sind, allerdings haben sie weniger positive Einstellungen gegenüber Kindern mit Förderbedarf in der sozial-emotionalen Entwicklung als gegenüber solchen mit SPF im Bereich Lernen oder in der geistigen Entwicklung (Schwab, 2015). Ungeklärt ist zurzeit, wie Einstellungen von Kindern gegenüber Peers mit Förderbedarf im Bereich der sozial-emotionalen Entwicklung – im Sinne der Theorie der sozialen Referenzierung (Feinman, 1992) – durch das Vorbildverhalten ihrer Eltern oder Lehrerinnen und Lehrer beeinflusst werden (de Boer, Pijl, Post & Minnaert, 2012). Vor diesem Hintergrund untersuchen wir in der vorliegenden Studie, ob und inwiefern Einstellungen von Kindern gegenüber Peers mit einem SPF in der sozial-emotionalen Entwicklung und im Bereich Lernen durch das von ihnen wahrgenommene Verhalten ihrer Grundschullehrkräfte gegenüber Kindern mit Förderbedarfen erklärt werden können.

In unserer Untersuchung haben wir N=743 Kinder der dritten und vierten Jahrgangsstufe anhand eines Fragebogens gebeten, Stellung zu ihren Einstellungen gegenüber Peers mit Förderbedarf in der sozial-emotionalen Entwicklung und im Bereich Lernen zu nehmen. Um die Einstellungen der Kinder erfassen zu können, haben wir ihnen ein Fallbeispiel zu einem Kind mit einem Förderbedarf im Bereich Lernen sowie ein weiteres zu einem Peer mit einem Förderbedarf im Bereich der sozial-emotionalen Entwicklung präsentiert (Schwab, 2015). Zu den Fallbeispielen beantworteten die Kinder jeweils Fragebogenitems, die an der gekürzten CATCH-Skala („Chedoke-McMaster Attitudes Towards Children with Handicaps scale“) von Bossaert und Petry (2013) orientiert waren. Darüber hinaus wurden die Kinder gebeten, Auskunft zu ihrem bisherigen Kontakt zu Peers mit Förderbedarf, zu ihren Selbstwirksamkeitsüberzeugungen im Hinblick auf den Umgang mit Peers mit SPF sowie zu ihren Einschätzungen bezüglich des Verhaltens ihrer Lehrerinnen und Lehrer gegenüber Kindern mit Förderbedarf zu geben. Die Ergebnisse verdeutlichen, dass die Grundschul Kinder über positive Einstellungen gegenüber den dargestellten Peers mit Förderbedarf verfügen. Im Detail zeigt sich, dass die Kinder Peers mit Förderbedarf in der sozial-emotionalen Entwicklung weniger positiv begegnen als solchen mit einem Förderbedarf im Bereich Lernen. Im Rahmen eines in Mplus berechneten Strukturgleichungsmodell wurden die Einstellungen der Kinder gegenüber Peers mit SPF in der sozial-emotionalen Entwicklung und im Bereich Lernen als abhängige Variablen sowie die bisherigen Kontakte zu Peers mit SPF, die Selbstwirksamkeitsüberzeugungen sowie das wahrgenommene Verhalten ihrer Lehrkräfte als unabhängige Variablen betrachtet. Die Ergebnisse verdeutlichen, dass die Einstellungen der Kinder gegenüber Peers mit SPF in der sozial-

emotionalen Entwicklung sowie im Bereich Lernen jeweils durch ihre bisherigen Kontakte zu Peers mit SPF, ihre Selbstwirksamkeitsüberzeugungen sowie das wahrgenommene Verhalten ihrer Lehrkräfte erklärt werden können. Im Detail wird dabei deutlich, dass der Effekt des wahrgenommenen Lehrkraftverhaltens auf die Einstellungen der an der Studie beteiligten Kinder gegenüber Peers mit SPF im Bereich Lernen durch ihre Selbstwirksamkeitsüberzeugungen mediiert wird. Die vorliegenden Befunde geben zusammenfassend Hinweise auf die Bedeutung des Vorbildverhaltens von Lehrerinnen und Lehrern gegenüber Kindern mit SPF im Klassenzimmer für die Entwicklung von Einstellungen bei Grundschulkindern gegenüber Peers mit Förderbedarf.

### **Einflussfaktoren auf die individualisierte Förder- und Unterrichtsgestaltung seitens der Grundschullehrkräfte**

**Hannelore Knauder<sup>1</sup>, Corinna Koschmieder<sup>2</sup>**

<sup>1</sup>Kirchliche Pädagogische Hochschule Graz, Österreich; <sup>2</sup>Karl-Franzens-Universität Graz, Österreich;  
[hannelore.knauder@kphgraz.at](mailto:hannelore.knauder@kphgraz.at)

Im erziehungswissenschaftlichen Diskurs wird häufig betont, dass für die Realisierung von individuellem Lernen und Lehren im schulischen Unterricht die Einstellungen, Haltungen und Überzeugungen der Lehrkräfte entscheidend sind, und, dass „individualisierende Förderung zuallererst und zutiefst mit einer Idee, einem Menschenbild und pädagogischen Zielvorstellungen zu tun hat“ (Schäfers, 2009, S. 77). Unter dieser Prämisse kann davon ausgegangen werden, dass einerseits die Einstellungen als Grundlage des professionellen Handelns von Lehrkräften gelten, andererseits besteht weitgehende Übereinstimmung darüber, dass Wissen und Können bzw. Überzeugungen ebenso zu den zentralen Komponenten der professionellen Handlungskompetenz von Lehrkräften zählen (z.B. Baumert & Kunter, 2006). Dementsprechend stellen Einstellungen, Überzeugungen sowie die Bereitschaft zur individualisierten Förder- und Unterrichtsgestaltung auf Seiten der Lehrkräfte wesentliche Basisbedingungen für die konkrete Realisierung von individueller Förderung in der Schule dar. Inwiefern diese zusammenhängen und/oder vorhergesagt werden können, ist bislang kaum untersucht.

Die vorliegende Studie verfolgt das Ziel, die Komponenten der Theorie des geplanten Verhaltens (Theory of Planned Behavior; Ajzen 1985) hinsichtlich ihrer Vorhersagekraft in Bezug auf die Bereitschaft und tatsächliche Durchführung individueller Fördermaßnahmen im Unterricht, zu überprüfen. Hierzu wird gemäß der Theorie des geplanten Verhaltens angenommen, dass einerseits die Förderintention und Gestaltung von individueller Förderung im Unterricht einer Lehrperson positiv vorhergesagt werden kann, und dass andererseits ein Zusammenhang zwischen Förderintention und/oder wahrgenommener Verhaltenskontrolle sowie tatsächlichem Verhalten besteht. Außerdem wird angenommen, dass die konkrete Umsetzung der individuellen Förderung im Unterricht zusätzlich von den motivationalen Orientierungen, die nach Baumert und Kunter (2006) zu den Kernelementen der Lehrerinnen- und Lehrerkompetenz gehören, abhängig ist. Diverse Untersuchungen geben Hinweise darauf, dass die intrinsische Motivation seitens der (zukünftigen) Lehrkräfte eine wichtige Einflussgröße auf die Einstellung und Intention zur Realisierung von individueller Förderung im Unterricht sein kann (Knauder & Koschmieder, 2017; 2018; Knoster, Villa & Thousand, 2000; Hellmich & Görel, 2014). Demnach wurde gemäß der Selbstbestimmungstheorie von Ryan und Deci (2000) zwischen intrinsisch und extrinsisch motivierten Personen zur Umsetzung individueller Förderung im Unterricht unterschieden. Erstere sind dann motiviert ein Ziel zu erreichen, wenn sie selbstbestimmt und kontrolliert handeln können, während Zweitere erst nach Aufforderung tätig werden.

Zur Prüfung der Zusammenhänge wurden Einstellungen, subjektive Norm, Einfluss der Schule und Selbstwirksamkeitsüberzeugung von 488 Lehrkräften der Grundschule bezüglich der individuellen Förderung im Unterricht erfasst und hinsichtlich ihrer Vorhersagekraft für die Bereitschaft und tatsächliche Durchführung untersucht. Ferner wurde neben der Erfassung der Komponenten der Einstellungen, vor allem auf den Zusammenhang der Intention und das konkrete Verhalten der Lehrkräfte in Bezug auf individuelle Förderung im Unterricht, das das Kernstück der Theorie des geplanten Verhaltens ausmacht (Ajzen, 1985, 1991; Ajzen & Fishbein, 2000), eingegangen. Die Daten hierfür wurden per Onlinefragebogen an Grundschullehrkräfte im Mai 2017 erhoben. Die Ergebnisse der Strukturgleichungsmodelle verdeutlichen für diese Stichprobe, dass die Einstellungen als signifikanter Prädiktor der intrinsischen Förderintention gelten, für die extrinsische Förderintention ist es der Einfluss der Schule. Die konkrete Umsetzung individualisierter Fördermaßnahmen kann direkt durch die wahrgenommene Verhaltenskontrolle und die intrinsische Förderintention als auch indirekt durch die Einstellung vorhergesagt werden. Somit konnte der angenommene Zusammenhang von Intention und Verhalten in der vorliegenden Studie nur teilweise bestätigt werden. Zudem zeigten die Ergebnisse des  $\chi^2$ -Differenzentests, dass die konkrete Umsetzung der individuellen Förderung nicht von den berufsspezifischen Zusatzqualifikationen der Lehrkraft beeinflusst wird. In Bezug auf das Dienstalter zeigen sich auf sehr schwachem Niveau signifikante Zusammenhänge mit der konkreten Durchführung von individuellen Fördermaßnahmen im Unterricht der Grundschule, der subjektiven Norm, dem Einfluss der Schule und den Selbstwirksamkeitsüberzeugungen.

Insgesamt verdeutlichen die Ergebnisse der vorliegenden Studie, dass Unterschiede in der Anwendung konkreter Maßnahmen zur individuellen Förderung im Unterricht der Grundschullehrkräfte, sowie deren diesbezügliche Intentionen mittels der Theorie des geplanten Verhaltens erklärt bzw. vorhergesagt werden können.

### **How to Design Educational Material for Inclusive Classes. First Results of a Mixed Method Study**

**Anna Noll<sup>1</sup>, Jürgen Roth<sup>1</sup>, Markus Scholz<sup>2</sup>**

<sup>1</sup>Universität Koblenz-Landau, Deutschland; <sup>2</sup>Pädagogische Hochschule Ludwigsburg; [noll@uni-landau.de](mailto:noll@uni-landau.de)

Inclusion is based on international legitimate developments like the Salamanca statement, which emphasizes the right of education for all (UNESCO, 1994). But to include all students regardless of their physical, intellectual, social or other abilities the educational framework conditions need to be adopted at first. One important step into this direction consists in providing all pupils access to the assignments by enhancing their readability.

#### **Theoretical Background**

Readability can be enhanced by linguistic simplifications like the application of easy-to-read language (Netzwerk Leichte Sprache, 2006). Easy-to-read language has been established to facilitate understanding in everyday life for people with disabilities. So far, these guidelines are not verified scientifically, but show several similarities to empirically based linguistic simplification rules, like the Hamburger Modell (Langer, Schulz von Thun, & Tausch, 2011). The authors state, among others, that only one statement should be made per sentence.

Another possibility to facilitate the comprehension of assignments is the use of symbols. A symbol can be defined as a graphical image conveying a single idea or concept (Detheridge & Detheridge, 2002). Little empirical data about the use of symbols to foster the readability of texts exists (Jones, Long & Finley, 2007; Poncelas & Murphy, 2007). Nevertheless, a positive influence can be assumed, e.g. because of the multimedia principle (Mayer, 2009). It indicates that people learn better from words and pictures than from words alone.

#### Methodology and First Results

How do students with and without SEN perceive and utilize symbols in learning material? This research question was focused on by a qualitative study, conducted in 2016. The sample consisted of students with learning difficulties and students without SEN. The students worked on the tasks individually and participated in a subsequent interview. The interviews were evaluated according to the coding manual by Saldana (2009) with a two rater evaluative coding approach (cf. *ibid.*) and a Cohens kappa of .75 was reached. Data was also collected by use of eye tracking. One result of the study is that symbols are used without a prior explanation by the students. This result can be exemplified by the students' reaction to one interview question: "Did the little pictures help or did they disturb you? Why?". Most students' answers were coded as "helped". The following excerpt of the interview is an example of this code:

Student 1: They helped me, they did not disturb me at all.

Interviewer: And why did they help you?

Student 1: You don't have to understand everything by reading. You also see the things you have to do.

The corresponding eye-tracking data and further transcripts of the interview will be presented at the conference. The tasks of the qualitative as well as of the quantitative study deal with introducing fractions. These are taught in activity-oriented manner with hands-on material.

For the quantitative study, which will be conducted in November and December 2017, a posttest-only design, which includes two experimental as well as one control group, is planned. While experimental group 1 receives a linguistically and pictorially simplified version of the tasks, experimental group 2 works with a variation which is linguistically simplified only. The control group receives a not simplified version. After working with the exercises, the students' knowledge about fractions will be measured. The participants' reading ability and their IQ will be elevated beforehand. These control variables shall help to build comparable groups. The following research question will be answered in this study: Does the use of easy-to-read language and/or enriching text with symbols facilitate students' performance in mathematical tasks? First results of the quantitative study will be presented at the GEBF 2018.

## F03

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus S105

### Ready to read in two languages? Deutsch- und partnersprachiges Leseverständnis von dual-immersiv unterrichteten Grundschulkindern

**Sandra Preusler<sup>1</sup>, Isabell Paulick<sup>1</sup>, Johanna Fleckenstein<sup>2</sup>, Jürgen Baumert<sup>3</sup>, Jens Möller<sup>1</sup>**

<sup>1</sup>CAU Kiel, Deutschland; <sup>2</sup>Fachhochschule Nordwestschweiz, Schweiz; <sup>3</sup>Max-Planck-Institut für Bildungsforschung, Deutschland; [spreusler@ipl.uni-kiel.de](mailto:spreusler@ipl.uni-kiel.de)

In der heutigen globalisierten Welt wird es zunehmend wichtig, mehr als eine Sprache zu sprechen. Dies gilt sowohl für Schülerinnen und Schüler (SuS), die die Verkehrssprache des jeweiligen Landes als Erstsprache (L1) gelernt haben als auch für SuS, die mit einer anderen als der Verkehrssprache als L1 aufgewachsen sind. An der Staatlichen Europa-Schule Berlin (SESB) werden Kinder mit Deutsch als L1 gemeinsam mit Kindern mit einer anderen L1 unterrichtet, wobei beide Sprachen gleichwertig eingesetzt werden. Die SESB folgt dem Konzept der dualen Immersion (Baker, 2011).

Einen besonders interessanten Aspekt des bilingualen Unterrichts stellt der Zusammenhang zwischen den unterschiedlichen Kenntnissen in der L1 und der L2 dar. Cummins (1979; 1984) entwickelte dazu die Interdepenzhypothese. Nach dieser gibt es einen positiven Zusammenhang in der Entwicklung der L1 und der L2, da ihnen eine gemeinsame Sprachfähigkeit – die common underlying proficiency (Cummins, 1981) – zugrunde liegt.

Während aus Deutschland bislang kaum Erkenntnisse vorliegen, zeigten sich international positive Effekte dualer Immersion auf schulische Leistungen und auf die Entwicklung der sprachlichen Fähigkeiten sowohl in der L1 als auch in der L2 (s. im Überblick Kim, Hutchison, & Winsler, 2015; Krashen, 2005). Dabei fehlen weltweit Studien, die es erlauben, die Kompetenzen in beiden Sprachen vergleichend zu betrachten und zu überprüfen, inwiefern das Ziel des balancierten Bilingualismus von SuS mit unterschiedlichem sprachlichen Hintergrund (ausschließlich mit Deutsch vs. ausschließlich mit der Partnersprache vs. bilingual aufgewachsen) erreicht wird. Unsere Studie versucht diese Lücke zu schließen.

#### Fragestellungen

1. Wie schneiden die einzelnen Sprachgruppen (=monolingual deutschsprachig/bilingual/nichtdeutschsprachig aufgewachsen) im Leseverständnis in den beiden Testsprachen ab? In welcher Sprache werden höhere Testleistungen erzielt?
2. Wie hängen das deutsche und das partnersprachige Leseverständnis miteinander und mit den Hintergrundvariablen zusammen?

#### Methode

Grundlage für diese Arbeit stellen die Erhebungen der EUROPA-Studie am Ende der vierten Jahrgangsstufe dar. Es handelte sich um eine Vollerhebung aller SESB-Klassen. Insgesamt umfasst die Stichprobe 939 SuS (54.0 % weiblich; Alter: M = 10.10, SD = 0.41 Jahre) von 17 Standorten der SESB. 21.4 % der Kinder sind ausschließlich mit Deutsch als Erstsprache aufgewachsen, 41.8 % sind bilingual aufgewachsen, haben also Deutsch und die jeweilige Partnersprache als L1. 36.9 % der Kinder haben ausschließlich die nichtdeutsche Partnersprache als L1 gelernt.

Die Tests, mit denen die Leseleistungen in deutscher Sprache erfasst wurden, stammten aus der PIRL-Studie von 2001 (Mullis, Martin, Gonzalez, & Kennedy, 2003). Die Tests in der Partnersprache stammen aus der PIRL-Studie aus den Jahren 2011 und 2001 (Mullis et al., 2003; Mullis, Martin, Foy, & Drucker, 2012). Als Schätzer für die Personenfähigkeit wurden für die Analysen Weighted Likelihood Estimates (WLEs) nach Warm (1989) genutzt.

Hintergrundvariablen stellen die Erstsprache des Kindes, der sozioökonomische Status der Eltern (erfasst durch den HISEI; vgl. Ganzeboom & Treiman, 1996) und die kognitive Leistungsfähigkeit des Kindes (erfasst durch den KFT; Heller & Perleth, 2000) dar.

#### Ergebnisse

Die Ergebnisse der Pfadanalysen zeigen, dass dual-immersiv unterrichtete Kinder – unabhängig von ihrer Erstsprache – signifikant besser im deutschen als im partnersprachigen Leseverständnistest abschneiden. Dabei ist der Unterschied zwischen den beiden Testsprachen am größten bei den monolingual deutschsprachig Aufgewachsenen und am geringsten bei den monolingual nichtdeutschsprachig Aufgewachsenen. Die monolingual nichtdeutschsprachig Aufgewachsenen erreichen also am ehesten eine balancierte Bilingualität, wobei ihr Leseverständnis im Deutschen auch deutlich unter dem der monolingual nichtdeutschsprachig und bilingual Aufgewachsenen liegt.

Regressionsanalysen zeigen einen signifikant positiven Einfluss des partnersprachigen Leseverständnisses auf das Leseverständnis im Deutschen, der auch dann nachzuweisen ist, wenn für den Sprachhintergrund, den sozioökonomischen und kognitiven Hintergrund kontrolliert wird.

Zusammenfassend lässt sich feststellen, dass SESB-Schulkinder trotz zweisprachigen Unterrichts am Ende der vierten Klassenstufe noch keine balancierten Leseverständnisfähigkeiten in zwei Sprachen aufweisen. Zudem zeigt sich ein positiver Zusammenhang zwischen den Leseleistungen in den beiden Testsprachen.

### Implizite Fähigkeitstheorien und Lesemotivation von Kindern im Grundschulalter im Zusammenhang mit dem Feedback in Elternhaus und Schule

**Frank Hellmich, Fabian Hoya**

Universität Paderborn, Deutschland; [frank.hellmich@uni-paderborn.de](mailto:frank.hellmich@uni-paderborn.de)

Das Unterstützungsverhalten von Eltern sowie Lehrerinnen und Lehrern gilt als eine wichtige Voraussetzung für das Lernen von Kindern im Grundschulalter. Feedback zu Lernergebnissen und Lernprozessen kommt dabei eine besondere Bedeutung zu. Unter Feedback werden Informationen verstanden, die Eltern oder Lehrkräfte Kindern in der Folge von Lernleistungen bereitstellen, um Diskrepanzen zwischen aktuell erbrachten Leistungsergebnissen und den zu erreichenden Lernzielen aufzuzeigen (Hattie & Timperley, 2007). Die Bedeutung von Feedback wichtiger Bezugspersonen wie Eltern oder Lehrkräfte für implizite Fähigkeitstheorien von Kindern, ihre fähigkeitsbezogenen Selbstkonzepte oder ihre intrinsische und extrinsische Motivation konnte in den vergangenen Jahren verdeutlicht werden (Burnett, 2003; Cimpian, Arce, Markman & Dweck, 2007; Mueller & Dweck, 1998). Die Rolle spezifischer Feedbackformen wie beispielsweise anstrengungsbezogene Rückmeldungen – verstanden als Lob oder Kritik in Bezug auf das kindliche Anstrengungsverhalten (Burnett & Mandel, 2010; Kamins & Dweck, 1999) – für die Entwicklung impliziter Fähigkeitstheorien, domänenspezifischer Selbstkonzepte sowie die Lern- und

Leistungsmotivation von Kindern im Grundschulalter ist mit wenigen Ausnahmen (Gunderson, Gripshover, Romero, Dweck, Goldin-Meadow & Levine, 2013; Mueller & Dweck, 1998) nur wenig untersucht. So zeigen beispielsweise die Ergebnisse der Untersuchung von Gunderson et al. (2013), dass die impliziten Fähigkeitstheorien von Kindern signifikant über das anstrengungsbezogene Feedback ihrer Eltern erklärt werden.

In Hinblick auf die Genese selbstbezogener Kognitionen und Motivationen im Kindesalter ist gegenwärtig ebenfalls weitgehend unbeantwortet, ob dem anstrengungsbezogenen Feedback der Eltern oder demjenigen ihrer Lehrkräfte eine größere Bedeutung für die Entwicklung impliziter Fähigkeitstheorien, dem akademischen Selbstkonzept sowie der intrinsischen und extrinsischen Motivation von Lernenden zukommt. In unserer Studie wird vor diesem Hintergrund am Beispiel des Leseunterrichts in der Grundschule untersucht, ob und inwiefern Unterschiede in der intrinsischen und extrinsischen Lesemotivation von Kindern im Grundschulalter durch ihre Wahrnehmungen anstrengungsbezogenen Feedbacks in Elternhaus und Schule erklärt werden können. Dabei nehmen wir in Anlehnung an Eccles (2005) an, dass Effekte des wahrgenommenen Feedbacks auf die Lesemotivation von Schülerinnen und Schülern aus dem dritten und vierten Grundschuljahr jeweils durch ihre impliziten Fähigkeitstheorien sowie ihre lesebezogenen Selbstkonzepte mediiert werden.

Zur Überprüfung dieser Hypothesen wurden in unserer Studie N=685 Schülerinnen und Schüler der dritten und vierten Klassenstufe (350 Mädchen, 335 Jungen) anhand eines Fragebogens zu ihren impliziten Fähigkeitstheorien, ihrer intrinsischen und extrinsischen Lesemotivation, ihren lesebezogenen Selbstkonzepten sowie ihrem wahrgenommenen anstrengungsbezogenen Feedback ihrer Eltern und Lehrkräfte befragt. Die Ergebnisse unserer Untersuchung verdeutlichen, dass die an der Studie beteiligten Kinder häufiger anstrengungsbezogenes Feedback im Elternhaus wahrnehmen als von ihren Lehrkräften im Leseunterricht. Darüber hinaus belegen die Ergebnisse aus einem in Mplus berechneten Strukturgleichungsmodell, dass die lesebezogenen Selbstkonzepte der Kinder signifikant über das wahrgenommene anstrengungsbezogene Feedback ihrer Eltern sowie auch über dasjenige ihrer Lehrkräfte erklärt werden können. Dagegen können Unterschiede in den impliziten Fähigkeitstheorien der Schülerinnen und Schüler lediglich über das von ihnen perzipierte anstrengungsbezogene Feedback ihrer Eltern erklärt werden, nicht aber über das wahrgenommene Feedback ihrer Lehrkräfte. Die intrinsische Motivation der an der Studie beteiligten Kinder wird in dem Modell signifikant über das von ihnen wahrgenommene anstrengungsbezogene Feedback ihrer Eltern sowie über ihre impliziten Fähigkeitstheorien vorgesagt. Die extrinsische Motivation der Schülerinnen und Schüler wird signifikant über das von ihnen wahrgenommene Feedback der Eltern, ihre lesebezogenen Selbstkonzepte sowie ihre impliziten Fähigkeitstheorien erklärt. Entgegen den theoretischen Annahmen belegen die Ergebnisse, dass das durch die Kinder perzipierte anstrengungsbezogene Feedback ihrer Lehrkräfte keinen signifikanten Prädiktor für ihre impliziten Fähigkeitstheorien sowie ihre intrinsische und extrinsische Lesemotivation darstellt.

Die Befunde aus unserer Studie geben Hinweise auf die Bedeutsamkeit wahrgenommener elterlicher Rückmeldungen zu Leseprozessen im Zusammenhang mit der Lesemotivation von Kindern im Grundschulalter. Sie können bei der Weiterentwicklung von 'family literacy'-Programmen Berücksichtigung finden, bei denen Eltern an der Förderung der Lesekompetenzentwicklung ihrer Kinder beteiligt werden.

## **Sprachbezogene Disparitäten in der Notenvergabe am Ende der Grundschulzeit**

**Johannes Schult**

Landesinstitut für Schulentwicklung, Deutschland; [Johannes.Schult@ls.kv.bwl.de](mailto:Johannes.Schult@ls.kv.bwl.de)

Theoretischer Hintergrund: Die Integration von Schülerinnen und Schülern mit ausländischen Wurzeln gehört zu den Aufgaben von Bildungsinstitutionen (vgl. Blossfeld et al., 2016). Dabei zeigen sich große Unterschiede bei der Notenvergabe zugunsten einheimischer Kinder. Ähnliche Unterschiede finden sich auch bei der Kompetenzmessung in großen Bildungsstudien. Disparitäten bei der Notenvergabe können durch fachliche Kompetenzen jedoch nicht vollständig erklärt werden (Bonefeld et al., 2017; Stanat & Christensen, 2006).

Ein Teil der Ungleichheiten lässt sich auf die mangelnde Beherrschung der Unterrichtssprache zurückführen (vgl. Kempert et al., 2017). Geringe Deutschkenntnisse erschweren den Zugang zu Lerngelegenheiten und haben negative Konsequenzen für die weitere schulische Entwicklung. Das Leistungsdefizit in Bildungsstudien fällt entsprechend bis zu doppelt so groß aus, wenn Schülerinnen und Schüler mit Migrationshintergrund im Alltag nicht die Unterrichtssprache sprechen (Stanat & Christensen, 2006).

Fragestellung: Die Rolle der Alltagssprache sollte gerade in Grundschulen im Fach Deutsch stärker ausgeprägt sein als in Mathematik (vgl. Gresch, 2016). Denn während die Sprache in Wort und Schrift ein wesentlicher Bestandteil der Lehrinhalte in Deutsch ist, stehen in Mathematik numerische Inhalte im Mittelpunkt. Die Forschungshypothesen lauten deshalb: a) Der Notenunterschied zwischen Schülerinnen und Schülern mit deutscher und mit nichtdeutscher Alltagssprache ist in Deutsch größer als in Mathematik. b) Nach der Korrektur für fachspezifische Kompetenztestleistungen bleibt der Notenunterschied in Deutsch größer als in Mathematik.

Methode: Die Analysen basieren auf zwei Zufallsstichproben baden-württembergischer Grundschulen. Pro Schule nahm eine Klasse an der Testung teil. Die Deutsch-Stichprobe bestand aus 3234 Kindern aus 185 Schulen. Die Mathematik-Stichprobe bestand aus 4889 Kindern aus 275 Schulen. Die Schülerinnen und Schüler bearbeiteten am Ende des vierten Schuljahres Kompetenztests (konkret: Aufgaben des Verfahrens „Lernstand 5“; vgl. Fischer et al., im Druck; Schulz et al., 2017) in den Bereichen Lesegeschwindigkeit und Leseverständnis (Deutsch) bzw. Rechenverfahren, Operationsverständnis und Zahlverständnis (Mathematik). Die jeweiligen Lehrkräfte gaben an, welche Noten die Kinder im letzten Halbjahreszeugnis bekommen hatten und welche der Kinder eine nichtdeutsche Alltagssprache hatten (Deutsch: 12%, Mathematik: 11%).

Die jeweilige Fachnote wurde als abhängige Variable in einer Mehrebenenregression verwendet (Ebene 1: Kind, Ebene 2: Schule). Zuerst wurde nur die dichotome Variable „nichtdeutsche Alltagssprache“ als Prädiktor verwendet. Anschließend wurden die Kompetenztestleistungen als weitere Prädiktoren aufgenommen und geprüft, ob der sprachbezogene Unterschied weiter erhalten bleibt.

Ergebnisse: In Deutsch waren die Noten von Kindern mit deutscher Alltagssprache im Schnitt um 0,73 Notenpunkte besser als die Noten von Kindern mit nichtdeutscher Alltagssprache. Durch die Kontrolle der Testleistungen schrumpfte der Notenunterschied auf 0,30 Notenpunkte. In Mathematik betrug der geschätzte Gruppenunterschied 0,58 Notenpunkte. Die Kontrolle der Testleistungen reduzierte den Notenunterschied auf 0,18 Notenpunkte. Sowohl vor als auch nach der Kontrolle der Testleistungen waren die Gruppenunterschiede in Deutsch signifikant größer als in Mathematik ( $p < .01$ ).

Die Befunde bestätigen, dass die Beherrschung der Unterrichtssprache von zentraler Bedeutung für den Schulerfolg ist (vgl. Esser, 2006). Die Relevanz der Alltagssprache bezüglich der Notengebung zeigte sich besonders in Deutsch, im geringeren

Umfang aber auch in Mathematik. Die Ergebnisse werden im Kontext von sprachlichen Förderangeboten einerseits (vgl. Blossfeld et al., 2016) und möglichen Stereotyp-bedingten Urteilsfehlern bei der Notenvergabe andererseits (vgl. Lorenz, 2017) diskutiert.

## **Sprachliche Ausdrucksfähigkeit in Mathematik – Konzeption eines diagnostischen Instruments zur Messung der schriftsprachlichen Kompetenzen von Dritt- und Viertklässlern (SAMT)**

**Alexandra Merkert, Anja Wildemann, Gerlinde Lenske**

Universität Koblenz-Landau, Deutschland; [merkert@uni-landau.de](mailto:merkert@uni-landau.de)

### 1. Theoretischer Hintergrund

Aus PISA geht hervor, dass sprachliche Kompetenzen eine wichtige Voraussetzung für das Lernen und den Schulerfolg bilden (OECD, 2007). Empirische Studien belegen zudem den erheblichen Einfluss der Sprachkompetenz auf die Mathematikleistungen von Lernenden (Gürsoy, Benholz, Renk, Prediger & Büchter, 2013; Mücke, 2007). Kombiniert man diese Befunde mit der in den Bildungsstandards für die Grundschule formulierten Forderung, dass Schülerinnen und Schüler im Unterricht in angemessener Weise über mathematische Inhalte kommunizieren und argumentieren sollen (Kultusministerkonferenz der Länder, 2014) lässt sich daraus die Notwendigkeit einer gezielten Förderung und auch Diagnose der Sprachkompetenz in Mathematik ableiten. Dabei müssen im Mathematikunterricht neben dem Umgang mit unterschiedlichen symbolischen, verbalen und graphischen Darstellungsformen auch bildungs- und fachsprachliche Aspekte berücksichtigt werden (Prediger & Wessel, 2011). Der Begriff der Bildungssprache geht dabei auf Konzepte aus dem englischsprachigen Raum wie beispielsweise das „cognitive academic language proficiency“ (Cummins, 2000) zurück. Um die Wirksamkeit eingeleiteter Sprachfördermaßnahmen im Fach Mathematik evaluieren zu können, fehlt es bislang aber vor allem im Primarbereich an geeigneten diagnostischen Instrumenten. Daher wurde im Rahmen des Projekts „Eva-Prim: Evaluation im Primarbereich – Sprachförderung in alltäglichen und fachlichen Kontexten“, das innerhalb der bundesweiten „BiSS“-Initiative (Bildung durch Sprache und Schrift) durchgeführt wird, ein Verfahren zur Diagnose der mathematisch-schriftsprachlichen Kompetenzen entwickelt.

### 2. Fragestellung:

Wie kann ein sprachdiagnostisches Instrument konzipiert sein, um die fachsprachlichen Kompetenzen von Schülerinnen und Schülern der dritten und vierten Klasse unter Berücksichtigung bildungssprachlicher Kompetenzen zu messen?

- Welche Aufgaben eignen sich zur Erhebung der mathematisch-schriftsprachlichen Kompetenzen?

- Welche verschiedenen Niveaus mathematisch-schriftsprachlicher Kompetenzen lassen sich bei Dritt- und Viertklässlern identifizieren?

- Welche Zusammenhänge lassen sich zwischen den ermittelten schriftsprachlichen Kompetenzen von Schülerinnen und Schülern in Mathematik und ihren mathematischen und bildungssprachlichen Kompetenzen, ihren Lesefertigkeiten sowie ihrem sozioökonomischen Hintergrund feststellen?

### 3. Methode

Anhand festgelegter Kriterien wurden zunächst sprachensible Mathematikaufgaben aus den Bereichen Arithmetik, Geometrie und Sachrechnen konzipiert, die zur Verschriftlichungen unterschiedlicher Lösungsdarstellungen und -wege anregen sollen. Dabei wurde das Ziel der Evokation spezifischer sprachlicher Handlungen, unter Verwendung einer bestimmten Lexik, aber auch Morphosyntax, die im mathematischen Kontext funktional und differenzierend eingesetzt werden können, verfolgt. Nach ihrer dreischrittigen Pilotierung, in deren Fokus der sprachliche Anregungsgehalt, die Klarheit und Struktur sowie die Bearbeitungsdauer standen, wurden die Aufgaben zur Erhebung der mathematisch-schriftsprachlichen Kompetenzen von Dritt- und Viertklässlern eingesetzt. Die schriftsprachlichen Erzeugnisse wurden im Anschluss in Anlehnung an die strukturierende skalierende Qualitative Inhaltsanalyse nach Mayring (2015) ausgewertet, um verschiedene Niveaus der mathematisch-schriftsprachlichen Kompetenzen zu ermitteln. Im Anschluss erfolgte die Pilotierung der Skala auf der Grundlage der Daten von 50 Schülerinnen und Schülern der dritten Klassenstufe. Seit Januar 2017 wird das SAMT-Verfahren im Projekt Eva-Prim zur Messung der schriftsprachlichen Kompetenzen im Fach Mathematik von Dritt- und Viertklässlern aus Bayern, Baden-Württemberg, Nordrhein-Westfalen, Sachsen-Anhalt und Rheinland-Pfalz eingesetzt. Daneben werden die Mathematikleistungen (DEMAT), die bildungssprachlichen Kompetenzen (BiSpra), die Lesekompetenz (SLs), der Wortschatz (PPVT) und die Grundintelligenz (CFT1-R) erhoben. Der sozioökonomische Status wird auf der Basis von Elternbefragungen ebenfalls erfasst.

### 4. Ergebnisse

Durch SAMT als sprachdiagnostischem Instrument soll es zukünftig möglich sein, verschiedene Kompetenzniveaus im Bereich der Sprachkompetenz in Mathematik bei Dritt- und Viertklässlern abzubilden, um sowohl den Kompetenzstand der Lernenden ermitteln zu können sowie Entwicklungsprozesse nachzuzeichnen als auch die Wirksamkeit eingeleiteter Sprachfördermaßnahmen im Rahmen eines sprachsensiblen Mathematikunterrichts zu evaluieren. Erste Analysen mittels konfirmatorischer Faktorenanalyse bestätigen das durch die Skala abgebildete Konstrukt. Eine Prüfung der internen Konsistenz als Maß für die Reliabilität auf der Grundlage der Pilotierungsdaten (n=50) liefert zudem zufriedenstellende Cronbachs-Alpha-Werte ( $\alpha = .762$ ). Eine hohe Interraterreliabilität von  $r_s = 0.929$  deutet darüber hinaus auf eine hinreichende Objektivität des Verfahrens hin. Eine Validierung des Instruments wird anhand der im Projekt Eva-Prim gewonnenen Daten bezüglich der bildungssprachlichen Kompetenzen in der Rezeption, der basalen Lesefertigkeit, den mathematischen Kompetenzen sowie den Schulleistungen vorgenommen.



## F04

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Rosshofgasse, S02

### **Individuelle Merkmale von Lehramtsstudierenden: Eine differenzierte Überprüfung der Negativselektionshypothese**

**Julia-Carolin Brachem, Hilde Schaeper**

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Deutschland; [brachem@dzhw.eu](mailto:brachem@dzhw.eu)

Individuelle Merkmale von Lehrkräften, darunter kognitive Fähigkeiten, Persönlichkeitsmerkmale und berufliche Interessen, werden als bedeutsam für die Entwicklung professioneller Lehrerkompetenzen und infolgedessen als relevant für guten Unterricht und die Entwicklung von Schülerkompetenzen und Schülerleistungen erachtet (Klusmann, 2013; Kunter et al., 2011; Roloff Henoch et al., 2015; Rothland, 2014).

Diverse internationale Studien zeigen jedoch auf, dass der Lehrerberuf häufig von Personen mit ungünstigeren kognitiven und psychologischen Merkmalen ergriffen wird (Denzler & Wolter, 2009; Eder et al., 2015; Guarino et al., 2006; Zumwalt & Craig, 2008). Im deutschen Kontext gibt es hingegen Studien, die keinen Hinweis für eine generelle Negativselektion ins Lehramt finden, sobald Lehramtstypen und Studienfächer in die Betrachtung einbezogen werden (Klusmann et al., 2009; Roloff Henoch et al., 2015; Rothland, 2014). Eine gewisse „interne Selektion“ je nach gewähltem Lehramtstyp und Studienfach kann jedoch durchaus festgestellt werden. Unterschiedliche kognitive, Persönlichkeits- und Interessenskonstellationen der Lehramtskandidaten bringen diese zum Ausdruck (Brookhart & Freeman, 1992; Gold & Giesen, 1993; Kaub et al., 2012; König et al., 2013; Neugebauer, 2013; Retelsdorf & Möller, 2012; Watt et al., 2012).

Unter Rückgriff auf bewährte Berufswahltheorien wie das Konzept der Person-Umwelt-Passung (Holland, 1997) und die Werterwartungstheorie (Eccles 1983; Wigfield & Eccles, 2000) wird im Rahmen der vorliegenden Studie die angenommene Negativselektion ins Lehramt differenziert nach Lehramtstyp und Studienfach überprüft. Dabei wird untersucht, (1) welchen Einfluss die individuellen Merkmale von Lehramtsstudierenden (kognitive Fähigkeiten, Persönlichkeitsmerkmale, berufliche Interessen) auf die Wahl eines Studiengangs (Lehramt vs. Nicht-Lehramt), (2) auf die Wahl eines Lehramtstyps (Primarbereich vs. Sekundarbereich I vs. Sekundarbereich II) und (3) auf die Wahl eines Studienfachs (MINT vs. Nicht-MINT) haben.

Auf der Basis umfangreicher Daten des Nationalen Bildungspanels (Startkohorte 5: Studierende (doi:10.5157/NEPS:SC5:9.0.0)) ist es im Rahmen dieser Studie möglich, eine größere Stichprobe zu untersuchen, die sowohl Lehramtsstudierende (N = 5.332) als auch Studierende anderer, fachlich vergleichbarer Studiengänge (N = 3.826) umfasst. Vor diesem Hintergrund kann die Gruppe der Lehramtsstudierenden differenziert betrachtet und Lehramtstypen und Studienfächer in den Analysen gleichzeitig berücksichtigt werden. Bei den Befragungsteilnehmenden handelt es sich um Studienanfängerinnen und Studienanfänger des Wintersemesters 2010/2011 an deutschen Hochschulen (Aschinger et al., 2011; Maurice et al., 2016).

Über multinomiale logistische Regressionsmodelle wird der Einfluss individueller Merkmale auf die Wahrscheinlichkeit, in unterschiedlichen Kontexten (Studiengang, Lehramtstyp, Studienfach) zu studieren, geschätzt. Dabei werden getrennte Schätzungen für Studierende mit und ohne MINT-Fach durchgeführt, fehlende Werte imputiert und für sozio-demografische und bundeslandspezifische Studienzugangsbedingungen kontrolliert. Zur Operationalisierung der individuellen Merkmale werden die Abiturnote, Kompetenztestwerte aus den Bereichen Lesen und Mathematik sowie das Big Five- und das RIASEC-Inventar herangezogen.

Im Hinblick auf die kognitiven Fähigkeiten der Studierenden zeigt sich, dass Lehramtskandidaten des Sekundarbereichs I im Vergleich zu Studierenden anderer Studiengänge signifikant schlechtere Abiturnoten und Kompetenzwerte im Bereich Mathematik aufweisen. Ein ähnliches Bild zeigt sich im Vergleich zu Lehramtsstudierenden des Sekundarbereichs II. Persönlichkeitsmerkmale haben nach Kontrolle anderer individueller Merkmale keinen nennenswerten Einfluss auf die Wahl der betrachteten Studiengänge. Die beruflichen Interessen der Studierenden erscheinen hingegen äußerst bedeutsam. Während insbesondere soziale Interessen die Wahrscheinlichkeit, ein Lehramtsstudium zu wählen, signifikant erhöhen, wird dies umso unwahrscheinlicher, je ausgeprägter das Forschungsinteresse ist. Dieses Muster spiegelt sich am deutlichsten im Primar- und Sekundarbereich I wider. Darüber hinaus zeigen die Analysen, dass die Wahl eines MINT-Studienfachs umso wahrscheinlicher ist, je ausgeprägter die Mathematik-Kompetenz und das Forschungsinteresse einer Person sind.

Grundsätzlich zeigt die Studie, dass individuelle Merkmale mit der Wahl von Studienkontexten im Zusammenhang stehen und dass es relevant ist, die Heterogenität von (Lehramts-) Studierenden bei der Studienplanung (und Lehrerbildung) zu berücksichtigen. Die Studie gibt keinen Hinweis auf eine generelle Negativselektion ins Lehramt, wenngleich deutliche Unterschiede zwischen den verschiedenen Lehramtstypen vorhanden sind.

### **Analyse von Profilen der Studienwahlmotivation von Lehramtsstudierenden und Referendaren: Zusammenhang zu Lehrerselbstwirksamkeit, selbstreguliertem Lernen und Einstellungen zum Lehren**

**Franziska Perels<sup>1</sup>, Charlotte Dignath-van Ewijk<sup>2</sup>, Laura Dörrenbächer<sup>1</sup>**

<sup>1</sup>Universität des Saarlandes, Deutschland; <sup>2</sup>Goethe-Universität Frankfurt; [f.perels@mx.uni-saarland.de](mailto:f.perels@mx.uni-saarland.de)

Die Studienwahlmotivation angehender Lehrkräfte stellt ein wichtiges Konstrukt zur Vorhersage von Engagement im Lehrerberuf sowie der damit einhergehenden Unterrichtsqualität dar (Kunter & Pohlmann, 2009; Watt & Richardson, 2008). Im Rahmen von Erwartungs-Wert-Ansätzen lassen sich dabei verschiedene Motivationskomponenten beschreiben, die die Studienwahl beeinflussen und die sich in eher intrinsische und eher extrinsische Motivationskomplexe untergliedern lassen (Pohlmann & Möller, 2010). So geht eine intrinsische Studienwahlmotivation z. B. mit erhöhter Studienstrategienutzung, einer höheren beruflichen Zufriedenheit und höherem beruflichen Engagement einher (Künsting & Lipowsky, 2011; Lipowsky, 2003). Untersuchungen im Rahmen eines personenzentrierten Ansatzes weisen darauf hin, dass unterschiedliche Motivationsprofile bei angehenden Lehrkräften vorliegen (z. B. Chin & Young, 2007). So fanden Billich-Knapp, Künsting und Lipowsky (2012) bei Grundschullehramtsstudierenden drei Studienwahlmotivprofile, die sich in der Ausprägung intrinsischer und extrinsischer Motivationen und im Hinblick auf motivational-affektive und persönlichkeitsbezogene Merkmale unterschieden.

Die vorliegende Studie zielt darauf ab, diese drei Profile weitergehend zu untersuchen. Dazu sollen sie in einem ersten Schritt anhand einer Stichprobe von Lehramtsstudierenden verschiedener Schulformen sowie einer Referendarstichprobe repliziert werden. In einem zweiten Schritt sollen die Zusammenhänge der Profile zu weiteren, für den Lehrerberuf relevanten Konstrukten wie Lehrerselbstwirksamkeit, selbstreguliertem Lernen und Einstellungen zum Lehren sowie zu Leistungsvariablen untersucht werden.

An der Untersuchung nahmen  $n = 127$  Lehramtsstudierende verschiedener Fächer ( $M_{\text{Alter}} = 24.39$ ,  $SD = 3.84$ ; 63% weiblich) und  $n = 42$  Grundschulreferendare ( $M_{\text{Alter}} = 26.23$ ,  $SD = 4.00$ ; 81% weiblich) teil. Die Studienwahlmotivation wurde mit dem FEMOLA (Pohlmann & Möller, 2010) erhoben. Zur Erfassung der Lehrereffektivität (Leidinger & Perels, 2015), des selbstregulierten Lernens (Dörrenbächer & Perels, 2016) sowie den Einstellungen zum Lernen und Lehren (Förderung selbstregulierten Lernens, Wissen über Strategieinstruktion; Dignath-van Ewijk & van de Werf, 2012) wurden eigene Instrumente entwickelt, die auf etablierten Skalen basieren. Alle Instrumente zeigen zufriedenstellende bis sehr gute Reliabilitäten. Als Leistungsmaß wurde bei den Studierenden die Durchschnittsnote im Studium, bei den Referendaren die Staatsexamensnote erfasst.

Eine latente Profilanalyse mit den Werten des FEMOLA (Marsh, Lüdtke, Trautwein & Morin, 2009) resultierte in einer Dreigruppen-Lösung, wobei die Gruppeneigenschaften denen von Billich-Knapp et al. (2012) entsprechen (intrinsisch, extrinsisch, Mischtyp). Diese Profile konnten unabhängig vom Geschlecht und von der Untersuchungsgruppe (Studierende – Referendare) gefunden werden. Univariate Varianzanalysen mit den Profilen als unabhängige Variablen zeigen signifikante, hypothesenkonforme Unterschiede in den abhängigen Variablen selbstreguliertes Lernen ( $F(2, 166) = 3.85$ ,  $p < .05$ ), Lehrereffektivität ( $F(2, 165) = 9.99$ ,  $p < .001$ ) und Einstellungen zum Lernen und Lehren ( $F(2, 161) = 6.88$ ,  $p < .01$ ), insofern, als Studierende mit vorrangig intrinsischer Motivation bzw. mit einem gemischten Motivationsprofil höhere Werte aufweisen, die in allen drei Maßen als positiv und wünschenswert zu werten sind. Bezogen auf die Durchschnittsnote im Studium zeigte sich kein signifikanter Unterschied zwischen den Profilen; bezogen auf die Staatsexamensnote sprechen die Ergebnisse einer univariaten Varianzanalyse bei den Referendaren für Vorteile des intrinsischen Motivationsprofils ( $F(2, 32) = 3.34$ ,  $p < .05$ ). Implikationen für weitere Forschung in Bezug auf die Studienwahlmotivation und daraus resultierende Konsequenzen für die Hochschullehre werden diskutiert.

## **Annahmen angehender Lehrkräfte zum Lernen mit multiplen Repräsentationen**

**Alexander Eitel, Anna Dersch, Alexander Renkl**

Universität Freiburg, Deutschland; [alexander.eitel@psychologie.uni-freiburg.de](mailto:alexander.eitel@psychologie.uni-freiburg.de)

### Theoretischer Hintergrund & Fragestellungen

Unterrichtsmaterialien in Form von Schulbuchaufgaben oder Arbeitsblättern enthalten zumeist multiple Repräsentationen. Abbildungen und Gleichungen werden zusätzlich zu einem geschriebenen Text verwendet, um einen Sachverhalt zu verdeutlichen. Laut empirischer Forschung ist das besonders dann förderlich, wenn die zusätzliche Repräsentation (z.B. das Bild) kognitive Prozesse des Wissenserwerbs (z.B. aus einem Text) unterstützt (Renkl & Scheiter, 2017). Insbesondere sind laut Ainsworth (2006) multiple Repräsentationen dann verständnisförderlich, wenn sie komplementäre Rollen erfüllen (complement), den Raum an falschen Interpretationsmöglichkeiten gegenseitig einschränken (constrain), oder den Aufbau eines tieferen Verständnisses für den Sachverhalt fördern (construct). Für (angehende) Lehrkräfte ist es entsprechend wichtig zu wissen, wann multiple Repräsentationen so gestaltet sind, dass sie diese kognitiven Funktionen erfüllen und sich mit höherer Wahrscheinlichkeit verständnisförderlich auswirken. Gegenstand der vorliegenden Untersuchung ist daher, inwiefern angehende Lehrkräfte solche kognitiven Funktionen in Unterrichtsmaterialien akkurat identifizieren können. Darüber hinaus ist Gegenstand der Untersuchung, ob angehende Lehrkräfte Misskonzepte über das Lernen mit multiplen Repräsentationen aufweisen, wie sie in der populärwissenschaftlichen Literatur im Bildungsbereich zu finden sind (siehe Beispiele in Weidenmann, 2002).

### Methode

Insgesamt nahmen 88 Lehramtsstudierende (62 weiblich;  $M = 23.94$  Jahre;  $M = 4.80$  Semester;  $M = 2.62$  Studien-/Unterrichtsfächer) der Pädagogischen Hochschule Freiburg an der Befragungsstudie teil (korrelatives Querschnittsdesign). Den Studierenden wurde zunächst ein Misskonzept-Fragebogen vorgelegt. Er enthielt fünf ja/nein-Aussagen zu verbreiteten Misskonzepten im Bereich des Lernens mit multiplen Repräsentationen, z.B. „Über multiple Sinneskanäle lernt man besser als über einen Sinneskanal“, „Das Lernen mit Texten und Bildern spricht beide Gehirnhälften an, wobei das Lernen mit Texten vor allem eine anspricht“ (vgl. Weidenmann, 2002). Im Anschluss sahen Studierende pro Seite ein text-bildbasiertes Lernmaterial (z.B. über Modelle in der Chemie), inklusive Informationen zur Zielgruppe und Unterrichtssituation, mit der offenen Frage „Welche Funktion erfüllt das Bild?“. Studierende sollten auf die Frage, welche Funktion das Bild erfülle, so ausführlich wie möglich antworten. Die Freitext-Antworten der Studierenden wurden entsprechend der Ainsworth-Taxonomie (2006) pro Material danach kodiert, wie häufig die kognitiven Funktionen complement, construct, constrain darin enthalten waren. Die Kodierung erfolgte durch zwei Rater (Inter-Rater Übereinstimmung:  $.61 > ICC < .82$ ). Die Häufigkeit der Nennung der einzelnen kognitiven Funktionen durch die Studierenden wurde pro Material im Anschluss mit der Expertenmeinung verglichen, um die Akkuratess der Nennung der kognitiven Funktionen zu ermitteln.

### Ergebnisse & Schlussfolgerungen

Binomialtests gegen den Skalenmittelpunkt von 0.5 (d.h. Ratewahrscheinlichkeit) ergaben, dass die Zustimmung zu drei der fünf Aussagen, nämlich zu „Über multiple Sinneskanäle lernt man besser als über einen Sinneskanal“, „Aus Texten und Bildern lernt man immer besser als aus Texten“ sowie zu „Das Lernen mit Texten und Bildern spricht beide Gehirnhälften an, wobei das Lernen mit Texten vor allem eine anspricht“, überzufällig hoch ausgeprägt war, alle  $ps \leq .003$ . Da fehlendes Wissen zu diesen Aussagen mit Antworten auf Rateniveau einhergehen müsste (wie es für die restlichen beiden Aussagen auch der Fall war), sprechen diese Daten für das Vorhanden fehlerhaften Wissens bzw. spezifischer Misskonzepte der angehenden Lehrkräfte in diesem Bereich. In den Freitext-Antworten ergaben sich ferner signifikante Unterschiede in der Benennung kognitiver Funktionen durch Studierende versus Experten. Insbesondere zeigen die Ergebnisse, dass auf die kognitiven Funktionen complement und constrain sehr selten durch die Studierenden hingewiesen wurde, selbst wenn sie in der Expertenantwort enthalten waren. Die mittlere Abweichung zur Expertenantwort fiel für die constrain-Funktion höher aus als für complement-Funktion,  $p < .001$ , welche wiederum höher ausfiel als für die construct-Funktion,  $p = .006$ . Dieses Ergebnis weist auf spezifische Wissenslücken in der Identifikation kognitiver Funktionen von multiplen Repräsentationen durch angehende Lehrkräfte hin. Die Gestaltung effektiver Interventionen zur Schließung dieser Wissenslücken sowie zur Adressierung der spezifischen Misskonzepte ist entsprechend Gegenstand geplanter Forschung.

## F05

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Rosshofgasse, S01

### Studien zur Modellierung des Professionswissens in der Teilchenphysik

**Michaela Oettle, Silke Mikelskis-Seifert**

Pädagogische Hochschule Freiburg, Deutschland; [Michaela.Oettle@ph-freiburg.de](mailto:Michaela.Oettle@ph-freiburg.de)

Das Projekt zielt auf eine Modellierung des Professionswissens ab, welches - vorrangig von gymnasialen - Lehrpersonen für den Unterricht teilchenphysikalischer Themen benötigt wird. Im Zuge der Modellierung soll insbesondere ein Einblick in die tiefere Struktur des Fachwissens als Teil des Professionswissens gegeben werden.

Teilchenphysik als Bereich der modernen Physik übt aufgrund einer rasant fortschreitenden Forschung und eines hohen Medieninteresses eine Faszination auf Schüler\*innen aus. Gleichzeitig kann die Teilchenphysik große Herausforderungen an die Lehrenden stellen in Folge der Komplexität und des fortgeschrittenen Niveaus des Themas innerhalb der Physik.

Das Professionswissen, welches Lehrpersonen für diese Aufgabe benötigen, wird als kognitiver Anteil im Modell professioneller Handlungskompetenz von Baumert und Kunter (2006) verortet. Diesem Modell folgend gliedert sich das Professionswissen weiter auf in die Wissensbereiche des Fachwissens, des Pädagogischen Wissens, des Fachdidaktischen Wissens, des Organisations- sowie des Beratungswissens.

Im Rahmen des vorgestellten Projektes wird ein Fokus auf die Untersuchung des fachwissenschaftlichen Teilbereichs gelegt. Studien zur Konzeptualisierung des Fachwissens von Lehrpersonen bzw. von Lehramtsstudierenden in der Physik konzentrieren sich hauptsächlich auf die Benennung von Wissenskomponenten (physikalische Themen) sowie auf die Identifikation von Wissensstufen (siehe unter anderem Riese et al. 2015). Während kürzlich auch weiter ausdifferenzierte Dimensionen wie die Hierarchische Komplexität dem Modell hinzugefügt wurden (Woitkowski & Riese, 2017), fokussierte sich die Forschung bislang nur auf die Bereiche Mechanik und Elektrodynamik.

Für das Gebiet der Teilchenphysik liegen noch keine Modellierungsergebnisse zum Fachwissen von Lehrpersonen vor. Infolgedessen wird im Rahmen dieses Projektes ein neues Modell für das Fachwissen entwickelt, mit dem Ziel, nicht nur eine Aufstellung der relevanten Themenbereiche zu geben, sondern eine tiefere inhaltliche Strukturierung und Vernetzung zu dokumentieren.

Im Zuge der Kohärenz von Lehre und fachdidaktischer Forschung lehnt sich das Forschungsdesign für eine erste Konzeptualisierung – die Benennung der relevanten Wissenskategorien - an folgende Form der Delphi-Methode an: In der ersten von mehreren konsekutiven Befragungsrunden wurden internationale Experten aus Forschung, Lehre und Öffentlichkeitsarbeit gebeten, die ihrer Ansicht nach zentralen Konzepte (engl. key concepts bzw. big ideas) in der Teilchenphysik zu nennen. Methodisch wurde hierzu ein Online-Fragebogen mit offenen Antwortformaten verwendet. Die Angaben der Experten wurden nach den Richtlinien einer zusammenfassenden induktiven Inhaltsanalyse (Mayring, 2015) in Kategorien und Unterkategorien eingeordnet und diese wurden wiederum mit Standardlehrwerken der Teilchenphysik abgeglichen. Zur Überprüfung des Systems wurden die Antworten von zwei unabhängigen Beurteilern kodiert, was in einer sehr hohen Interrater-Reliabilität resultierte (Cohens  $\kappa = 0.95$ ). Häufigkeitsanalysen der von den Experten genannten Aspekte wurden für eine erste Reduktion des Kategoriensystems verwendet. Als Resultat ergab sich ein System von 10 Hauptkategorien mit insgesamt 46 Unterkategorien.

In der zweiten Runde der Delphi-Befragung wurde das konstruierte Kategoriensystem derselben Population an Experten vorgelegt und eine Inhaltsvalidierung auf Basis der von den Experten abgegebenen Akzeptanzratings durchgeführt.

Dem Ablauf einer Delphi-Befragung folgend, soll eine weitere Befragungsrunde zu einem Konsens der Experten über die Relevanz der vorgeschlagenen (Unter-) Kategorien für das Fachwissen von Lehrpersonen führen bevor im Anschluss eine tiefergehende Analyse der Sachstruktur aufbauend auf den bislang rein begrifflich definierten Kategorien stattfinden kann. Für diesen letzten Schritt bietet sich die Methode des Concept-Mapping an, um unter anderem Zusammenhänge und Teilstrukturen sichtbar zu machen.

Die Ergebnisse dieser strukturierenden Modellierung können als Basis für die Entwicklung eines Testinstruments zur Erfassung professioneller Kompetenzen von Lehrkräften im Bereich der Teilchenphysik dienen. Und dieses Wissen unter anderem über die bereits bestehenden konzeptuellen Vorstellungen der Lehrenden gilt wiederum als Voraussetzung, um die Struktur und die Inhalte von zukünftigen Lehreraus- und fortbildungsprogrammen im Bereich der Teilchenphysik zu bestimmen.

Im Vortrag werden die Ergebnisse aus der ersten Befragungsrunde in Form des Kategoriensystems präsentiert, ein Einblick in die Modellvalidierung sowie ein Ausblick auf die folgenden Befragungsrunden gegeben.

### Quasi-experimentelle Haltung von Lehrpersonen als integraler Teil professioneller Handlungskompetenz

**Andreas Seifert<sup>1</sup>, Daniel Scholl<sup>2</sup>, Wilfried Plöger<sup>2</sup>**

<sup>1</sup>Universität Paderborn, Deutschland; <sup>2</sup>Universität zu Köln, Deutschland; [seifert@plaz.upb.de](mailto:seifert@plaz.upb.de)

Lehrpersonen verbinden mit ihrer firsthand experience eine hohe Wertschätzung und oft die Überzeugung, that “they have learned from classroom experience most of what they know about teaching” (Buchmann & Schwille, 1983, p. 30). Diese Überzeugung impliziert zugleich die Annahme, dass sich die Kompetenz einer Lehrperson in Abhängigkeit von der Dauer ihrer Berufserfahrung kontinuierlich weiterentwickelt. Im Gegensatz zu dieser Annahme lieferten empirische Studien bisher jedoch keine Evidenz für den Zusammenhang zwischen der Dauer der Berufserfahrung einer Lehrperson und der Ausprägung ihrer Kompetenzen.

Die Ergebnisse solcher Studien legen demgegenüber eher den Schluss nahe, dass das Sammeln von Erfahrungen nicht automatisch zu einer Verbesserung von Kompetenzen führt. Hinzukommen muss – so wird angenommen – eine kritische Haltung gegenüber der eigenen Erfahrung, um aus ihr systematisch zu lernen und dadurch seine Kompetenzen ständig zu verbessern (Hiebert, Morris, Berk, & Jansen, 2007; Loughran, 2002).

Diese kritische Haltung bezeichnen wir als quasi-experimentelle Haltung (QEH). Sie zeigt sich in der Bereitschaft von Lehrpersonen, ihre Kerntätigkeit – die Planung, Durchführung und Analyse von Unterricht – in Analogie zum Experimentieren in der Forschung wahrzunehmen und zu gestalten, aus diesem zyklischen Prozess bedeutungsvolle Erfahrungen zu gewinnen, diese Erfahrungen kritisch zu evaluieren und aus ihnen Konsequenzen zur Verbesserung ihres künftigen Unterrichts herzuleiten.

Bisher vorliegende Instrumente, die die Einstellung gegenüber kritischem Denken oder kritischer Reflexion messen, zeigen zwei Desiderate: Einige Instrumente (z.B. Kember & Leung, 2000) sind zu breit angelegt, weil sie solche Einstellungen berufsübergreifend definieren und daher nicht auf die professionellen Kernaktivitäten von Lehrpersonen zielen. Andere Instrumente, die enger auf das Berufsfeld von Lehrpersonen abgestimmt sind, fokussieren einseitig auf Aspekte des Denkens (z.B. Sobral, 2000) oder des Handelns (z.B. Kayapinar & Erkus, 2009), anstatt diese beiden Aspekte als einen zusammenhängenden Prozess (Mena Marcos et al., 2009) zu modellieren.

Vor dem Hintergrund dieser Desiderate haben wir eine Studie durchgeführt mit dem Ziel, ein Instrument zur Messung der QEH von Lehrpersonen zu entwickeln und zu erproben.

An der Studie nahmen 512 Personen aus sechs Studienseminaren teil, die im Rahmen der zweiten Ausbildungsphase ihr Referendariat absolvierten und zum Zeitpunkt der Untersuchung über eine Unterrichtserfahrung von durchschnittlich 14,8 Monaten (SD = 1.17) verfügten.

Den Teilnehmer(innen) wurde ein zuvor mehrfach pilotierter Fragebogen vorgelegt, mit dem die Ausprägung der QEH anhand von 12 Items gemessen wurde. Zusätzlich wurden fünf Variablen zur Beurteilung der kriterialen Validität der Skala zur QEH erhoben (Selbstwirksamkeit, Belastungserleben, Einstellungen zum Eingangs- und Perspektivengespräch, zum Coaching und zum Führen von Portfolios).

Die Auswertung der Daten folgte den üblichen statistischen Prozeduren (Bestimmung von Trennschärfeindizes, der Reliabilität der Skala, der Korrelationen zwischen den Skalen). Darüber hinaus wurden konfirmatorische Faktorenanalysen durchgeführt, um die faktorielle Struktur des Konstrukts quasi-experimentelle Haltung zu bestimmen.

Die Erprobung des Instruments zeigte, dass die QEH von Lehrpersonen mit einer kurzen Skala auf reliable und valide Weise gemessen werden konnte. Die Reliabilität der Skala lag bei .75. Die durchgehend signifikanten Korrelationen ( $p < .01$ ) zwischen der QEH-Skala und den anderen fünf Skalen stützen die Validität des Konstrukts QEH, wobei die QEH-Skala mit der Skala zum Belastungserleben negativ und mit den anderen vier Kriterien positiv korrelierte.

Die konfirmatorischen Faktorenanalysen decken zwei verschiedene Modelle: (1) ein first-order-Modell, bei dem die Items verteilt auf vier interkorrelierten Faktoren laden (akribische Planung von Unterricht, Zuversicht in die Vorhersagbarkeit des späteren Unterrichtsverlaufs, evidenz-basierte Analyse von Unterricht, Skepsis gegenüber unreflektierter Erfahrung) und (2) ein second-order-Modell, in dem die vier Faktoren auf einem übergreifenden Faktor QEH laden.

Vor dem Hintergrund dieser positiven Ergebnisse nutzen wir die QEH-Skala derzeit in einer von der DFG geförderten Längsschnittstudie, um die Bedeutung der QEH für die Entwicklung professioneller Kompetenzen zu bestimmen.

## **Die Fertigkeit zur Beurteilung der Formulierung von Lernzielen als Teil des professionellen Unterrichtshandelns von Lehrkräften**

**Joerg Wittwer, Linda Kratschmayr**

Universität Freiburg, Deutschland; [joerg.wittwer@ezw.uni-freiburg.de](mailto:joerg.wittwer@ezw.uni-freiburg.de)

### Theoretischer Hintergrund

Lernziele sind Aussagen, die angeben, was Lernende nach Abschluss eines Lernprozesses wissen und können sollen. Sie spielen aus instruktionspsychologischer Perspektive eine zentrale Rolle bei der Planung des Unterrichts und bei der Überprüfung des Lernerfolgs von Schülerinnen und Schülern (Klauer & Leutner, 2012), da nur auf der Grundlage von festgelegten Lernzielen passende Lehr-Lern-Formen und passende Prüfungsformen ausgewählt werden können (Prinzip des constructive alignment, Biggs & Tang, 2011). Entsprechend weisen metaanalytische Befunde nach, dass Lernziele eine hohe Lernwirksamkeit besitzen (z.B. Hattie, 2009). Damit Lernziele als pädagogisches Werkzeug effektiv eingesetzt werden, ist es allerdings wichtig, dass sie eindeutig formuliert werden (Gronlund & Brookhart, 2009), da ansonsten unklar bleibt, was Schülerinnen und Schüler nach Beendigung eines Lernprozesses leisten sollen.

### Fragestellung

Trotz der Bedeutung von Lernzielen ist bislang empirisch wenig bekannt, was Lehrkräfte über Lernziele und speziell über ihre Formulierung wissen und wie sie gegenüber Lernzielen eingestellt sind. Deshalb untersuchten wir, (1) wie Lehrkräfte gut und schlecht formulierte Lernziele einschätzen, (2) ob Lehrkräfte typische Fehler in der Formulierung von Lernzielen erkennen, (3) welche Einstellung Lehrkräfte gegenüber der Verwendung von Lernzielen aufweisen und (4) welcher Zusammenhang zwischen der Einstellung gegenüber Lernzielen und der Einschätzung der Formulierung von Lernzielen besteht.

### Methode

An der Untersuchung nahmen N = 122 Lehrkräfte von Gymnasien, Realschulen, Gesamtschulen und Berufsschulen teil. Von den 122 Lehrkräften waren 60 weiblich und 57 männlich (5 Lehrkräfte gaben ihr Geschlecht nicht an). Im Mittel waren die Lehrkräfte M = 44.64 Jahre (SD = 10.83) alt. In der Untersuchung schätzten die Lehrkräfte die Formulierung von 13 vorgegebenen Lernzielen auf einer 4-stufigen Skala von 1 = schlecht bis 4 = gut ein. Dabei konnten sie zusätzlich eine Begründung für ihre jeweilige Einschätzung abgeben. Die Lernziele bezogen sich auf Inhalte aus unterschiedlichen Fächern (z.B. Deutsch, Mathematik). Drei Lernziele waren ohne absichtlichen Fehler formuliert. Zusätzlich wurden jeweils zwei Lernziele zu den folgenden fünf Fehlertypen formuliert: (1) Das Lernziel besteht aus einer Aussage mit mehreren Lernzielen, (2) Das Lernziel betont den Lernprozess und nicht das Lernergebnis, (3) Das Lernziel ist kein Lernziel, sondern ein Lernziel, (4) Das Lernziel ist unspezifisch, und (5) Das Lernziel enthält kein Tätigkeitsverb. Dann bearbeiteten die Lehrkräfte sechs Aussagen, mit denen ihre Einstellung gegenüber der Verwendung von Lernzielen erhoben wurde. Schließlich machten die Lehrkräfte Angaben zu ihrer Person.

### Ergebnisse

Die Ergebnisse zeigten, dass die Lehrkräfte zwar gut formulierte Lernziele von schlecht formulierten Lernzielen signifikant unterscheiden konnten, aber ihre Beurteilung der schlecht formulierten Lernziele im Mittel eher milde ausfiel. Bei allen Lernzielen erstreckte sich die Spannweite für die Beurteilung über die gesamte Skala, was eine hohe Varianz in der Einschätzung der Formulierung der Lernziele zwischen den Lehrkräften widerspiegelt. Auch gaben die Lehrkräfte in ihren Begründungen für ihre Einschätzungen häufig nicht die Fehler an, die Grundlage für die Formulierung der Lernziele waren, sondern nannten Aspekte, die nicht unmittelbar mit der Angemessenheit der Formulierung von Lernzielen zu tun hatten (z.B. Relevanz eines Lernziels für das Lernen). Insgesamt waren die Lehrkräfte vergleichsweise positiv gegenüber der Verwendung von Lernzielen eingestellt. Zugleich zeigte sich, dass die Lehrkräfte die schlecht formulierten Lernziele signifikant umso kritischer beurteilten, je positiver ihre Einstellung gegenüber der Verwendung von Lernzielen war.

## Diskussion

Insgesamt deuten die Befunde darauf hin, dass Lehrkräfte Lernzielen eine hohe Bedeutung zumessen, sie aber gleichzeitig offensichtlich kein einheitliches Verständnis darüber besitzen, wie ein Lernziel gut formuliert werden sollte und welche typischen Fehler bei der Formulierung von Lernzielen auftreten können. Zudem legen die Ergebnisse nahe, dass eine positivere Einstellung gegenüber der Verwendung von Lernzielen dazu führt, sich intensiver mit Lernzielen und deren angemessenen Formulierung auseinanderzusetzen.

## Das Wissen über Lehr-Lernmethoden von Lehrkräften der Erwachsenen- und Weiterbildung – Zur validen Messung eines mehrdimensionalen Konstrukts

**Christian Marx<sup>1</sup>, Annika Goeze<sup>1</sup>, Thamar Voss<sup>2</sup>, Viola Deutscher<sup>3</sup>, Verena Hoehne<sup>4</sup>, Josef Schrader<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen e.V., Deutschland; <sup>2</sup>Universität Freiburg; <sup>3</sup>Universität Mannheim; <sup>4</sup>Hector-Institut für Empirische Bildungsforschung; [marx@die-bonn.de](mailto:marx@die-bonn.de)

### Theoretischer Hintergrund

Wie für den Schulkontext vielfach empirisch gesichert (Hattie, 2009), wird auch für die Erwachsenenbildung (EB) angenommen, dass Lehrkräfte durch ihre professionelle Kompetenz maßgeblich für den Lernerfolg von TeilnehmerInnen mitverantwortlich sind (Schrader, 2010). Der Berufszugang und die Berufsausübung in der EB sind kaum staatlich oder tarifvertraglich reglementiert und damit seit jeher „offen“ für Lehrkräfte mit unterschiedlichen Bildungsbiographien. Vor diesem Hintergrund ist es verständlich, dass in den vergangenen Jahren das Interesse an der Erfassung der Kompetenzen des Lehrpersonals gestiegen ist, bislang liegen hierzu allerdings kaum gesicherte Befunde über die Kompetenzen von Lehrkräften der EB vor. Vielmehr hat die Entwicklung von Instrumenten zur Erfassung professioneller Kompetenzen erst begonnen (Marx et al., 2017; Rohs et al., 2017).

Vermehrt auch für den Bereich der EB adaptiert (z.B. Lencer & Strauch, 2016; Marx et al., 2017) wird das ursprünglich für den Schulkontext entwickelte und dort empirisch bewährte Rahmenmodell der professionellen Kompetenz von Lehrkräften von Baumert und Kunter (2011), weil es generische und damit für die EB relevant erscheinende Bereiche der professionellen Kompetenz thematisiert. Dieses Modell liegt auch dem Projekt Think zugrunde, das an die Arbeit von Voss, Kunter und Baumert (2011) anschließt und in dem das pädagogisch-psychologische Wissen (PPK) als Teil des professionellen Wissens in acht Facetten untergliedert und anschließend an diese ein Instrument zur Erfassung des PPK von Lehrkräften verschiedener Bildungsbereiche entwickelt wird (Marx et al., 2017).

Über verschiedene Projekte hinweg stützen Befunde zur Frage nach der Dimensionalität des PPK die Annahme, dass die Facetten einem gemeinsamen Konstrukt - bestehend aus mehreren Facetten - zugeordnet werden können (Voss et al., 2015). Zugleich stellt sich allerdings die Frage, welches Wissen konkret das allen Facetten gemeinsame und welches das ihnen spezifische ist; in nuce stellt sich diese Frage auch für die einzelnen Facetten des PPK.

### Fragestellung

Der vorliegende Beitrag greift diese Frage auf und führt die in Marx et al. (2017) begonnene Entwicklung eines Instruments zur Erfassung des PPK durch die Erfassung der Facette „Wissen über Lehr-Lernmethoden und -konzepte“ fort, die für die Gestaltung von Lehr-Lernprozessen in der EB allgemein als zentral betrachtet wird. Am Beispiel dieser Facette, erfasst durch fünf Aufgaben zu verschiedenen Lehr-Lernmethoden, werden drei Forschungsfragen gestellt:

- 1) Welche Struktur zeigt sich für die fünf Aufgaben dieser Facette?
- 2) Was ist das gemeinsame bzw. spezifische Wissen der Aufgaben?
- 3) Zeigen sich DIF-Effekte für Lehrkräfte mit vs. ohne pädagogischer Vorbildung?

### Methode

Alle Analysen wurden mit der R Statistical Software (R Development Core Team, 2017) durchgeführt an einer Stichprobe von N = 212 EB-Lehrkräften. Zur Beantwortung der Frage 1 wurden ein eindimensionales Modell, ein Correlated Factors Modell, ein bifactor und ein Rasch testlet Modell konfirmatorisch mit dem TAM Package (Robitzsch et al. 2017) auf die Daten angepasst und anhand der Informationskriterien AIC, AICc, BIC und CAIC auf ihre Anpassungsgüte hin verglichen. Dabei wurden fünf Correlated Factors bzw. fünf Gruppenfaktoren (pro Aufgabe einer) und ein Generalfaktor spezifiziert. Zur Beantwortung der dritten Fragestellung wurden DIF-Analysen mit dem TAM Package durchgeführt (Robitzsch et al., 2017).

### Ergebnisse

Die informationstheoretischen Maße AIC, AICc, BIC und CAIC deuten auf die Überlegenheit des Rasch testlet Modells hin, für das die Aufgaben mit Weighted Mean Square-Werten (WMNSQ) zwischen 0.92 und 1.12 eine gute Modellanpassung zeigten, die Streuung der Aufgabenschwierigkeiten kann in Relation zu den Personenfähigkeiten als gut bezeichnet werden (-4.01 logits ≤ Aufgabenschwierigkeit ≤ +0.97 logits). Der Generalfaktor kann als Wissen über tiefenstrukturelle Qualitätskriterien interpretiert werden und soll exemplifiziert an Aufgaben, Gegenstand der Diskussion sein; alternative Erklärungen des Generalfaktors, z.B. allgemeine kognitive Fähigkeiten, können nicht ausgeschlossen werden. Nur für eine Antwortoption zeigte sich ein moderater DIF-Effekt. Mit dem vorgestellten Test kann das Wissen über Lehr-Lernmethoden bei typischerweise unterschiedlich vorgebildeten Lehrkräften der EB angemessen erfasst werden.

## F06

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus S212

### Klausurmodellierung in der Technischen Thermodynamik mittels Item Response Theory – Ergebnisse einer hochschuldidaktischen Begleitforschung

**Thorsten Braun**

Universität Stuttgart, Deutschland; [thorsten.braun@zlw.uni-stuttgart.de](mailto:thorsten.braun@zlw.uni-stuttgart.de)

Zur Professionalisierung der grundlegenden Ingenieurausbildung an Hochschulen gehört zweifelsohne die Gestaltung von Prüfungen (Huber 2007; Coates 2014). Ein zentraler Faktor ist dabei, welche qualitativen Merkmale Prüfungen im Grundlagenstudium besitzen und welche Güte sie hinsichtlich der Beurteilung von Personenfähigkeiten und Kompetenzen aufweisen (Nickolaus 2013). Der hier vorgestellte Beitrag behandelt diese Fragen für eine Modulabschlussprüfung der Technischen Thermodynamik. Die Prüfung ist für Studierende aus sechs Studiengängen an einer deutschen Universität obligatorisch und schließt ein zweisemestriges Studienmodul ab, das in der Regel im dritten und vierten Semester besucht wird.

Prüfungen kommt eine ganz besondere Bedeutung für die Gestaltung von Lernprozessen (Biggs.2009), die Sicherung von Lernerfolg und die Beurteilung erworbener Kompetenzen (Schindler.2015; Jaudzims.2013) zu.

Ein häufiger Kritikpunkt an einer weit verbreiteten Prüfungspraxis betrifft das Korrekturverfahren, wenn es sich auf gewichtete Summenscores stützt und dabei die Möglichkeiten der testtheoretischen Modellierung nicht berücksichtigt (Bühner 2011; Strobl 2015). Es ist also aus praktischer wie theoretischer Perspektive wünschenswert, wenn die Gestaltung von Hochschulprüfungen im Rahmen eines reflexiven Prozesses stattfindet, der die Leistungsfähigkeit der Prüfungen untersucht und auch die Lehrenden bei der Ausschöpfung von wissenschaftlich fundierten Gestaltungsmöglichkeiten unterstützt.

Dabei wird von theoretischer Seite angenommen, dass eine Rekonstruktion von Kompetenzstrukturen in Prüfungen prinzipiell möglich ist und dass sich Ergebnisse dieser Rekonstruktion sinnvoll auf Learning Outcomes und Lernprozesse beziehen lassen (Schindler 2015). Wenngleich die damit verbundenen, notwendigen Anstrengungen beachtlich sind (Nickolaus 2011), gewinnt die Prüfung durch eingehende Analyse an Transparenz für Lehrende wie Studierende gleichermaßen. Dies kommt einer Professionalisierung von Lehrenden und Studierenden im Umgang mit Prüfungsgestaltung und Prüfungsbewältigung zugute. Diese Professionalisierung zu begleiten und anzuregen, ist Aufgabe der empirischen Bildungsforschung.

Zentrale Fragestellungen für das vorgestellte Forschungsvorhaben sind: A) Wie lässt sich die von den Lehrenden ausgearbeitete Prüfung mit Hilfe der Item Response Theory (IRT) angemessen modellieren? B) Welche Erkenntnisse zur Prüfungsoptimierung können aus der Modellierung gewonnen werden? C) Welche persönlichen und studienbezogenen Merkmale weisen erfolgreiche und nicht erfolgreiche Studierende auf?

Als Ausgangspunkt für die Modellierung der Prüfung diente ein einfaches IRT-Modell (Rasch/1PL, Rost 2004; Bühner 2011). Sukzessive wurden komplexere Modelle angewendet (2PL; Rasch Testlet Model; Wang 2005) und untereinander hinsichtlich ihrer Güte und ihres Erkenntnisertrags verglichen. Dabei war insbesondere von Interesse, die Auswirkungen verschiedener Codier- und Modellierungsverfahren im Detail zu betrachten, um deren praktische Bedeutung genauer abschätzen zu können.

Die IRT-Modellierung fand im Kontext eines größeren, methodenintegrativen Designs statt (mixed method; Kelle 2008, 2014), in dem die Modellierungsergebnisse zur Studierendenfähigkeit und den Aufgabeneigenschaften auf weitere, quantitative und qualitative, Daten der untersuchten Studierendenkohorte (N=322) bezogen wurden.

[Ergebnisse]

Die Ergebnisse zeigen, dass bei der Modellierung der Thermodynamik-Prüfung eine Codierstrategie am vorteilhaftesten ist, die sowohl statistische wie inhaltliche Merkmale der Korrekturbewertungen berücksichtigt.

Die beste Modellpassung wird durch ein 1PL-Modell mit WML-Schätzern erzielt (Warm 1989). Dies ist ein überraschender Befund. Obwohl die Testlet-Struktur (Wainer 2007) der Prüfung durch explorative wie auch konfirmatorische Verfahren nachgewiesen wird (Chen 1997; Zhang 1999), kann eine essentielle Eindimensionalität (Reise 2016) der Prüfung festgestellt werden. Das hat enorme praktische Konsequenzen für die weitere Arbeit. Zum einen können die Parameterschätzungen des 1PL-Modells als robust und gut angepasst verwendet werden, zum anderen ergeben sich durch die Analyse der Haupt- und Nebenfaktoren der Testlet-Struktur tiefere Erkenntnisse über den Beitrag, den einzelne Aufgabenbündel für die Feststellung der Personenfähigkeit leisten. Dies ergibt einen reichhaltigen Fundus von Anknüpfungspunkten für eine Weiterentwicklung der Prüfungsgestaltung.

Die Reliabilität des bevorzugten 1PL-Modells beträgt 0.77. Es mussten im Verlauf der Kalibrierung zwei von zweiundzwanzig Items aus statistischen (schlechter Itemfit) sowie inhaltlichen Gründen (mehrdeutige Aufgabenstellung) ausgeschlossen werden. Bezüglich der Eindimensionalität sind 87% der im Modell erklärten Varianz auf Unterschiede im Hauptfaktor zurückzuführen (OmegaH 0.70; ECV 0.87; PUC 0.79).

### Themen- und Erfolgsbias in Datenbeständen der empirischen Bildungsforschung? Eine Analyse auf Basis der NEPS-Daten

**Uta Liebeskind<sup>1</sup>, Marie Kühn<sup>2</sup>**

<sup>1</sup>DZHW, Deutschland; <sup>2</sup>DZHW, Deutschland; [liebeskind@dzhw.eu](mailto:liebeskind@dzhw.eu), [kuehn@dzhw.eu](mailto:kuehn@dzhw.eu)

Erkenntnisse aus standardisierter empirischer Bildungsforschung sind zu einem Schlüssel für professionelles Handeln geworden: Erstens ist es ganz allgemein ein zentrales Merkmal professionellen Handelns, auf akademisch generiertes Wissen zurückzugreifen (siehe z.B. Mieg 2003; Pfadenhauer 2003), welches in erziehungs-, sozial- und wirtschaftswissenschaftlichen Disziplinen zu einem Gutteil auf Ergebnissen aus standardisierter Datenerhebung fußt. Zweitens greifen politische Entscheidungsträger auf Daten und Ergebnisse von bildungsbezogenen Forschungsinstituten zurück (DJI, BIBB, DZHW, etc.) und die Wissenschafts- und Bildungspolitik ermöglicht große large scale-Datenerhebungen wie die Beteiligung am Programme for the International Assessment of Adult Competencies (PIACC) oder der Etablierung des Nationalen Bildungspanels (NEPS).

Umso wichtiger ist es, ganz prominent auch die Bedingungen zu benennen, unter denen der Rückgriff auf große Datenbestände zulässige Schlüsse erlaubt. Eine zentrale Frage für die Einschätzung der Qualität von large scale-Datenbeständen ist diejenige, ob bei der Datensammlung Verzerrung durch systematischen Ausfall auftritt: Über wen erzählen die Daten überhaupt etwas, über wen nicht?

Für die Datenbestände des NEPS, eine der zentralen Datengrundlagen für empirische Bildungsforschung in Deutschland fragen wir: Gibt es einen Teilnahmebias 1) im Blick auf persönlichen Erfolg oder 2) im Blick auf die persönliche Nähe zum Thema der Studie, also Bildung? Bislang liegen nur wenig empirische Arbeiten vor, die diesen beiden Fragen nachgehen. Für das Thema der Studie zeigt die Non-Response-Forschung für Befragungen im Querschnitt ganz klar, dass fehlendes Interesse am Thema ein wichtiger Verweigerungsgrund ist (Loosveldt und Joye 2016b: 70; Durrant und Steele 2009). Für Erfolg ist die empirische Befundlage karg und in ihren Ergebnissen uneinheitlich (Birkelbach 1998; Windzio und Grotheer 2002). Misst man Erfolg über Kompetenzen, so lässt sich für die Erwachsenenkohorte des NEPS klar zeigen, dass hohe Lesekompetenz das Teilnahmeverhalten positiv beeinflusst (Stöckinger et al. ohne Jahr). Die Forschung zu diesen beiden Fragen kann sich bislang nur auf wenig ausgearbeitete theoretische Erklärungsmuster stützen. Konzeptionelle Ansatzpunkte liegen allerdings vor, etwa zur Bedeutung persönlicher Merkmale (Hox et al. 1995) oder zur Bedeutung des survey climate (Loosveldt und Joye 2016a), dessen Bedeutung für das Teilnahmeverhalten mit dem Einfluss des Themas der Studie über dessen Öffentlichkeitswirksamkeit interagiert.

Unser Beitrag erweitert die bereits vorliegenden Ergebnisse von Liebeskind und Vietgen (2017) zur NEPS-Studierendenkohorte, in dem zwei weitere Startkohorten des NEPS in die Analyse einbezogen werden: die Klasse-9-Kohorte und die Erwachsenenkohorte. Wir konzentrieren uns auf den Ausfallgrund Verweigerung (im Gegensatz zu Nicht-Erreichbarkeit und Ausfall aus der Panelpopulation durch Tod oder Umzug, zur Unterscheidung siehe z.B. Groves et al. 2009, S. 129 oder Schnell 2012, S. 329). Es wird der Einfluss von verschiedenen Erfolgsmaßen (Schul- und Abschlussnoten, individuelle Kompetenzmaße, Einkommen aus Erwerbsarbeit) ebenso wie der Einfluss des Interesses am Thema (über Ausbildungsberufe, Studienfächer, ausgeübte Berufe) auf das Teilnahmeverhalten modelliert. Dazu greifen wir auf längsschnittliche Analyseverfahren (Sequenzanalysen und Fixed Effects Models) zurück. Die ersten Ergebnisse zur NEPS-Studierendenkohorte (binäres logistisches Regressionsmodell für Teilnahmeverweigerung zwischen erster und zweiter CATI-Befragungswelle, siehe Liebeskind und Vietgen 2017) zeigen, dass ein besserer Schulabschluss ebenso wie das Studium eines bildungsnahen Studiengangs die Wahrscheinlichkeit der Teilnahmeverweigerung signifikant senkt.

### **Implementation and Implications of a new nonparametric estimation approach of a latent variable model**

**Tim Fabian Schaffland, Stefano Noventa, Augustin Kelava**

Universität Tübingen, Deutschland; [tim-fabian.schaffland@uni-tuebingen.de](mailto:tim-fabian.schaffland@uni-tuebingen.de)

A new nonparametric approach to estimate factor scores and their relationship – using nonparametric regression – will be discussed (Kelava, A., Kohler, M., Krzyżak, A. and Schaffland, T. F. (2017)). First the implementation in R will be presented and its application will be explained. Then the method will be investigated in two simulation studies. In a first simulation study it will be compared to two different ways to estimate structural equation models, i.e. the Structural Equation Mixture Modeling (SEMM) (Bauer (2005)) approach and the Latent Moderated Structural Equations (LMS) (Klein and Moosbrugger (2000)) approach. We will look at their performance for normal distributions with varying degrees of skewness and kurtosis and different functional relations between the latent variables. In a second simulation study the new approach will be compared to different factor score estimators, e.g. Bartlett's method (Bartlett (1937)) and the regression method (Thurstone (1935)). Here we will look at their performance for different normal and non-normal distributions to compare their biases and their consistency properties. Implications for applied research in education science will be discussed by giving an example for its usage in a substantive research context.

### **Zwei Verfahren zur Analyse von Messinvarianz im Vergleich: Mehrgruppen-CFA vs. IRT-basierter Item-Fit**

**Janine Buchholz, Johannes Hartig, He Jia**

Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Deutschland; [buchholz@dipf.de](mailto:buchholz@dipf.de)

Theoretischer Hintergrund

Ein Ziel internationaler Large-Scale Assessments (LSA) besteht darin, die anhand von Leistungstests und Fragebögen gemessenen Konstrukte zwischen einer großen Anzahl an Ländern zu vergleichen. Solche Vergleiche sind jedoch nur dann zulässig, wenn in allen Ländern tatsächlich das gleiche Konstrukt gemessen wurde („Messinvarianz“). Da zur Skalierung in LSAs häufig Modelle der Item-Response-Theorie (IRT) Anwendung finden, betrifft Messinvarianz die Gleichheit von Itemparametern des IRT-Modells. Zur Analyse der Messinvarianz wurde in PISA 2015 (OECD, 2016) ein innovativer Ansatz gewählt: Auf die gemeinsame Skalierung der Daten aus allen Ländern im Rahmen eines Generalized-Partial-Credit-Modells (GPCM; Muraki, 1992) folgte die Berechnung der RMSD (Root Mean Square Deviance)-Item-Fit Statistik für jedes Item in jedem Land. Sie zeigt die Abweichung zwischen modellimplizierter und beobachteter Item Characteristic Curve und somit die Passung der internationalen Parameter auf die Daten eines einzelnen Landes an (OECD, in press); gleichzeitig informiert sie jedoch nicht, welcher Modellparameter (Schwierigkeit vs. Diskrimination) vom internationalen Parameter abweicht. Im Unterschied zu dieser innovativen Methode wird Messinvarianz traditionell am häufigsten im Rahmen einer Mehrgruppen-CFA (MG-CFA; Jöreskog, 1971) untersucht (Cieciuch, Davidov, Schmidt, Algesheimer, & Schwartz, 2014): In einer Serie ineinander geschachtelter Modelle werden Sets von Parametern zwischen Gruppen gleichgesetzt und Veränderungen in der Modellgüte beobachtet, um so auf das Vorliegen gleicher Itemparameter zwischen Gruppen schließen zu können (Meredith, 1993). Im LSA-Kontext birgt dieser Ansatz jedoch zwei wesentliche Schwächen: zum einen ist die Anzahl der Gruppen häufig sehr groß, sodass die üblichen Kriterien über die zulässige Höhe der Veränderung in der Modellgüte nicht mehr gültig sind (Rutkowski & Svetina, 2014). Zum anderen liefert die Güte auf der Modellebene naturgemäß keine Information über die Ursache der Invarianzverletzung, also dasjenige Indikatoritem, dessen Parameterrestriktion mit Gruppenunterschieden einhergeht; gleichzeitig ist dies jedoch häufig von großem Interesse für gruppenvergleichende Analysen.

Fragestellung

Die beiden präsentierten Ansätze zur Untersuchung von Messinvarianz unterscheiden sich im Wesentlichen im zugrundeliegenden Messmodell (IRT vs. CFA) und der Analyseebene (Item-Fit vs. Modellgüte), dennoch können sie auf die gleichen Daten angewendet werden. Es stellt sich die Frage, ob die beiden Verfahren zu gleichen Entscheidungen hinsichtlich des Vorliegens von Messinvarianz für ein gegebenes Datenbeispiel liefern.

Methode

Zur Beantwortung der Fragestellung wurden Daten des Elternfragebogens aus PISA 2015 (OECD, 2016) herangezogen (N=86378). Alle acht Skalen wurden zunächst entlang der Dokumentation (OECD, in press) und anhand der in PISA 2015 verwendeten Software mdltm (von Davier, 2008) skaliert. Die RMSD-Statistik für jede Land\*Item-Interaktion wurde grafisch

aufbereitet, um so die Datenmenge zu reduzieren und eine Beispielskala zur Illustration identifizieren zu können. Die hier aufbereiteten Ergebnisse beziehen sich auf die Skala „Parents perceived school quality“, die anhand von sechs vierstufigen Items („strongly disagree“ bis „strongly agree“) gemessen wurde. Item PA007Q13NA („My child's school offers parent education (e.g. <courses on family literacy>) or family support programmes (e.g. <to assist with health, nutrition>).“) zeigte dabei als einziger Indikator dieser Skala einen besonders schlechten Item-Fit für Hongkong an (RMSD= 0.2623). Diese Skala wurde anschließend in einer MG-CFA in Mplus (Muthén & Muthén, 1998-2017) unter Annahme skalarer Messinvarianz geschätzt und die Modifikationsindizes analysiert.

#### Ergebnisse

Die Modifikationsindizes (ab einem Wert=10) weisen einen Median von M.I.= 68.18 auf. Der höchste Wert (M.I.=2131.49) bezieht sich auf das Intercept von PA007Q13NA in Hongkong; der Wert für die Faktorladung in Hongkong beträgt M.I.=580.99. Die Ergebnisse zeigen also in Übereinstimmung mit jenen des RMSD an, dass die Annahme gleicher Parameter für Hongkong mit schlechterer Passung einhergeht, wobei dies scheinbar stärker durch die Annahme des internationalen Intercepts verursacht ist. Die letztere Information stammt dabei aus den Mplus-Modifikationsindizes und verdeutlicht damit, wie beide Ansätze unterschiedliche wertvolle Information für die inhaltliche Interpretation beitragen können. Im Vortrag werden die Ergebnisse grafisch und in aggregierter Form über diverse Skalen präsentiert.



## F07

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Bernoullianum 32, H120

### Früherkennung abbruchgefährdeter Studierender und experimentelle Studien zur Wirksamkeit von Maßnahmen

**Johannes Berens, Simon Oster, Kerstin Prof. Dr. Schneider**

Wuppertaler Institut für bildungsökonomische Forschung (WIB), Bergische Universität Wuppertal, Deutschland;  
[berens.johannes@gmail.com](mailto:berens.johannes@gmail.com)

Studienabbrüche belasten die Hochschulen, die öffentlichen Haushalte und insbesondere die Studierenden. Aus diesem Grund unterhalten Hochschulen eine Vielzahl von Maßnahmen, die Studienzweifler unterstützen sollen (Tinto, 1975; Heublein & Wolter, 2011; Larsen, et al., 2013). Diese können jedoch nur dann wirksam sein, wenn sie von der Zielgruppe nachgefragt werden. Ist die Maßnahme den StudienzweiflerInnen unbekannt, schätzen sie den eigenen Bedarf falsch ein oder wird die Maßnahme nur von nicht abbruchgefährdeten Studierenden nachgefragt, kommt es zur Fehlallokation. Weiterhin ist die Effektivität der Maßnahmen weitestgehend unbekannt. Möglich ist, dass ein Studienabbruch lediglich verzögert wird. Beides, die nicht zielgerichtete Maßnahmenallokation und die unbekannte Effektivität der vorgehaltenen Maßnahmen, lassen Handlungsbedarf für die Hochschulsteuerung erkennen.

In dem zweistufigen Forschungsprojekt „FragSte“ wurde in einem ersten Schritt ein weitestgehend automatisiertes Frühwarnsystem zur Identifikation abbruchgefährdeter Studierender entwickelt (Arulampalam 2005; Sara et al., 2015). Anschließend wird die Maßnahmenallokation und Effizienz von Maßnahmen experimentell getestet.

Da das Frühwarnsystem ausschließlich Studierendendaten verwendet, die nach dem Hochschulstatistikgesetz erhoben werden, ist es an allen deutschen Hochschulen ohne Datenerhebungs- oder Adaptionsaufwand einsetzbar. Die Studierendendaten bestehen aus demografischen und voruniversitären Informationen sowie akademischen Leistungsdaten. Mittels ökonomischer Verfahren und Machine-Learning-Algorithmen werden basierend auf Daten ehemaliger Studierender (Absolventen und Studienabbrecher) die Merkmale gesucht, die Absolventen von Nichtabsolventen unterscheiden lassen. Die automatisierte Berücksichtigung hochschulindividueller Besonderheiten, die die Relevanz einzelner Prädiktoren verändern – wie besonders gute Hochschulzugangsnote an Eliteuniversitäten, heterogene Zusammensetzung der Studierenden oder kurze Studiendauern an Hochschulen mit hohen Studiengebühren – macht das Frühwarnsystem zu einem selbstjustierenden System.

In dem Frühwarnsystem kommen neuronale Netze, verschiedene Entscheidungsbäume sowie Regressionsanalysen zum Einsatz. Den unterschiedlichen „blind spots“ der Verfahren geschuldet, werden identische Studierende gegebenenfalls unterschiedlich beurteilt. Daher werden die eingesetzten Verfahren im Wege des „adaptive Boosting“ zu einem noch präziseren und robusteren Ergebnis zusammengeführt (Schapire & Freund, 2012).

Das Frühwarnsystem wurde an einer staatlichen Universität mit etwa 23.000 Studierende bei 90 Bachelor-Studiengänge und einer privaten Fachhochschule mit 6.500 Studierende, 30 Bachelorstudiengänge und durchschnittlichen Studiengebühren von 400,- Euro im Monat getestet. Untersucht wurden ehemalige Studierende von Bachelorstudiengängen der Jahre 2007 bis 2017. Diese wurden in Lernkohorten (Uni: 15.973 | FH: 7.169 Studierende) und eine Testkohorte (Uni: 2.127 | FH: 1.398 Studierende) aufgeteilt.

Zum Zeitpunkt der Immatrikulation können basierend auf demografischen und voruniversitären Informationen bereits 68,08% (Uni) und 67,17% (FH) der Studienabschlüsse korrekt prognostiziert werden. Unter zusätzlicher Verwendung der akademischen Leistungen steigt die Genauigkeit im ersten Semester deutlich auf 78,87% (Uni) und 84,49% (FH) und bis zum vierten Semester auf 89,71% (Uni) und 95,51% (FH). Die höhere Genauigkeit der privaten Fachhochschule ist den früheren Abbruchzeitpunkten geschuldet. So können im sechsten Semester 91,85% (Uni) der Studienergebnisse korrekt prognostiziert werden.

Insgesamt zeigt sich, dass ab dem ersten Semester die akademische Leistung als Prädiktor eine stetig zunehmende und die demografischen oder voruniversitären Informationen eine stetig abnehmende Aussagekraft haben. Selbst die Information des Studienganges, des Migrationshintergrundes oder die Hochschulzugangsnote ist im multivariaten Kontext entbehrlich, sobald Leistungsdaten verfügbar sind. Sehr gute Prädiktoren sind hingegen die Studiengeschwindigkeit (CP/Semester) und die Art der absolvierten Prüfungen.

In der anlaufenden zweiten Projektphase, werden die Maßnahmenallokation und die Interventionen an beiden Hochschulen experimentell getestet. Mit weiteren Universitäten werden derzeit Gespräche geführt. In einem ersten Schritt werden abbruchgefährdete Studierende mithilfe des Frühwarnsystems identifiziert und in eine Kontroll-, Warn- und Interventionsgruppe eingeteilt. Studierende der Warn- und Interventionsgruppe werden automatisiert informiert. Ob dieser Hinweis eine Wirkung auf anschließende Studienabbrüche hat, stellt die erste zu testende Intervention dar. Studierende der Interventionsgruppe werden zum Zwecke der Problemidentifikation und Auswahl eines bestehenden passgenauen Hilfsangebotes eingeladen und in der Maßnahmenteilnahme begleitet. Dadurch werden die Maßnahmenallokation und Maßnahmeneffektivität experimentell evaluiert. Da auch die Evaluation weitestgehend automatisiert ist, ist auch dieser Projektteil für andere Hochschulen nutzbar.

### Monetäre und nicht-monetäre Promotionserträge im Zeitverlauf

**Thorsten Euler, Trennt Fabian**

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung DZHW, Deutschland; [euler@dzhw.eu](mailto:euler@dzhw.eu)

Soziale Ungleichheit findet sich im besonderen Maße im Zusammenhang mit Bildung in Gesellschaften wieder. Nicht nur führt bestehende soziale Ungleichheit zur Reproduktion derselben im Bildungsbereich, auch Ungleichheit im Output des Bildungssystems produziert neue Ungleichheiten. Dies gilt auch für den Arbeitsmarkt (Sesselmeier & Blauermeier 1997).

Gemäß des Strukturfunktionalismus Parsons' lässt sich in der arbeitsteiligen Gesellschaft eine u.a. bildungsbezogene soziale Schichtung erklären (Davis & Moore 1973). Unter humankapitalistischen Gesichtspunkten stellt dabei die Verteilung von Bildungsabschlüssen einen Vergabemechanismus knapper begehrter Positionen dar. Signaltheoretisch gesehen (Spence 1973) besteht bei Hochqualifizierten zudem die geringere Wahrscheinlichkeit einer Fehlbesetzung.

Infolgedessen wird erwartet, dass mit steigendem Bildungsniveau als Teil des Humankapitals die Chance auf Erlangung beruflich attraktiver Positionen zunimmt (Abele 2003; Kühne 2009).

Die fortlaufende Bildungsexpansion in Deutschland seit den 1970er Jahren und das damit einhergehende Phänomen der Bildungsinflation (Boudon 1974) sind bei den tertiären Abschlüssen angekommen. Ein immer größerer Anteil junger Menschen erlangt die Hochschulzugangsberechtigung und einen anschließenden Hochschulabschluss. Dieser Prozess wird durch die Bologna-Reform noch weiter verstärkt. Eine attraktive berufliche Positionierung ist daher immer weniger durch einen Hochschulabschluss an sich gesichert (trotz höherer Chancen im Vergleich zu Nicht-Absolvent(inn)en). Entsprechend steigt die Notwendigkeit einer Promotion als wichtigeres Schichtungsmerkmal – auch jenseits des Wissenschaftssystems.

Hieraus ergibt sich die Frage, ob sich mit der steigenden Bedeutung der Promotion als Differenzierungsmerkmal auch die Differenzen zwischen Promovierten und Nicht-Promovierten hinsichtlich der Attraktivität ihrer Stellen erhöht haben. Attraktive Stellen für Promovierte bieten dabei höhere monetäre, aber auch nicht-monetäre Erträge im Vergleich zu Stellen von Nicht-Promovierten. Diese Erträge werden in der Regel über das Berufseinkommen, die Adäquanz der Beschäftigung, die berufliche Position (insb. Führungspositionen) oder die berufliche Zufriedenheit hinsichtlich spezifischer Tätigkeitsaspekte erfasst.

Frühere Studien zeigten einen monetären Vorteil einer Promotion (Heineck & Matthes 2012; Fabian 2013; Flöther 2015), eine höhere Zufriedenheit im Beruf bei Promovierten (Falk & Küpper 2013) und eine höhere Chance auf Besetzung einer Führungsposition bei diesen (Fabian 2013). Allerdings variieren die Effekte stark zwischen den einzelnen Fachrichtungen und hinsichtlich des Geschlechts (Heineck & Matthes 2012). Alle diese Studien basieren allerdings auf Beobachtungen eines einzelnen Prüfungsjahrgangs und können entsprechend weder Kohorteneffekte aufzeigen, noch eine zeitliche Entwicklung im Rahmen der Bildungsinflation.

Der geplante Vortrag ist Teil der derzeit noch laufenden Auswertungen und Berichtserstellung im Rahmen der dritten Erhebung des Prüfungsjahrgangs 2005 des DZHW-Absolventenpanels für den Projektträger (Bundesministerium für Bildung und Forschung). Er versucht diese bestehende Forschungslücke mittels der Betrachtung monetärer und nicht-monetärer Promotionserträge mehrerer Prüfungskohorten zu schließen.

Als Grundlage dienen die Daten des DZHW(ehemals HIS–HF)-Absolventenpanels. In diesem werden seit 1989 alle vier Jahre repräsentative Stichproben der Absolvent(inn)en eines Prüfungsjahres an deutschen Hochschulen im Abstand von jeweils circa einem, fünf und zehn Jahren nach Studienabschluss befragt. Sie ermöglichen einen fundierten Einblick in den Übergang von Absolvent(inn)en in berufliche und wissenschaftliche Karrieren. Die Analyse stützt sich auf Absolvent(inn)en traditioneller Universitätsabschlüsse in den jeweils dritten Befragungen der Prüfungsjahrgänge 1997 (n=4.033), 2001 (n=3.235) und 2005 (n=2.552), jeweils circa zehn Jahre nach Studienabschluss.

Die Analysen erfolgen mittels logistischer Regressionsmodelle unter Einbezug demographischer, vorhochschulischer, studien- und berufsbezogener Kontrollvariablen. Variationen hinsichtlich der spezifischen Fachrichtung und Geschlechtereffekte werden separat dargestellt.

Erste einfache Auswertungen deuten auf eine Zunahme der monetären Erträge im Zeitlauf hin. So steigt die rein deskriptive Differenz unter Absolvent(inn)en traditioneller Universitätsabschlüsse zwischen Promovierten und Nicht-Promovierten von knapp unter 13.000 € für den Jahrgang 1997 auf etwa 23.000 € für den Jahrgang 2005 an. Unter Kontrolle der Fachrichtungen und Geschlechtereffekte sinken die Höhen der Differenzen, bleiben jedoch in ihrem Trend einer Zunahme des monetären Vorteils einer Promotion stabil. Im Bereich nicht-monetärer Erträge einer Promotion ist die Entwicklung weniger eindeutig. Zielsetzung des Beitrags ist es, mittels weiterführender Modellierung die eindeutigen Veränderung in Promotionserträgen herauszuarbeiten und zu präsentieren.

## **Wirtschaftsbürgerliche Kompetenz von Gymnasialschülerinnen und -schülern in der Schweiz und in Deutschland**

**Nicole Ackermann<sup>1</sup>, Christin Siegfried<sup>2</sup>**

<sup>1</sup>Universität Zürich, Schweiz; <sup>2</sup>Goethe-Universität Frankfurt, Deutschland; [nicole.ackermann@ife.uzh.ch](mailto:nicole.ackermann@ife.uzh.ch), [Siegfried@em.uni-frankfurt.de](mailto:Siegfried@em.uni-frankfurt.de)

### Theoretischer Hintergrund

In einem demokratischen Rechtsstaat mit einer liberalen Gesellschafts- und Wirtschaftsordnung sind die Staatsbürger in private, berufliche und politische Meinungsbildungs- und Entscheidungsprozesse eingebunden. Sie werden beispielsweise mittels öffentlicher Debatten und Abstimmungen regelmässig aufgefordert, sich zu aktuellen wirtschafts- und gesellschaftspolitischen Problemstellungen bzw. zu deren Lösungsvorschlägen zu äussern. Charakteristisch für diese Problemstellungen ist, dass sie abstrakt, komplex und multiperspektiv sind (Dubs, 2001; Eberle, 2006). Um diese sozio-ökonomischen Lebenssituationen erfolgreich und verantwortungsvoll zu bewältigen, also „professionell“ zu handeln, benötigen junge Menschen wirtschaftsbürgerliche Kompetenz (WBK). Dies umfasst die Fähigkeit, Fertigkeit und Bereitschaft des Individuums wirtschaftliche und politische Grundlagen zu verstehen sowie aktuelle und komplexe sozio-ökonomische Problemstellungen zu bearbeiten (vgl. Weinert, 2001, Klieme, 2004, 2007; Eberle, 2015).

Die Förderung der WBK geschieht an Schweizer Gymnasien vorrangig in den Fächern „Wirtschaft & Recht“ (W&R) und „Geschichte & Staatskunde“, an Deutschen Gymnasien in „Wirtschaftswissenschaften“ (Wiwi) und „Politik & Wirtschaft“ (PoWi). In der Schweiz muss das Fach W&R von allen Gymnasialschülern als obligatorische Einführung besucht werden und kann als Schwerpunktkurs gewählt werden (MAR Art. 9, 1995). In Deutschland sind ökonomische Inhalte – wenn auch sehr divers und teilweise stark unterrepräsentiert- mittlerweile in allen Bundesländern implementiert (Weber, 2007). In Hessen muss z.B. das Fach PoWi von allen Gymnasialschülern in der Sekundarstufe I und II obligatorisch besucht werden (Hessisches Kultusministerium, 2011) und kann in der Sekundarstufe II durch den Leistungskurs Wiwi ergänzt werden (OAVO § 13, 2009). Die Lehrpläne (Lerninhalte, Studentafeln) sind jeweils unterschiedlich für den obligatorischen Einführungskurs und für den wählbaren Schwerpunktkurs.

Bisherige empirische Untersuchungen, die sich auf eine curricular-orientierte ökonomische Kompetenz beziehen, zeigen bei Schweizer Maturanden einen signifikanten Unterschied in der Testleistung zugunsten der Testpersonen mit Schwerpunkt W&R (Schumann & Eberle, 2014). Eine kürzlich erfolgte Untersuchung zur wirtschaftsbürgerlichen Kompetenz von kaufmännischen Lernenden in der Schweiz und in Deutschland zeigt eine nicht-signifikant höhere Testleistung der Schweizer Lernenden gegenüber den Deutschen; die Testleistung ist signifikant höher, wenn nur Testpersonen ohne Matura bzw. Abitur berücksichtigt werden (Schumann, Kaufmann, Eberle, Jüttler, & Ackermann, 2017). Inwiefern sich länderbezogene Unterschiede in der WBK-Testleistung auch für Gymnasialschüler zeigen, bedarf einer empirischen Überprüfung.

### Fragestellung

Das Ziel dieses Einzelbeitrages ist es, Ausprägungen und Unterschiede in der wirtschaftsbürgerlichen Kompetenz von Gymnasialschülern in der Schweiz und in Deutschland zu analysieren. Dabei sollen folgende Fragestellungen beantwortet werden: (FF1) Ländervergleich: Wie unterscheidet sich die Testleistung von Schülern der beiden Länder? (FF2) Profilvergleich: Wie unterscheidet sich die Testleistung von Schülern im Schwerpunktkurs Wirtschaft von denjenigen in einem anderen Schwerpunktkurs?

#### Methode

Zur Messung der kognitiven Facette der WBK wurden zwei Themen (Altersvorsorge und Staatsverschuldung) des revidierten WBK-Tests (Eberle, Schumann, Kaufmann, Jüttler & Ackermann, 2016; Ackermann, in Arbeit) sowie ein Kontextfragebogen zur Messung sozio-demographischer Merkmale der Testpersonen eingesetzt. Die Datenerhebungen fanden im Winter/Frühling 2017 statt. Die Schweizer Stichprobe umfasst 375 Schüler (W&R 42 %, Nicht-W&R 58 %) von gymnasialen Maturitätsschulen des Kantons St. Gallen. Die Deutsche Stichprobe umfasst 130 Schüler (Wiwi 66 %, Nicht-Wiwi 34 %) des Gymnasiums im Bundesland Hessen. Für die Datenanalyse werden T-Tests für unabhängige Stichproben gerechnet.

#### Ergebnisse

Beim Vergleich der Testleistung, welche maximal 30 Punkte erreichen konnte, zeigt sich beim Vergleich beider Länder ein signifikanter Unterschied zugunsten der Schweizer Gymnasialschüler ( $M=20.04$   $SD=3.59$ ) gegenüber den Deutschen ( $M=15.54$ ,  $SD=4.14$ ) ( $t(503)=-11.024$ ,  $p<.001$ ,  $d=-1.25$ ). In der Schweizer Stichprobe haben Schüler der Gruppe W&R ( $M=21.04$   $SD=3.28$ ) signifikant höhere Testwerte als diejenigen der Gruppe Nicht-W&R ( $M=19.29$   $SD=3.64$ ) ( $t(373)=4.820$ ,  $p<.001$ ,  $d=0.503$ ) erreicht. In der Deutschen Stichprobe haben Schüler der Gruppe Wiwi ( $M=14.7$ ,  $SD=4.17$ ) signifikant geringere Testwerte als diejenigen der Gruppe Nicht-Wiwi ( $M=17.18$   $SD=3.59$ ) ( $t(128)=-3.360$ ,  $p=.001$ ,  $d=-0.622$ ).

## Monetäre Erträge non-formaler beruflicher Weiterbildung – Revisited

**Benjamin Schimke**

Bergische Universität Wuppertal, Deutschland; [schimke@uni-wuppertal.de](mailto:schimke@uni-wuppertal.de)

Das wissenschaftliche Erkenntnisinteresse an beruflicher Weiterbildung (kurz: bW) ist nicht zuletzt wegen dem paradigmatisch anmutenden Begriff des lebenslangen Lernens groß. Trotz des anhaltend hohen Interesses und der stetig wachsenden Zahl an Publikationen ist der Erkenntnisstand begrenzt, da mit unterschiedlichen Weiterbildungskonzepten, Datensätzen und Methoden, teils stark voneinander abweichende Analyseergebnisse produziert werden. Neben den Fragen über die sozial ungleiche Verteilung der Teilnahme an bW und den subjektiv bewerteten Nutzen der Teilnehmer/innen sind insbesondere die (objektiven) Erträge bW von zentralem wissenschaftlichem Interesse.

Ziel des Beitrags ist es, die monetäre Rendite bW zu quantifizieren. Dabei wird der Versuch unternommen, die in diesem Forschungsfeld bestehende Erkenntnislage um eine such- und matchingtheoretische Perspektive zu erweitern. Es wird zu zeigen sein, dass erstens die bislang in vielen Forschungsarbeiten fehlende Berücksichtigung unbeobachteter job-match-spezifischer Heterogenität zu einer verzerrungsanfälligen Schätzung von Weiterbildungserträgen geführt hat. Zweitens die statistische Kontrolle dieser Heterogenitätsquelle über eine Instrumentalvariablen-schätzung einen Erkenntnisgewinn zur ökonomischen Weiterbildungsforschung liefert.

Die Basis des theoretischen Rahmens bildet die Humankapitaltheorie nach Becker (1964) und die daran anknüpfende Formulierung einer Einkommensgleichung nach Mincer (1974), sowie Ansätze aus der Such- und Matchingtheorie nach Jovanovic (1979a, b). Die Erweiterung der Mincer'schen Einkommenschätzfunktion spielt in zahlreichen Forschungsarbeiten zu den Erträgen bW eine elementare Rolle (bspw. Ehlert 2017; Büchel & Pannenberg 2004). Die übliche Herangehensweise in diesem Zusammenhang ist das Fixed-Effects-Modell (kurz: FE). Sowohl Ehlert (2017) als auch Büchel und Pannenberg (2004), sowie Wolter und Schiener (2008) wenden diese Modellierungstechnik auf Daten des NEPS, des SOEP und den Mikrozensus an. Sie alle ermitteln niedrige, wenn auch signifikante Weiterbildungserträgen mindestens für spezifische Teilgruppen am Arbeitsmarkt und nutzen die elementaren Vorteile von FE-Regressionen gegenüber anderen Modellierungsstrategien. Diese sind u.a. die Konsistenzigenschaften; selbst wenn klassische Modellannahmen (wie z.B. die Exogenitätsannahme) verletzt werden (Andreß et al. 2013; Brüderl 2010). Bei Vorliegen unbeobachteter zeitvarianter Heterogenität liefert aber auch ein FE-Schätzer keine konsistenten Schätzergebnisse.

Dieser Fall scheint bei der bisherigen Forschung zu monetären Weiterbildungserträgen gegeben zu sein. Der Renditeschätzer bW sollte mit der job-match-spezifischen Fehlerkomponente zusammenhängen (vgl. Altonji & Shakotko 1987; Altonji & Williams 2005). In diesen Teil des Fehlerterms geht die unbeobachtete Qualität der Arbeitnehmer-Arbeitgeber-Passung ein, die wiederum maßgeblich für das Weiterbildungsverhalten von abhängig Beschäftigten sein dürfte. Insbesondere Erwerbstätige mit verhältnismäßig schlechten Job-Matches werden am Arbeitsmarkt, auf der Suche nach einer besseren Passung mobiler sein (vgl. Jovanovic 1979a, b; Mincer & Jovanovic 1981) und dadurch eine geringere – insbesondere arbeitgeberinduzierte – Weiterbildungshäufigkeit aufweisen. Diese Zusammenhangsvermutung legt den Schluss nahe, dass die Weiterbildungserträge bislang überschätzt wurden.

Das Hauptanliegen des Vortrags ist es daher, die aus der Such- und Matchingtheorie abgeleitete Implikation für die Datenanalyse mittels einer Instrumentalvariablen-schätzung (kurz: IV) zu berücksichtigen. Dabei dienen die innerhalb eines Job-Matches zentrierten Weiterbildungshäufigkeiten und -intensitäten als Instrumentalvariablen für die beobachteten Häufigkeiten und Intensitäten. Dadurch wird die notwendige Orthogonalität zwischen Regressoren der bW und den individualspezifischen und job-match-spezifischen Fehlertermkomponenten erreicht (vgl. Altonji & Shakotko 1987).

Die Analysen werden zurzeit auf Basis der Erwachsenenkohorte des NEPS (Version 7.0.0) mithilfe der von 2009 bis 2015 erhobenen Panelwellen 1 bis 6 durchgeführt (Blossfeld et al. 2011). Die bereinigte Stichprobe umfasst 10.718 Erwerbstätige, welche zu 35.550 Erhebungszeitpunkten einen Bruttostundenlohn berichten ( $M=17,94$ ,  $SD=8,99$ ) und im gesamten Analysezeitraum an durchschnittlich 3,16 Maßnahmen der bW ( $SD=5,22$ ) mit einer mittleren Kursdauer von 54,62 Stunden ( $SD=120,37$ ) teilnehmen. Fehlende Werte in für die Untersuchung zentralen unabhängigen Variablen werden multipel imputiert. Erste multivariate Analysen bestätigen die Vermutung eines, durch unbeobachtete job-match-spezifische Heterogenität, verzerrten FE-Schätzers. Der Hausman-Spezifikations-Test (Hausman & Taylor 1981) zum Vergleich von FE- und IV-Schätzung stützt die theoretischen Überlegungen zugunsten des IV-Renditeschätzers ebenfalls. Zusätzlich sollen in einem weiteren Analyseschritt segmentierungstheoretische Überlegungen integriert werden, um die Ergebnisse für unterschiedliche Beschäftigtengruppen des deutschen Arbeitsmarktes zu untersuchen (Sengenberger 1987).

## F08

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Bernoullianum 32, H103

### **Transformationale Führung in Schulen und Lehrpersonen-Selbstwirksamkeit: eine Überprüfung des dual-effects-Modells auf individueller und kollektiver Ebene**

**Regula Windlinger<sup>1</sup>, Julia Warwas<sup>2</sup>**

<sup>1</sup>PHBern, Schweiz; <sup>2</sup>Universität Bamberg, Deutschland; [regula.windlinger@phbern.ch](mailto:regula.windlinger@phbern.ch)

Hintergrund:

Untersuchungen zur Wirksamkeit schulischen Leitungshandelns rücken transformationale Führungsaktivitäten in den Fokus (vgl. Harazd & van Ophuysen, 2011). Meta-Analysen dokumentieren übereinstimmend positive Zusammenhänge der transformationalen Führung (TF) mit Merkmalen der Einzelschule, wie bspw. geteilten Schulentwicklungszielen oder innovationsoffener Organisationskultur, sowie mit Einstellungen, Selbstwahrnehmungen und Verhaltensweisen der Lehrkräfte, die ihrerseits förderlich für eine produktive Schulentwicklung sind und indirekte Zusammenhänge der TF mit Schülerleistungen begründen (Hendriks & Scheerens, 2013; Leithwood & Sun, 2012). Die Stärkung der Selbstwirksamkeitserwartungen von Lehrkräften stellt in dieser Betrachtung eine wichtige Zielgröße transformationaler Führung dar, weil dieses Führungskonzept per definitionem eine Veränderung von Einstellungen und Selbstwahrnehmungen der Mitarbeiter/innen anstrebt, um deren Engagement in organisationalen Entwicklungsprozessen zu erhöhen (z.B. Hoffmann et al., 2011).

Auf weiteren Forschungsbedarf verweist aber die hohe Schwankungsbreite ermittelter Zusammenhänge zwischen TF und diversen Zielvariablen, welche Robinson et al. (2008) auf operationale Vielfalt und konzeptionelle Unschärfe zurückführen. TF wird meist als Globalkonstrukt aufgefasst; seine statistischen Effekte werden wahlweise auf aggregierter Schulebene oder auf Ebene der einzelnen Lehrkräfte spezifiziert. Unberücksichtigt blieben bislang Diskussionen um sog. duale Effekte und damit eine konzeptionelle wie auch konsequent empirische Differenzierung transformationaler Führungsaktivitäten, die sich individuenzentriert an die einzelne Lehrkraft richten (Individuelle Förderung; Intellektuelle Stimulierung), und solchen, die sich gruppenzentriert an das Kollegium wenden (Idealisierter Einfluss, Inspirierende Motivierung) (z.B. Kark & Shamir, 2002; Schriesheim, et al., 2009). Für eine Überprüfung dualer Effekte im Schulbereich bieten sich Selbstwirksamkeitserwartungen der Lehrenden an, da sich sowohl in der Schulentwicklungsforschung (Goddard et al., 2000) als auch in der Lehrergesundheitsforschung (Skaalvik & Skaalvik, 2007) individuelle und kollektive Wirksamkeitserwartungen als analytisch trennbare Konstrukte bewährt haben.

Empirisch unterbelichtet ist weiterhin die Frage, ob TF unter allen Umständen gleichermassen wirksam ist. Obwohl die Kontextabhängigkeit schulischer Führung wiederkehrend betont wird (z.B. Brauckmann, 2012; Moos, Krejsler & Kofod, 2008), wird selten überprüft, inwieweit variierende Zusammenhänge zwischen der TF und Zielgrößen durch schulspezifische Bedingungsmerkmale erklärt werden können. Die vorliegende Studie berücksichtigt als Moderatoren zum einen die Schulgröße, die als Indikator der Nähe bzw. Distanz zwischen der Führungskraft und den einzelnen Kollegiumsmitgliedern dient (vgl. Chun et al., 2009), zum anderen einen standortspezifischen Schulsozialindex, mit dessen Hilfe Schulen in schwierigen Lagen (vgl. Racherbäumer et al., 2013) identifiziert werden können.

Fragestellungen:

1. Lassen sich individuen- und gruppenbezogene Dimensionen transformationaler Führungsaktivitäten auch empirisch unterscheiden?
2. Bestehen Zusammenhänge zwischen gruppenbezogener TF und kollektiver Selbstwirksamkeit sowie zwischen individuenbezogener TF und individueller Selbstwirksamkeit?
3. Unterscheiden sich diese Effekte je nach Kontext der Schulen (Größe, Schulsozialindex)?

Methode:

Die Datenerhebung (schriftliche Befragung) erfolgte im Rahmen einer vom Schweizerischen Nationalfonds geförderten Studie (<http://p3.snf.ch/project-136877>) bei 1702 Lehrpersonen in 118 Schulen im Kanton Bern, Schweiz.

Erfasst wurden transformationale Führung (MLQ 5x short, Felfe, 2006), individuelle Selbstwirksamkeitserwartungen (Tschannen-Moran & Woolfolk Hoy, 2001) und kollektive Selbstwirksamkeitserwartungen (Schwarzer & Jerusalem, 1999). Reliabilitätsanalysen ergaben durchweg zufriedenstellende bis gute Skalen- und Itemkennwerte. Kontextvariablen wurden amtlichen Statistiken als schulgenaue Informationen entnommen.

Zur Auswertung wurden Mehrebenen-Strukturgleichungsanalysen in Mplus (Muthén & Muthén, 1998-2012) berechnet. Durch Anwendung des „doubly latent“-Verfahrens wurden dabei sowohl Mess- als auch Stichprobenfehler kontrolliert (Marsh et al., 2012).

Ergebnisse:

Der Vergleich verschiedener Messmodelle (Zwei-Ebenen-CFAs) weist das Modell mit der theoriekonformen Unterscheidung individuen- und gruppenbezogener Dimensionen transformationaler Führung als vorzugswürdig gegenüber einem Globalfaktor-Modell aus. Auf der Schulebene ist die gruppenbezogene TF erwartungsgemäß positiv mit der kollektiven Selbstwirksamkeit assoziiert ( $\beta = .503^{**}$ ); auf der Lehrendenebene hängt die individuenbezogene TF systematisch mit individuellen Selbstwirksamkeitserwartungen zusammen ( $\beta = .162^{**}$ ).

Moderatormodelle zeigen, dass der Zusammenhang zwischen gruppenbezogener TF und kollektiver Selbstwirksamkeit in kleineren Schulen stärker ausgeprägt, also abhängig von der Schulgröße, ist. Moderierende Effekte des Schulsozialindex lassen sich nicht zufallskritisch absichern, wohl aber direkte Effekte auf die kollektive Selbstwirksamkeit.

### **Offen für Bildungsinnovationen? – Einflussfaktoren auf die Akzeptanz von Reformmaßnahmen aus der Perspektive von Schulleiterinnen und Schulleitern**

**Eunji Lee, Susanne Böse, Marko Neumann, Kai Maaz**

Deutsches Institut für Internationale Pädagogische Forschung, Deutschland; [lee@dipf.de](mailto:lee@dipf.de)

Theoretischer Hintergrund

Mit Blick auf die gegenwärtige Situation des deutschen Bildungssystems lässt sich derzeit eine Vielzahl von Veränderungsprozessen ausmachen. Diese Veränderungen beziehen sich oftmals auf die Implementation von umfassenden

Reformvorhaben und die Einführung von Bildungsinnovationen mit den Zielsetzungen der Verbesserung der Qualität von Schule und der Reduktion sozialer Ungleichheiten (Hasselhorn, Köller, Maaz & Zimmer, 2014). Dabei sind in jüngster Zeit zunehmend auch Schulen in schwieriger Lage in den Fokus gerückt. Das Land Berlin hat im Jahr 2014 das sogenannte Bonus-Programm initiiert (Senatsverwaltung für Bildung, Jugend und Wissenschaft, 2016), im Rahmen dessen Schulen in schwieriger Lage zusätzliche finanzielle Mittel (jährlich bis zu 100.000 EUR pro Schule) erhalten, über die sie weitgehend frei verfügen können. Die Mittelbewilligung ist hierbei an das Schließen von Zielvereinbarungen mit der zuständigen Schulaufsicht gekoppelt.

Die Implementation von Reformen bzw. Innovationen geht mit zahlreichen Herausforderungen einher (Rogers, 2003), die unter anderem diejenigen Akteure betreffen, deren Aufgabe es ist, diese Innovationen umzusetzen (Gräsel, 2010; Klein & Knight, 2005). Dabei wird betont, dass es für den Erfolg von Reformen wichtig ist, dass die beteiligten Akteure die betreffende Innovation als vorteilhaft bewerten – ihr also mit einer grundlegenden Akzeptanz gegenüberstehen und die Reform mittragen (z.B. Euler & Sloane, 1998; Gräsel & Parchmann, 2004; Hasselhorn et al., 2014; Rogers, 2003). Im Kontext von schulischen Innovationen betrifft dies insbesondere Schulleiterinnen und Schulleiter (Altricher & Wiesinger, 2005), da sie letztendlich verantwortlich für die Umsetzung an ihren Schulen sind. Allerdings sind die empirischen Befunde bezüglich der Akzeptanz von Innovationen und ihrer möglichen Bedingungsfaktoren im Schulkontext noch rar (Böse, Neumann, Becker, Maaz & Baumert, eingereicht).

#### Fragestellung

Im vorliegenden Beitrag wird daher die Akzeptanz des Bonus-Programms als Voraussetzung für eine erfolgreiche Implementation näher beleuchtet. Auf Grundlage der Daten der BONUS-Studie (Böse, Neumann, Gesswein & Maaz, 2017), der wissenschaftlichen Begleituntersuchung des Bonus-Programms, soll der Fokus auf die Schulleiterinnen und Schulleiter, die als Hauptakteure im Implementationsprozess an ihren Schulen agieren, gelegt werden. Im Zentrum steht die Frage, welche Faktoren auf der Individualebene der Schulleiterinnen und Schulleiter sowie auf der Ebene der Schulen als Determinanten für die Akzeptanz des Programms identifiziert werden können.

#### Methode

Im Rahmen der BONUS-Studie wurden 191 Schulleiterinnen und Schulleiter kurz nach der Einführung des Programms befragt. Die Akzeptanz bezüglich des Bonus-Programms wurde durch Fragen, die zum einen die Bewertung des Programms (Beispielitem „Insgesamt halte ich das Bonus-Programm für sehr sinnvoll“) und zum anderen die Erwartungen an das Programm (Beispielitem: „Das Bonus-Programm wird dazu beitragen, Bildungsbenachteiligungen von Schülerinnen und Schülern auszugleichen“) erfragten, operationalisiert. Als potentielle Einflussfaktoren wurden das Kohärenzgefühl hinsichtlich des Programms (wird das Programm als verständlich und bewältigbar wahrgenommen?), Kenntnisse und Vorerfahrungen in eigenverantwortlicher Mittelbewirtschaftung, die Selbsteinschätzung zur effektiven Verwendung der Bonus-Mittel, die Nutzung und Bewertung von Unterstützungs- und Vorbereitungsangeboten sowie verschiedene personenbezogene Merkmale (z.B. Geschlecht, Alter, Berufserfahrung) und Schulmerkmale (z.B. soziale Zusammensetzung der Schülerschaft, Innovationsbereitschaft im Kollegium) betrachtet. Neben der bivariaten Analyse der Zusammenhänge erfolgte auch eine multivariate Vorhersage der Akzeptanzeinschätzungen durch die verschiedenen Einflussfaktoren. Die Analysen erfolgten mithilfe des Statistikprogramms R. Fehlende Werte in den Analysevariablen wurden über Full-Information-Maximum-Likelihood-Schätzungen berücksichtigt.

#### Ergebnisse

Die bivariate Analysen zeigten deutliche positive Zusammenhänge zwischen den Akzeptanzeinschätzungen und dem Kohärenzgefühl. Weitere positive Zusammenhänge konnten unter anderem mit der Einschätzung zur effektiven Mittelverwendung und der Bewertung der Vorbereitungsangebote festgestellt werden. Für die Erwartungen an das Programm fanden sich zudem positive Zusammenhänge mit der Innovationsbereitschaft des Kollegiums und den gemachten Vorerfahrungen in der selbstständigen Mittelbewirtschaftung. In den multivariaten Analysen erwies sich insbesondere das Kohärenzgefühl als prädiktiv für die Akzeptanzeinschätzungen der Schulleiterinnen und Schulleiter.

Für zukünftige Analysen stellt sich die Frage, inwieweit sich die Akzeptanz des Programms auf dessen Umsetzung sowie das Erreichen der an den einzelnen Schulen angestrebten Zielsetzungen und Veränderungen auswirkt.

## **Einfluss von Projektleitung und Vertrauen der Projektbeteiligten auf die Kooperation in Bildungslandschaften**

**Marius Schwander, Giang Pham, Stephan Gerhard Huber**

PH Zug, Schweiz; [marius.schwander@phzg.ch](mailto:marius.schwander@phzg.ch)

#### Theoretischer Hintergrund

Seit 2014 läuft in der Schweiz eine Initiative mit dem Ziel, die Vernetzung schulischer und ausserschulischer Akteure in Bildungslandschaften zu etablieren zur Verbesserung der Chancengerechtigkeit in der Bildung. Anhand längsschnittlicher Evaluationsdaten dieses Projekts werden hypothetische fördernde Bedingungen der Kooperation zwischen Netzwerkakteuren und die Rolle der Projektleitung untersucht. Gemäss Gronn (2002) sind die Gelingensbedingungen unter anderem ausreichende Zeitfenster, Vereinbarungen über pädagogische Vorstellungen und Strategien, gegenseitiges Vertrauen sowie offene Kommunikation. Im vorliegenden Beitrag liegt der Fokus auf das gegenseitige Vertrauen, da dieses neben einem gemeinsame Zielverständnis für die Kooperation in Netzwerken als unabdingbar angesehen wird, da die Zusammenarbeit in der Regel ohne vertragliche Regelungen geschieht (Hajer & Versteeg, 2005).

Vertrauensgemeinschaften sind durch geteilte Meinungen und Haltungen charakterisiert (Seligman, 2012), was ein gemeinsames Zielverständnis begünstigen dürfte. Vertrauen gilt weiter als Schlüsselement für eine funktionierende Kommunikation (Hoy, Smith, & Sweetland, 2002). Entsprechend hat Vertrauen auch einen Einfluss auf das Teilen von Informationen (vgl. Butler, 1999). Fehlt hingegen das Vertrauen, ist es unwahrscheinlich, dass Lehrpersonen die eigene professionelle Praxis zugunsten gemeinsamer Durchführung oder Reflexion riskieren (Tschannen-Moran, 2014).

Die Wirksamkeitsforschung im Bildungskontext zeigt, dass die Schulleitung genau hier eine besondere Rolle spielt, indem sie (unter anderem) die Akzeptanz gemeinsamer Ziele fördert, eine Schulkultur stärkt und die Zusammenarbeit etabliert (Leithwood & Riehl, 2003). Die Schulleitung kann das Vertrauen des Kollegiums gewinnen, indem sie Strukturen für die Entscheidungsfindung schafft und Diskretion hinsichtlich Entscheidungen, die Rückschlüsse auf die Expertise der beteiligten Personen zulassen, gewährleistet (Tschannen-Moran, 2014). Das Vertrauen in die Schulleitung hängt zudem stark davon ab, wie ihre Kompetenz eingeschätzt wird (Tschannen-Moran, 2014).

#### Fragestellung

Die Literatur zu Vertrauen im Bildungskontext bezieht sich überwiegend auf Schulleitungen und Lehrpersonen und wird hier auf Bildungsnetzwerke übertragen. Entsprechend gehen wir in diesem Beitrag folgender Fragestellung nach: Welchen Einfluss hat Vertrauen auf die Kooperation der Projektbeteiligten und wird dieses durch die Projektleitung beeinflusst? Als hypothetische Strukturmodelle werden die genannten theoretischen Kausalzusammenhänge berücksichtigt.

#### Methode

Die verwendeten Daten beruhen auf Akteure-Fragebogendaten von 275 Personen in 2015 (T2) und 331 Personen in 2016 (T3). Die Handlungskoordination der Projektleitung wurde durch die Projektbeteiligten eingeschätzt und umfasst Fragen zur inhaltlichen und organisatorischen Kompetenz, zur Unterstützung bei Planung und Durchführung oder zur Beteiligung und Einflussnahme bei Entscheidungen. Zur Erhebung der Kooperationsbedingungen wurden die Skalen «Zeitstruktur» und «Gemeinsames Ziel» von Fussangel (2008) sowie die Skalen «Vertrauen» und «Kommunikationsklima» aus Schulen im Team verwendet. Die Skala «Informationsweitergabe» wurde berechnet aus den adaptierten Items der Pädagogischen Entwicklungsbilanzen und stellt eine einfache Form der Zusammenarbeit dar. Die verwendete Kooperationsart «Kokonstruktion und Reflexion» gilt als höchste Stufe von Kooperation (Fussangel, Dizinger, Böhm-Kasper, & Gräsel, 2010).

Zur Bestätigung der theoretischen Strukturmodelle wird explorativ vorgegangen. Zunächst werden die Wirkrichtungen zwischen jeweils zwei Bedingungsfaktoren mittels Cross-Lagged-Modellen anhand Daten zweier Messzeitpunkte bestimmt. Darauf basierend werden die resultierenden Ergebnisse in einem Pfadmodell zusammengefasst und die Projektleitung als unabhängige sowie die «Kokonstruktion und Reflexion» als abhängige Variable ins Modell ergänzt. Die resultierenden Ergebnisse werden inhaltlich interpretiert und mit der vorliegenden Theorie verglichen.

#### Ergebnisse

Ergebnisse der Cross-Lagged Modelle weisen auf die Wirkung einer Zeitstruktur auf ein gemeinsames Ziel und auf das Vertrauen der Projektbeteiligten. Widererwartet begünstigt weder Vertrauen ein gemeinsames Ziel, noch führt ein gemeinsames Ziel zu höherem Vertrauen. Beide Faktoren korrelieren aber zu beiden Messzeitpunkten hoch ( $.54 < r < .61$ ), was zur Charakterisierung von Vertrauensgemeinschaften nach Seligman (2012) passt. Übereinstimmend mit Butler (1999) führt Vertrauen zur Informationsweitergabe.

Im Pfadmodell mit allen verfügbaren Variablen und Wirkrichtungen bleibt der Effekt von Vertrauen auf Informationsweitergabe weiterhin signifikant ( $R^2=49\%$ ). Übereinstimmend mit Tschannen-Moran (2014) wird «Kokonstruktion und Reflexion» ebenfalls durch Vertrauen vorhergesagt ( $R^2=7\%$ ). Die Ergebnisse bestätigen somit die zentrale Bedeutung von Vertrauen, welches wiederum direkt und indirekt durch die Projektleitung beeinflusst wird ( $R^2=33\%$ ).

## **School Turnaround: School Leadership and Coordinated System Efforts as a Key to Radical School Development?**

**Stephan Gerhard Huber, Guri Skedsmo**

Pädagogische Hochschule Zug, Schweiz; [stephan.huber@phzg.ch](mailto:stephan.huber@phzg.ch), [guri.skedsmo@phzg.ch](mailto:guri.skedsmo@phzg.ch)

#### Theoretical Perspectives

First, research on school effectiveness, which since the 1970's has focused on the influence of various school related factors on student academic outcomes, and increasing looking at "ineffective" or "failing" school contexts (e.g. Baumertetal2006; Murphy&Meyers2008; Author2012a,b).

Second, research on school improvement with a particular focus on "School Turnaround", adapting initiatives to the needs of the individual schools (e.g. Reynoldsetal1996). Models include school closure and staff retention, others concentrate on professional development of staff or even school fusion between "failing" and "successful" schools (Author2012c; Racherbäumeretal2013).

#### Objectives

This paper focuses on identifying success factors of school turnaround but also factors hindering school improvement.

#### Methods

This paper explores school development processes in 20 schools in difficult circumstances in a large city in Germany. Ten schools took part in interventions initiated as by the ministry, the school authority and a foundation. The others form the control group (located in the same community with similar demographic and quality characteristics).

Data was gathered over four years using a mixed-methods longitudinal design which includes: i. document analysis (eg. school inspection reports, school development planning documents, the logs of the coaches for the schools), ii. three surveys among school leaders and the whole staff: 2014RR52%, fall2015RR38%, and spring 017RR44%, iii. each time 80-90 semi-structured interviews with principals, teachers, social workers and the school governing system as well as actors involved in interventions (e.g. school development coaches). Finally, further statistical information from the ministry about the schools (e.g. illness of staff, the number of lesson drop out, student absence and academic completion rates).and VERA 3/8 data is analysed. SEM is used as analytic technique to test out latent variables and investigate relationships between latent variables. Growth models are applied to explore changes over time where different types of intervention, input, throughput and output variables.

#### Results

Findings from quantitative analysis indicate in some of the schools changes with respect to cooperation, degree of consensus about strategies, job satisfaction and capacity to deal with work pressure.

During the three first project years, most of the school leaders have mainly concentrated on changing the school structure and communication processes with staff. Then they concentrated on improving teaching and learning.

In general, strategic leadership is a main key for various improvements. In all schools, except one, there has been a change of principals, either right before the start of the project or during the second and third project year, which seems to have a major impact on the school culture and work processes in the schools resulting in decreases of lesson drop out and student absence. In several cases, the school leaders have also managed to cooperate with key actors in the local community as well as media in order to change the school reputation. This has resulted in increased numbers of students having the school as their first choice. Only a few schools have managed to accomplish higher completion rates.

Findings show that it is highly problematic for schools to do a "turnaround" on their own. Cooperation with local authorities and coordinated system efforts are needed in order for the schools to improve in a sustainable way.

In comparison with countries such as the US and England "school turnaround" in the German context has limitations in terms of strong regulations, which make the exchange of school leaders and staff difficult. Radical improvement has not been observed but we can form a clear thesis that a concerted action of different actors and a professional, profound and persistent approach on all levels of leadership is a necessary condition for school turnaround where unwillingness and incompetence to improve and dysfunctionality in organisational culture exist.

## F09

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Bernoullianum 32, H223

### Buchcharakteristika und Lesemotivation

**Franziska Maria Locher, Maximilian Pfof**

Universität Bamberg, Deutschland; [franziska.locher@uni-bamberg.de](mailto:franziska.locher@uni-bamberg.de)

#### Theoretischer Hintergrund

Ältere Jugendliche sind im Vergleich zu Kindern und jüngeren Jugendlichen oftmals weniger motiviert zu lesen (z.B. Chapman & Tunmer, 1995; McKenna, Kear, & Ellsworth, 1995; Smith, Smith, Gilmore, & Jameson, 2012). Diese Tatsache wirft die Frage auf, welche Faktoren zum Absinken der Lesemotivation mit zunehmendem Alter beitragen. Diesbezüglich wird häufig der wenig motivierende Deutschunterricht diskutiert (Garbe, 2010). Trotzdem fokussieren sich die meisten Untersuchungen, welche sich mit der Analyse von individuellen Unterschieden in der Lesemotivation beschäftigen, meist auf das freizeithliche Lesen und lassen die Rolle der Schule unbeachtet. Dabei bietet diese wichtige Gelegenheiten, um Kindern und Jugendlichen das Lesen nahezubringen.

#### Fragestellung

Beim schulischen Lesen entscheiden eher die Lehrkräfte, was gelesen wird und ein Nichterfüllen der Leseaufgaben führt gegebenenfalls zu negativen Konsequenzen für die SchülerInnen, d.h. schulisches Lesen ist oftmals external reguliert. Beim freizeithlichen Lesen entscheiden die SchülerInnen was, und ob sie überhaupt lesen wollen. Sie suchen sich das Genre aus, welches ihnen gefällt oder Texte, die ihrem Kompetenzniveau entsprechen. Daher ist anzunehmen, dass SchülerInnen beim freizeithlichen Lesen ein höheres Kompetenz- und Autonomieerleben verspüren und ihre Lesemotivation daher größer ist als im schulischen Kontext. Mit Bezug auf die Selbst-Bestimmungs-Theorie (Deci & Ryan, 1985, 2000) und Modelle zur situativen und habituellen Motivation (Guay, Vallerand, & Blanchard, 2000; Guthrie, Hoa, Wigfield, Tonks, & Perencevich, 2005) wurden folgende Forschungsfragen formuliert: Welche Unterschiede zwischen freizeithlichem und schulischem Lesen lassen sich hinsichtlich Genre, Lesbarkeit und Erscheinungsjahr finden? Ist die situative Motivation beim freizeithlichen Lesen höher als beim schulischen Lesen? Wie hängen Genre und andere Buchcharakteristika mit der situativen Motivation zusammen? Und schließlich: Sagt die situative Motivation die habituelle Motivation vorher?

#### Methode

Zur Bearbeitung der Fragestellung lagen Buchangaben von insgesamt 401 NeuntklässlerInnen vor. Buchtitel und Autorennamen wurden anhand der Frage „Bitte gib den Titel und Autor von höchstens drei der Bücher an, die du im letzten halben Jahr für den Deutschunterricht lesen musstest/ ...zum Vergnügen in deiner Freizeit gelesen hast.“ erhoben. Die 1142 Buchangaben wurden verwendet und anhand eines systematischen Kodierschemas bewertet. Die Inter-Rater-Reliabilität war gut (Genre:  $\kappa = .83$ ; Erscheinungsjahr:  $\kappa = .71$ ). Darüber hinaus wurden Buchausschnitte verwendet, um den Lesbarkeitsindex (LIX) zu bestimmen, ein Kriterium für die linguistische Oberflächenstruktur/Schwierigkeit eines Textes (Lenhard & Lenhard, 2014-2017). Die situative („Bitte gib für jedes dieser Bücher an, wie gern du auch weiterhin Bücher mit ähnlichem Inhalt lesen möchtest.“) und habituelle Motivation (4 Items z.B. „Lesen ist eins meiner liebsten Hobbies.“;  $\alpha = .89$ ) wurden im Rahmen der Fragebogen erfasst. Es liegt eine genestete Datenstruktur vor – Bücher (Level 1) geclustert in Personen (Level 2). Aus diesem Grund wurden Mehrebenenmodelle verwendet. Durch dieses Vorgehen kann untersucht werden, ob die Motivation der SchülerInnen zwischen Büchern und/ oder zwischen Personen variiert.

#### Ergebnisse

Erstens zeigte sich, dass die Lesemotivation beim freizeithlichen Lesen ( $M = 3.67$ ;  $SE = .03$ ) höher ist als beim schulischen Lesen ( $M = 2.19$ ;  $SE = .04$ ). Zweitens, lesen SchülerInnen im Schulkontext ältere Bücher ( $M = 3.40$ ;  $SE = .05$ ) und Bücher mit höheren Lesbarkeitsindices ( $M = 43.29$ ;  $SE = .63$ ), d.h. schwierigere Bücher als in ihrer Freizeit ( $M = 1.83$ ;  $SE = .05$ ;  $M = 38.84$ ;  $SE = .45$ ). Drittens scheint besonders das Lesen von Klassikern ( $B = -.60$  und  $-.55$ ;  $p < .01$ ) negativ mit der situativen Motivation zusammenzuhängen. Dieses Genre wird jedoch im Schulkontext besonders häufig gelesen. Und viertens konnte ein Zusammenhang zwischen situativer und habitueller Lesemotivation gefunden werden. Dies gilt sowohl für das freizeithliche ( $B = .42$ ;  $p < .01$ ) als auch das schulische Lesen ( $B = .51$ ;  $p < .01$ ).

Anhand dieser Ergebnisse kann gefragt werden, ob zu Gunsten der Lesemotivation eine Anpassung des Literaturunterrichts an die Präferenzen der SchülerInnen vorgenommen werden sollte. Andererseits ist unklar, ob nicht gerade klassische Texte solche Leseherausforderungen beinhalten, die für die Lesekompetenzentwicklung besonders wichtig sind. Dies macht das gleichzeitige verfolgen beider Ziele zu einer schwierigen Aufgabe (Spinner, 2009).

### Der Einfluss von Quellenexpertise auf die Rezeption von Evidenz

**Eva Thomm<sup>1</sup>, Rainer Bromme<sup>2</sup>**

<sup>1</sup>Universität Erfurt; <sup>2</sup>Westfälische Wilhelms-Universität Münster; [eva.thomm@uni-erfurt.de](mailto:eva.thomm@uni-erfurt.de)

#### Theoretischer Hintergrund

Laien müssen sich im Alltag zunehmend mit wissenschaftlicher Evidenz auseinandersetzen, um fundierte Entscheidungen treffen zu können (z.B. bzgl. gesundheitsbezogene Frage). Die Fähigkeit, wissenschaftliche Argumente anhand ihrer Evidenz bewerten zu können, ist ein zentrales Ziel der Entwicklung und Förderung von Scientific Literacy (Kind & Osborne, 2017; PISA, 2012). Hierfür ist es informativ, auch die von Laien bereits angewendeten Heuristiken bei der Beurteilung von Evidenz zu betrachten. Dies lässt sich insbesondere anhand der Rezeption gesundheitsbezogener Aussagen untersuchen. Empirische Studien im Gesundheitskontext zeigen, dass sowohl empirische als auch anekdotische Evidenz die Bewertung gesundheitsbezogener Aussagen beeinflussen (z.B. Hornikx, 2005). Es stellt sich jedoch die Frage, wie Laien diese tatsächlich bewerten können? Zu vielen Fragen verfügen sie oftmals nur über lückenhaftes Vorwissen, so dass sie die Gültigkeit der Evidenz nicht direkt bewerten können. Stattdessen müssen Laien auch die Quelle, die die Evidenz hervorbringt, berücksichtigen, um zu beurteilen, wie verlässlich die Evidenz und somit ein Argument ist (Bromme et al., 2010; Chinn et al., 2011). In der Persuasionsforschung wurde zwar der Einfluss von Eigenschaften des Arguments und der Quelle betrachtet (z.B. Petty & Cacioppo, 1986), dies geschah jedoch meist unabhängig voneinander, so dass bisher unklar bleibt, ob die Überzeugungskraft dieser Evidenzarten in Abhängigkeit ihrer Quelle variiert (Hahn et al., 2009).

#### Fragestellung



Ziel dieser Studie war es zu untersuchen, wie Laien Informationen über Evidenzart und die Expertise der Quelle, die die Evidenz liefert verknüpfen, um die Glaubwürdigkeit gesundheitsbezogener Aussagen zu bewerten. Wir nehmen an, dass eine Aussage dann als besonders glaubwürdig bewertet wird, wenn ihre Quelle Evidenz vorbringt, für die sie auch als zuständig und kompetent erachtet wird.

#### Methode

Die Studie folgte einem zweifaktoriellen Design mit den Innersubjektfaktoren Quellenexpertise (hohe vs. niedrige Expertise) und Evidenzart (empirische vs. anekdotische Evidenz). Studierende (N = 72) mit geringem Vorwissen lasen in jeder Bedingung eine kausale gesundheitsbezogene Aussage, die entweder durch empirische oder anekdotische Evidenz gestützt wurde. Um dies zu vermeiden, erhielten die Versuchspersonen als empirische Evidenz das Ergebnis einer Studie, die das jeweilige Thema anhand eines Kontrollgruppendesigns untersuchte. Im Kontrast hierzu beschrieb die anekdotische Evidenz eine beobachtete Erfahrung der jeweiligen Quelle. Quellenexpertise wurde variiert, indem der Autor des Arguments entweder ein Laie oder ein Experte war. Es wurden folglich parallele Argumente zu vier verschiedenen gesundheitsbezogenen Themen entwickelt, die über die Bedingungen hinweg ausbalanciert präsentiert wurden. Wir erfassten mittels Likert-Skalen das Vorwissen, die persönliche Zustimmung zur jeweiligen Aussage sowie die Glaubwürdigkeit der Aussage und der Quelle.

#### Ergebnisse

Effekte der Faktoren wurde mittels Varianzanalysen mit Messwiederholung getestet. Interaktionseffekte wurden durch t-Tests für abhängige Stichproben analysiert. Die Analysen belegen einen Interaktionseffekt der Evidenzart und Quellenexpertise in der Bewertung der Aussagenglaubwürdigkeit. Eine Aussage wurde dann als besonders glaubwürdig bewertet, wenn sie durch empirische Evidenz gestützt und zugleich von einem Experten hervorgebracht wurde. Gleichzeitig beeinflusste die Evidenzart nicht die Bewertung, wenn eine Laienquelle die Aussage traf. Analog ergab sich ein Interaktionseffekt auf die Bewertung der Quellenglaubwürdigkeit. Die Expertenquelle wurde als glaubwürdiger eingeschätzt, wenn sie empirische Evidenz anführte. Die präsentierte Evidenzart hatte jedoch keinen Einfluss darauf, wie die Laienquelle bewertet wurde. Die Analyse der Zustimmung zur Aussage bestätigte Haupteffekte der Faktoren Quellenexpertise, und Evidenzart. Die Versuchspersonen stimmten der Aussage eher zu, wenn die Quelle eher ein Experte war. Zudem stimmten sie der Aussage eher zu, wenn sie durch empirische Evidenz gestützt wurde.

Unsere Ergebnisse zeigen, dass Laien Informationen über Evidenzart und Quellenexpertise aufeinander abstimmen, wenn sie die Glaubwürdigkeit von Quelle und Aussage bewerten. Allerdings scheinen Sie diese Informationen nicht in gleicher Weise zu berücksichtigen, wenn sie ihre Zustimmung zu Aussagen einschätzen. Dennoch zeigt sich, dass Laien Aussagen, die empirische Evidenz enthalten bzw. von einem Experten stammen, eher zu stimmen.

## **Sprachförderliches Lehrerhandeln im Mathematikunterricht**

**Magdalena Schulte, Anita Schilcher**

Universität Regensburg, Deutschland; [magdalena.schulte@ur.de](mailto:magdalena.schulte@ur.de)

#### Theoretischer Hintergrund

Die Erkenntnis, dass Sprache im Fachunterricht als Schlüssel für den schulischen Erfolg gilt (Becker-Mrotzek et al. 2013; Gogolin 2014), stellt nicht zuletzt aufgrund der anhaltenden Migrationsbewegungen hohe Ansprüche an das professionelle Handeln von Lehrkräften aller Fachrichtungen. Doch auch für Kinder mit Deutsch als Erstsprache erweist sich der sprachensible Fachunterricht als gewinnbringend, da auch sie das bildungssprachliche Register der Institution Schule im Laufe ihres Bildungsweges erwerben bzw. ausbauen müssen (Vollmer et al. 2010). Lehrkräften wird somit auf der einen Seite ein hohes Professionswissen hinsichtlich Sprachförderung abverlangt; andererseits müssen sie in der Lage sein, ihr theoretisches Wissen umzusetzen, indem sie im Unterricht gute Sprachfördersituationen schaffen (Hopp et al. 2010). Die im Rahmen des Projekts Eva-Prim eingebettete Video-Studie rückt das sprachförderliche Lehrerhandeln in den Fokus. Im Zentrum steht hierbei der Mathematikunterricht, dessen sprachliche Hürden oftmals die mathematische Performanz der Lerner einschränken (Abshagen 2015). Als besonders förderlich für den Ausbau des bildungssprachlichen Registers erweist sich nach Fürstenau (2012) eine möglichst umfangreiche Produktion eigener Äußerungen in diesem Register. Für den Unterricht bieten sich also vor allem reziproke Interaktionsmuster an, in denen sich die Lehrkraft stärker zurücknimmt (Fürstenau 2012). Große Bedeutung kommt dabei den durch die Lehrerfrage intendierten Sprachprozeduren der Schülerantworten zu. Die Bildungsstandards für das Fach Mathematik fordern hier mit Sprachprozeduren des Beschreibens und Begründens sprachintensive Diskursfunktionen (KMK 2005).

#### Fragestellung

Anhand von Videografien wird die sprachliche Anregungsqualität durch Lehrkräfte im Mathematikunterricht unter Fokussierung auf die Lehrerfrage und die damit intendierten Diskursfunktionen deskriptiv erfasst. Zudem erfolgt eine Analyse der sprachlichen Qualität der Schüleräußerungen in Abhängigkeit von der Lehrerfrage. Darüber hinaus wird untersucht, ob zwischen dem fachdidaktischen Sprachförderwissen und dem sprachförderlichen Lehrerhandeln im Unterricht ein Zusammenhang besteht.

#### Methode

Als Datengrundlage dieser qualitativ angelegten Studie dienen acht Videografien, die Einführungsstunden in sprachintensive Themenkomplexe des Mathematikunterrichts der dritten Jahrgangsstufe mithilfe eines Zwei-Kamera-Designs im Zuge des zweiten Messzeitpunktes der Evaluation im Februar und März 2017 erfassen. Die Daten wurden mithilfe von niedrig bzw. mittel inferenten Kategoriensystemen analysiert. Die Analyse der Transkripte erfolgte zunächst hinsichtlich der auftretenden Sozialformen sowie der Verteilung der Redebeitragsdauern. Hieran schloss eine Kodierung der Lehrerfragen an, die in Hinblick auf ihre intendierten Diskursfunktionen klassifiziert wurden. Den Abschluss bildet eine Analyse der sprachlichen Qualität der Schülerantworten, die auf die kodierten Lehrerfragen folgen.

Da im Zuge des zweiten Messzeitpunktes der Evaluation zudem das Professionswissen der Lehrkräfte in puncto Sprachförderung mittels des im Projekt entwickelten Tests „FaWi-S“ getestet wurde, wird im Anschluss das unterrichtliche Handeln der videografierten Lehrkräfte mit ihrem Sprachförderwissen in Beziehung gesetzt.

#### Ergebnisse

Erste Ergebnisse zeigen, dass der analysierte Mathematikunterricht zu 68,5% im Klassenplenum stattfindet (31,5%: Nicht-Plenumsphasen), also der fragend-entwickelnde Unterricht dominiert. Da die Plenumsphasen einen durchschnittlichen Redebeitragsumfang der Lehrkräfte von 68,4% aufweisen, die Schülerbeiträge hingegen lediglich 21,6% ausmachen (Rest:

Schweigen), lässt dies auf sehr lehrerzentrierte Diskurse schließen. Diese verbale Dominanz der Lehrkräfte rührt in der untersuchten Stichprobe nicht zuletzt aus dem Frageverhalten der Lehrenden: Bei durchschnittlich 64,5% der aufgezeichneten Lehrerfragen handelt es sich um ENTSCHEIDEN- bzw. BENENNEN-Fragen, auf die der Schüleroutput erwartungsgemäß kurz ausfällt (durchschnittlicher Umfang der Ein-Wort-Antworten nach ENTSCHEIDEN-Fragen: 72,1%; nach BENENNEN-Fragen: 53,3%). Die in den Bildungsstandards geforderten komplexen Sprachprozeduren hingegen werden in der untersuchten Stichprobe deutlich seltener eingefordert (BESCHREIBEN: 23,7%; ERKLÄREN/BEGRÜNDEN: 7,1%), böten jedoch weitaus komplexere Sprechansätze. So weisen 66% der Schülerantworten nach Fragen, die auf Erklärprozesse abzielen, eine minimale satzwertige Einheit (bestehend aus Subjekt und Finitum; vgl. Griefhaber 2013) auf. Um sprachlich komplexere Schülerantworten zu elizitieren, sind also Lehrerfragen erforderlich, die auf komplexere Sprachhandlungen abzielen.

Abschließend wird eine Untersuchung des Zusammenhangs zwischen Lehrerhandeln und -wissen vorgenommen. Erste Ergebnisse des FaWi-S (maximal erreichbar: 34 Punkte) zeigen bereits im fachdidaktischen Wissen der acht Lehrkräfte eine große Streuung (Min: 16,46 Punkte; Max: 22,58 Punkte).

## **Einstellungen von LektorInnen zu mehrsprachigkeitsdidaktischen Ansätzen und deren Einsatz in der universitären Fremdsprachenlehre – eine Begleitforschungsstudie in der "Sprachpraxis Spanisch" in der ersten Phase der Lehrerbildung**

**Isabelle Mordellet-Roggenbuck, Eva Maria Nelz**

Institut für Romanistik, Pädagogische Hochschule Freiburg, Deutschland; [isabelle.mordelletroggenbuck@ph-freiburg.de](mailto:isabelle.mordelletroggenbuck@ph-freiburg.de),  
[eva.nelz@ph-freiburg.de](mailto:eva.nelz@ph-freiburg.de)

Seit den 90er Jahren hat sich innerhalb der romanistischen Spracherwerbs- und fremdsprachendidaktischen Forschung die Mehrsprachigkeitsdidaktik (Meißner/Reinfried 1998; Martinez et al. 2006) sowie die zugehörige Interkomprehensionsdidaktik (Mordellet-Roggenbuck 2011; Ollivier/Strasser 2013; Morkötter 2016) etabliert. Es konnte gezeigt werden, dass die Implementierung mehrsprachigkeitsdidaktischer Anteile im Fremdsprachenunterricht Aspekte, wie Language Awareness, Sprachenlernen, Lernmotivation oder Persönlichkeitsentwicklung fördern kann. Folgerichtig wird Mehrsprachigkeit in den neuen Bildungsplänen der Sekundarstufe I und II als eine individuelle Kompetenz definiert, deren Herausbildung die beiden Schulfremdsprachen Französisch und Spanisch, neben Englisch, erheblich fördern können. Dabei werden sowohl das große Transferpotential des Französischen „als Brückensprache [...] im Kontext der Mehrsprachigkeit“ (Bildungsplan Gymnasium F1 2016: 6) als auch der Beitrag des Spanischen „zur angestrebten Mehrsprachigkeit“ (Bildungsplan Gymnasium SPA3 2016: 6) betont.

Die Umsetzung mehrsprachigkeitsdidaktischer Verfahren wirft jedoch die Frage der Professionalisierung von Fremdsprachenlehrpersonen auf, da sie die Umorientierung von einer monolingualen hin zu einer multilingualen Ausrichtung der Lehre fordert und sowohl die Komponenten savoir, savoir-faire als auch savoir-être umfasst (vgl. Martinez 2015: 14).

Bislang durchgeführte Untersuchungen zur Mehrsprachigkeitsdidaktik nehmen fast ausschließlich auf den schulischen Kontext und die Lernerperspektive Bezug. Zweifelsohne ist der Forschungs- und Praxisbereich der Mehrsprachigkeitsdidaktik jedoch viel facettenreicher und tangiert die nachschulische, universitäre Lehrperspektive aus vielen Gründen gleichermaßen:

So legt zuallererst die Gegebenheit, dass der Großteil der (Lehramts-)Studierenden an deutschen Hochschulen selbst mehrsprachig ist (vgl. May 2012: 37) den Einsatz von mehrsprachigkeitsdidaktischen Ansätzen in den Sprachkursen nahe.

Genauso signalisiert das geringe Zeitvolumen, welches der "Sprachpraxis" in den Bachelor- und M.Ed.-Studiengängen zur Verfügung steht, die Notwendigkeit, neue Wege zu gehen und Synergien zu nutzen (vgl. Meißner 2011: 38).

Im Falle der sprachpraktischen Ausbildung angehender FremdsprachenlehrerInnen kommt zudem die didaktische Vorbildfunktion der LektorInnen zum Tragen, deren Lehrtätigkeit die Einstellungen der Studierenden bzgl. effektiven Lehrens, Lernens und Unterrichtspraktiken nachdrücklich prägen kann (vgl. Crandall 2000: 35).

Übergeordnet stellt sich also die Frage, inwieweit die universitäre "Sprachpraxis" die angehenden FremdsprachenlehrerInnen auf mehrsprachigkeitsdidaktische Ansätze vorbereitet.

Vor diesem Hintergrund und ausgehend von der Feststellung, dass die Lehrtätigkeit der LektorInnen bislang selten im Fokus fremdsprachendidaktischer Forschung stand, die Hochschullehre aber zunehmend in den Fokus bildungspolitischer Bemühungen und hochschulischer Organisationsprozesse rückt (Jütte et al. 2017: 1), ergeben sich zunächst folgende Fragestellungen:

Welche Einstellungen zum Themenkomplex Mehrsprachigkeit/Mehrsprachigkeitsdidaktik haben die LektorInnen? Werden gängige Konzepte und Methoden in der "Sprachpraxis" eingesetzt? Kann eine Fortbildungsmaßnahme zum Thema "Mehrsprachigkeitsdidaktik" sowohl auf die Einstellungen der LektorInnen als auch deren Lehrtätigkeit Einfluss nehmen?

In diesem work-in-progress-Beitrag soll eine Begleitforschungsstudie zu einer Fortbildungsmaßnahme der Erasmus+ Personalmobilität vorgestellt werden, die aus einer Fragebogenerhebung vor sowie einer Follow-up-Erhebung nach der Fortbildung besteht.

Ausgehend von einer einleitenden Befragung zu individuellen Sprach- und Sprachlernerfahrungen der LektorInnen sowie Fragen zur Förderung von Mehrsprachigkeit im Sinne europäischer Rahmentexte zur Hochschullehre, richtete sich das Interesse des Fragebogens zunächst auf die Einstellungen der TeilnehmerInnen zu Sprachtransfer, Sprachvergleich und Sprachreflexion in der universitären Fremdsprachenlehre. Es sollte überprüft werden, ob die LektorInnen mehrsprachigkeitsdidaktischen Verfahren gegenüber positiv eingestellt sind. Im Anschluss wurden die Erfahrungen der LektorInnen mit konkreten Verfahren zu diesen Aspekten erhoben.

Zwei Monate nach der Fortbildungsmaßnahme soll die Erhebung während der Vorlesungszeit wiederholt werden. Dabei gehen wir davon aus, dass die Fortbildung sowohl hinsichtlich der Einstellungen als auch der Verwendung der Verfahren Einfluss auf die LektorInnen haben wird.

## F10

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus S107

### Fachstudium in Mathematik: Zusammenhänge mit Mathematikleistungen von Grundschulkindern und ihrer Wahrnehmung der Unterrichtsqualität

**Raphaela Porsch<sup>1</sup>, Heike Wendt<sup>2</sup>**

<sup>1</sup>Carl von Ossietzky Universität Oldenburg, Deutschland; <sup>2</sup>TU Dortmund, Deutschland; [porsch@uni-muenster.de](mailto:porsch@uni-muenster.de)

Professionelles Handeln von Lehrkräften wird ermöglicht, da diese über berufsrelevante Kompetenzen verfügen, zu denen Fachwissen als auch fachdidaktisches Wissen gehören (Baumert & Kunter, 2006). Die Lehramtsausbildung bietet umfangreiche Lerngelegenheiten zum Erwerb von Professionswissen (Kunter et al., 2011). Die Verfügbarkeit professioneller Kompetenzen lässt erwarten, dass sie effektives Unterrichtshandeln unterstützen, das zu einem Lernzuwachs bei Schülerinnen und Schülern beiträgt. Respektive wäre zu erwarten, dass fehlendes Fachwissen weniger qualitätsvolleren Unterricht bedingt und Lernergebnisse geringer ausfallen. International (z.B. Dee & Cohodes, 2008) und national (z.B. Porsch & Wendt, 2017) liegen zahlreiche Arbeiten vor, die die Zusammenhänge der formalen Qualifikation von Lehrkräften auf Schülerleistungen bzw. den Output geprüft haben. Kaum im Fokus der Forschung stand bisher jedoch die Frage, ob sich die häufig schlechteren Lernergebnisse – im Vergleich der Lernenden, die bei Lehrkräften mit einer Fachausbildung unterrichtet werden – durch die Qualität des Unterrichts bzw. die Prozessebene erklären lassen. In der COACTIV-Studie (Kunter et al., 2013) zeigte sich, dass das Wissen von Mathematiklehrkräften vermittelt über die Unterrichtsqualität positiv auf die Schülerleistungen in der Sekundarstufe I wirkt. Blömeke et al. (2016) sind für die Grundschule auf Grundlage von Daten aus TIMSS-2011 der Frage nachgegangen, ob die Lehrerqualifikation Unterschiede in der Unterrichtsqualität – modelliert als ein eindimensionales Konstrukt – erklären. Es zeigt sich ein negativer Zusammenhang zwischen dem Studienschwerpunkt in Mathematik und Selbstauskünften zu Indikatoren der qualitätsvollen Unterrichtsgestaltung. Ergebnisse qualitativer Studien (z.B. Du Plessis, 2015) verweisen hingegen ebenfalls auf einen positiven Einfluss der Ausbildung auf die Qualität des Unterrichts. Unsere Fragen lauten daher: (1) Kann der Fachabschluss von Lehrkräften für das Fach Mathematik die Mathematikleistungen von Schülerinnen und Schülern in Klasse 4 erklären? (2) Liegt ein Zusammenhang von Schülerleistungen im Fach Mathematik und der Unterrichtsqualität vor? (3) Kann die Fachausbildung der Lehrkräfte Unterschiede in der wahrgenommenen Unterrichtsqualität erklären? Auf Grundlage der nationalen Erweiterung zu TIMSS-2015 soll der Zusammenhang von Lehrerqualifikation, Unterrichtsqualität und Schülerleistungen im Fach Mathematik der Grundschule analysiert werden. Im Vergleich zu TIMSS-2011 wurde in der Folgestudie die Lehrerqualifikation (n = 210) differenzierter operationalisiert (Mathematik als Schwerpunktfach: 64%, als Nebenfach: 17%, ohne Mathematikstudium: 19%). Zudem wurde die Unterrichtsqualität, die durch Schülerbewertungen (n = 3363) erfasst wurde, in drei Dimensionen unterschieden: Kognitive Aktivierung [Fauth et al., 2014; 7 Items,  $\alpha = .75$ ], Unterstützendes Klima [Fauth et al., 2014; 7 Items,  $\alpha = .75$ ] und Klassenführung [Baumert et al., 2014; 7 Items,  $\alpha = .79$ ]. Für die Analysen wurden die Daten imputiert (FIML) und gewichtet (TCHWGT). Modelle aus Mehrebenenanalysen wurden mit Hilfe von SAS spezifiziert, die neben der Lehrerqualifikation weitere Merkmale der Lehrkräfte (u.a. Fortbildungsteilnahme, Berufsjahre) sowie der Schülerinnen und Schüler (u.a. Mathematikleistungen, sozio-ökonomischer Hintergrund, kognitive Grundfähigkeiten) berücksichtigen. Die Ergebnisse zeigen, dass kein signifikanter Zusammenhang zwischen der Qualifikation der Mathematiklehrkräfte und den Mathematikleistungen der Schülerinnen und Schüler vorliegt (F1). „Kognitive Aktivierung“ und „Unterstützendes Klima“ stehen in einem negativen Zusammenhang mit den Mathematikleistungen, „Klassenführung“ korreliert positiv mit den Mathematikleistungen, jedoch ist der Zusammenhang gering. (F2). Schließlich zeigt sich ein signifikanter Zusammenhang zwischen der Qualifikation der Mathematiklehrkräfte und der Unterrichtsqualität (F3): Bei alleiniger Betrachtung der Fachausbildung kann das Merkmal „Hauptfach Mathematik im Studium“ Unterschiede in der wahrgenommenen Unterrichtsqualität, „Kognitive Aktivierung“ und „Unterstützendes Klima“, erklären. Bei Berücksichtigung weiterer Merkmale können „Kognitive Aktivierung“ und „Klassenführung“ durch das Merkmal „Hauptfach Mathematik im Studium“ erklärt werden. Zwar kann das „Unterstützende Klima“ bei Modellierung weiterer Prädiktoren nicht durch die Fachausbildung der Lehrkräfte erklärt werden, aber hier scheint ein weiteres Merkmal der Lehrkräfte – ihre Unterrichtserfahrung – prädiktiv zu sein. Die Ergebnisse deuten darauf hin, dass Unterschiede in der Fachausbildung der Mathematiklehrkräfte die Wahrnehmung der Unterrichtsqualität aus Sicht von Schülerinnen und Schülern an Grundschulen erklären können. Implikationen für weitere Forschungsarbeiten werden abgeleitet sowie eine Einordnung der Ergebnisse in das Forschungsfeld vorgenommen.

### Das nonverbale Verhalten Lernender: Unterschiede bei Lernenden in Abhängigkeit von ihren Kompetenzerwartungen an Lehrende

**Marcel Hackbart, Barbara Thies**

Technische Universität Braunschweig, Deutschland; [m.hackbart@tu-braunschweig.de](mailto:m.hackbart@tu-braunschweig.de)

Klassenführung als Merkmal von Professionalität (vgl. Baumert & Kunter, 2006) ist eine Interaktionsgestaltung zwischen einer Lehrperson und Lernenden (als Gruppe und als Individuum), die die Störungen im Unterricht minimiert und somit die Bereitstellung maximaler Lernzeit ermöglicht (vgl. Bildungsdirektion Kanton Zürich, 2011; Kiel, Frey & Weiß, 2013; Weinert, 1996). Die Lehr-Lern-Beziehung spielt eine wesentliche Rolle, so zeigen sich Schüler/innen in positiven, vertrauensvollen Beziehungen mit ihren Lehrkräften u. a. motivierter und haben ein positiveres Bild ihres Lernerfolgs (Schweer, 1997).

Erwartungen, die sich u.a. auf Stereotype beziehen, beeinflussen die Lehr-Lern-Beziehung selbsterfüllend, insbesondere bei beginnenden Interaktionen (Jussim & Harber, 2004). Die Kompetenzerwartungen, die Lehrende gegenüber Lernende halten, wurden bisher gut untersucht (vgl. Jussim & Harber, 2004; Tauber, 1998). Gemäß dem transaktionalen Modell von Nickol (1985) haben jedoch auch die Erwartungen der Lernenden einen Beitrag zur Interaktionsgestaltung. Lernende, die eine kompetente Lehrperson erwarten, haben höhere Leistungen und erleben den Unterricht positiver (Abediwura & Tayo, 2007; Edwards, Edwards, Shaver & Oak, 2009; Feldman & Prohaska, 1979) als Lernende, die neutrale oder inkompetente Lehrpersonen erwarten. Die Auswirkungen der Kompetenzerwartungen von Lernenden an Lehrende auf das Verhalten der Lernenden wurden in nur wenigen Studien fokussiert. So verhielten sich Lernende, die eine kompetente Lehrperson erwarten, zugewandter (Feldman & Prohaska, 1979) und angemessener (z.B. Ausführen der übertragenden Aufgaben, Unterlassen von Zwischenrufen; Jamieson, Lydon, Stewart & Zanna, 1987). Eine Fokussierung der Lehrpersonen-Wahrnehmung auf das nonverbale Verhalten der Lernenden (vgl. diagnostische Kompetenz; Baumert & Kunter, 2006) kann ein Ansatz für lehrbezogenes Feedback und somit für die Professionalisierung der Lehrperson sein. Es stellt sich also die Frage, wie das nonverbale Verhalten von Lernenden mit den Kompetenzerwartungen an Lehrende zusammenhängt.

In der vorgestellten Untersuchung wird der Einfluss der Kompetenzerwartungen von Lernenden an eine Lehrperson auf das nonverbale Verhalten der Lernenden in einem Laborsetting untersucht. Die Teilnehmenden (N=62; Geschlecht: 12 männlich, 50 weiblich; Alter: M=24.8, SD=8.9) wurden unter Verwendung der Cover-Story, es handle sich um eine Lehrkräfte-Supervision, eingeladen. Pro Durchgang konnten bis zu vier Personen teilnehmen. Die Kompetenzerwartungen an die Lehrkraft wurden bei der Untersuchung durch Vorlegen von Zeugnissen manipuliert: In der Kompetenz-Bedingung erhielten die Teilnehmenden ein Zeugnis mit einer 1.3 als Note und in der Inkompetenz-Bedingung ein Zeugnis mit einer 3.3. Anschließend sollten die Teilnehmenden einen Vortrag über das nicht-schulfachbezogene Thema Nägelkaue der jeweiligen Lehrkraft ansehen. Den Teilnehmenden wurde erzählt, dass die Lehrperson im Gebäude ist und eine Live-Übertragung erfolgt, wodurch die Aufnahme der Teilnehmenden auf Video gerechtfertigt wurde (fiktive Live-Übertragung). Tatsächlich handelte es sich um eines von zwei vorher aufgezeichneten Videos mit einer weiblichen oder männlichen Person (faktisch dieselbe Konföderierte in eher „weiblicher“ oder „männlicher“ Kleidung; die Stimmhöhe wurde je in androgynen Bereich transformiert). Durch die Variation des Geschlechts sollte die Generalisierbarkeit erhöht werden und mit den Videos der Effekt der Lehrkraft konstant gehalten werden. Nach Abschluss des Vortrags wurden die Teilnehmenden gebeten, einen Fragebogen zu beantworten, in dem sie u.a. die Kompetenz der Lehrperson beurteilten. Das nonverbale Verhalten wurde in Anlehnung an das Schema von Mehrabian (1969) ausgewertet. Dafür wurden pro Untersuchungsdurchgang 2:44 Minuten des Videomaterials mit den Teilnehmenden am Ende des Vortrags der Lehrkraft herangezogen, die durch zwei Auswerterinnen, die blind gegenüber der Fragestellung waren, ausgewertet wurden.

In ersten Analysen zeigt sich, dass die Kompetenzbedingung nicht mit der Kompetenz-Einschätzung zusammenhängt ( $r=-.03$ ,  $p>.10$ ). Nichtsdestotrotz hielten Teilnehmende, die eine kompetente Lehrkraft erwarteten, weniger lang Blickkontakt mit der Lehrkraft bzw. dem Monitor ( $r=-.25$ ,  $p<.10$ ), stützten den Kopf kürzere Zeit ab ( $r=-.26$ ,  $p<.05$ ) und verschränkten die Arme länger ( $r=.24$ ,  $p<.10$ ) als Teilnehmende, die eine inkompetente Lehrkraft erwarteten. Ein Zusammenhang zwischen Kompetenz-Einschätzung, Geschlechts-Bedingung und nonverbalen Verhalten konnte nicht ausgemacht werden. Dies spricht für zwei unterschiedliche „Langeweileprofile“ auf Seiten der Lernenden in Abhängigkeit ihrer Kompetenzerwartung an die Lehrkraft.

### **Wenn Regel- und SonderpädagogInnen in inklusiven Klassen unterrichten: Eine Studie zur Unterrichtsqualität im Kontext von „nested instruction“**

**Helena Krähenmann<sup>1</sup>, Franziska Felder<sup>1</sup>, Sarah Jandl<sup>1</sup>, Susanne Schnepel<sup>1</sup>, Rachel Sermier Dessemontet<sup>2</sup>, Elisabeth Moser Opitz<sup>1</sup>**

<sup>1</sup>Universität Zürich, Schweiz; <sup>2</sup>haute école pédagogique vaud, Schweiz; [hkraehenmann@ife.uzh.ch](mailto:hkraehenmann@ife.uzh.ch)

#### Theoretischer Hintergrund

Zur Qualität von inklusivem Unterricht liegen bislang erst wenige Untersuchungen vor. Das hängt unter anderem damit zusammen, dass solche Studien aufgrund der sogenannten „nested instruction“ (Jones & Brownell, 2014) anspruchsvoll sind. Wenn eine Klassenlehrperson (KLP) und eine Sonderpädagogin oder ein Sonderpädagoge (SP) gemeinsam für den Unterricht verantwortlich sind, ist es schwierig, Aussagen zum Einfluss der einzelnen Lehrkraft auf das Unterrichtsgeschehen zu machen, insbesondere weil auch davon auszugehen ist, dass bei KLP und SP aufgrund der unterschiedlichen Ausbildung und beruflichen Rollen andere Faktoren eine Rolle spielen.

Als zentral für die Gestaltung von inklusivem Unterricht wird häufig die Zusammenarbeit der Lehrkräfte genannt, ein Faktor, der beide Lehrkräfte in gleichem Maß betrifft (z.B. Lütje-Klose & Urban, 2014). Seitens der KLP wird die Einstellung zur Inklusion als zentral erachtet, insbesondere im Zusammenhang mit Selbstwirksamkeitsüberzeugungen (Ruhberg et al., 2017). Bei SP ist ein solcher Zusammenhang nicht nachgewiesen, zudem zeigen SP, die in inklusiven Klassen arbeiten eine positivere Einstellung zur Inklusion als Lehrkräfte in separativen Settings (Gebhardt et al., 2011). Da die Verantwortung für besondere Fördermaßnahmen in der Regel den SP übertragen wird (z.B. Steppacher, 2014; Preiß et al., 2016) kann zudem angenommen werden, dass bei den SP das Fachwissen zu besonderen Fördermaßnahmen eine Rolle spielt. Der Einfluss von den genannten Faktoren auf die Unterrichtsqualität ist allerdings bis anhin noch nicht untersucht worden.

Unterrichtsqualität kann mit unterschiedlichen Variablen wie Passung bzw. Umgang mit Heterogenität (Helmke, 2014) operationalisiert werden. Für die vorliegende Untersuchung wird „Differenzierung“ gewählt. Differenzierung ist ein zentraler Aspekt von inklusivem Unterricht (z.B. Broderick et al., 2015; Jordan & Stanovich, 2001) – insbesondere in Klassen, in denen Kinder mit dem Förderschwerpunkt geistige Entwicklung (FgE) unterrichtet werden.

#### Fragestellung

Welcher Zusammenhang besteht zwischen der Einstellung von KLP zur Inklusion von Kindern mit FgE, dem professionellen Wissen der SP zur mathematischen Förderung von Kindern mit Förderbedarf, der Zufriedenheit der Lehrkräfte mit ihrer Zusammenarbeit und der Umsetzung von zieldifferentem Lernen im inklusiven Mathematikunterricht der Grundschule?

#### Methode

Die Stichprobe setzt sich aus 34 inklusiven Grundschulklassen (Klasse 2 und 3), in denen mindestens ein Kind mit FgE unterrichtet wurde, zusammen. In allen Klassen wurde eine Arithmetikstunde videografiert, in der die KLP und die SP (N = je 34) unterrichteten.

Bei den KLP wurde die Einstellung zur Inklusion von Kindern mit FgE mit dem Fragebogen „Opinion Relative to Integration“ (ORI; Antonak & Larrivee, 1995) erhoben (25 Items, Cronbachs Alpha .86). Bei den SP wurde ein Fragebogen zum Wissen zur mathematischen Förderung von Kindern mit Förderbedarf eingesetzt (24 Items, Cronbachs Alpha .81). Die Zufriedenheit mit der Zusammenarbeit wurde mit einem Fragebogen (Eigenentwicklung) erfasst (22 Items, Cronbachs Alpha .91). Die Videoaufnahmen (Filmmaterial: 3329 min.) wurden mit einem Instrument u.a. hinsichtlich des zieldifferenten Lernens im Mathematikunterricht mit mehreren (Negativ-)Indikatoren hoch inferent geratet (Generalisierbarkeitskoeffizient relativ .94).

#### Ergebnisse

Die Ergebnisse einer Regressionsanalyse mit der abhängigen Variable „zieldifferentes Lernen“ und den Prädiktoren „Einstellung der KLP zur Inklusion von Kindern mit IB“, dem „Wissen der SP“ sowie der „Zufriedenheit bezüglich der Zusammenarbeit der Lehrkräfte“ ergeben eine Varianzaufklärung von  $R^2=.23$ . Einen signifikanten Beitrag dazu leistet ausschließlich der Prädiktor Wissen der SP ( $\beta=.51$ ,  $p<.01$ ).

#### Diskussion

Entgegen den Erwartungen erweisen sich weder die Einstellung der KLP noch die Zufriedenheit mit der Zusammenarbeit als signifikante Prädiktoren für die Umsetzung von zieldifferentem Lernen. Das Wissen der SP trägt zur Varianzaufklärung bei, was

durch die Verantwortungsübertragung bezüglich Fördermaßnahmen an die SP (Steppacher, 2014; Preiß et al., 2016) erklärt werden kann. Allerdings stellen sich auch Fragen zum gewählten Vorgehen, insbesondere weil bei den zwei Gruppen von Lehrkräften unterschiedliche Daten erhoben worden sind. Dies wird im Beitrag kritisch diskutiert.

## F11

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus, Aula 33

### Die Veränderung der Selbstwirksamkeitserwartungen und Kompetenzeinschätzungen von Lehramtsstudierenden durch die Praxisphase

Chair(s): **Katja Hartig** (Goethe-Universität Frankfurt, Deutschland)

DiskutantIn(nen): **Olga Kunina-Habenicht** (Pädagogische Hochschule Karlsruhe)

Selbstwirksamkeitserwartung bezeichnet „die subjektive Gewissheit, neue oder schwierige Anforderungssituationen aufgrund eigener Kompetenz bewältigen zu können“ (Schwarzer & Jerusalem, 2002, S. 35). Dabei muss der Glaube daran, eine Situation meistern zu können nicht unbedingt mit den eigenen Fähigkeiten übereinstimmen. Aber die positive Einstellung zur Situation und die Überzeugung, in einer Situation kompetent handeln zu können, hilft beim erfolgreichen Handeln in herausfordernden Situationen (Schmitz & Schwarzer, 2000; Schwarzer & Jerusalem, 2002; Schwarzer & Schmitz, 1999). Damit stellt die Selbstwirksamkeitserwartung einen wichtigen Aspekt professioneller Kompetenz für Lehrkräfte dar (Baumert & Kunter, 2011).

Praxisphasen spielen eine zentrale Rolle im Lehramtsstudium. Hier entwickeln Lehramtsstudierende ihr Können durch das Sammeln von Erfahrungen in der Schulpraxis. Dabei werden sie von schulischen und universitären Mentoren begleitet und bei der Reflexion beobachteter und selbst erfahrener Praxis unterstützt. Der Qualität der Begleitung hat dabei eine große Bedeutung für das Gelingen der Praxisphase (Gröschner, Schmitt & Seidel, 2013; Stiller, 2016). Ergebnisse zur Erforschung der Wirkung von Praxisphasen im Lehramtsstudium zeigen einen bedeutsamen Zuwachs in der Kompetenzeinschätzung der Studierenden während der Praxisphase. Überdies zeigt sich, dass dieser wahrgenommenen Zuwachs an Kompetenzen nicht nur mit der Qualität der schulischen und universitären Betreuung sondern auch mit der Selbstwirksamkeitserwartung der Studierenden in Zusammenhang steht (Schubarth, Gottmann & Krohn, 2014).

Das Symposium umfasst drei Beiträge, die sich mit der Veränderung von Selbstwirksamkeitserwartungen und selbst eingeschätzten Kompetenzen von angehenden LehrerInnen während verschiedener Praxisphasen beschäftigen.

Der erste Beitrag von Bartels et al. widmet sich der Frage, welchen Einfluss die Betreuung während der Praxisphase und die die Praxisphase begleitende Reflexion in Form von Portfolioarbeit auf die Selbstwirksamkeit im Hinblick auf inklusiven Unterricht und die Einstellung zur Gestaltung und den Effekten inklusiven Unterrichts haben. Ziel ist es, Folgerungen zu generieren für die eine qualitativ gute Betreuung während des Praktikums.

Auch der zweite Beitrag beschäftigt sich mit dem Einfluss der Betreuung während der Praxisphase. Klingebiel et al. gehen der Frage nach, welcher Zusammenhang zwischen der Kompetenzselbsteinschätzung und der Einschätzung der studentischen Kompetenzen durch schulische und universitäre Betreuer besteht.

Dem dritten Beitrag von Knuth-Herzig et al. liegt die Annahme zugrunde, dass sich Berufskonzepte und berufliches Handeln gegenseitig bedingen. Im Hinblick auf Selbstwirksamkeitserwartungen kann man davon ausgehen, dass aufgrund mangelnder Praxiserfahrung die Selbstwirksamkeitserwartung der Studierenden zu Beginn der Praxisphase wahrscheinlich eher „theoretisch“ und durch die Vorstellungen über den Beruf beeinflusst wird und sich im Laufe der Praxisphase mit zunehmender Erfahrung im Umgang mit den SchülerInnen und der eigenen Rolle als angehende Lehrkraft verändert. Es wird der Frage nachgegangen, ob sich der Einfluss von Berufskonzepten auf die Selbstwirksamkeitserwartungen der Studierenden über die Praxisphase hinweg verändert.

Die Beiträge werden diskutiert von Olga Kunina-Habenicht.

*Beiträge des Symposiums*

### Der Einfluss von Betreuung in der Praxisphase und Portfolioarbeit auf die Selbstwirksamkeit und Einstellung in Bezug auf Inklusion

**Antonia Bartels, Andrea Hopf, Katrin Knoll, Reinhilde Stöppler, Jochen Wissinger**  
Justus-Liebig-Universität Gießen

Der Ausbau inklusiver Beschulung gestaltet sich in den 16 Bundesländern Deutschlands unterschiedlich, doch gilt gemeinhin als langfristige Perspektive, „das bisher stark gegliederte und von seiner Ausrichtung her separierende deutsche Schulsystem in ein inklusives System zu überführen“ (Döttinger und Hollenbach-Biele 2015, S.317). Eine Aus- oder Weiterbildung von Lehrkräften, die die Komplexität wachsender Heterogenität berücksichtigt, ist Grundlage für professionelles Handeln in einem inklusiven Schulsetting. Darauf können individuelle Voraussetzungen für eine gelingende Realisierung, nämlich eine positive Einstellung zu Inklusion sowie die Zuversicht, „den Anforderungen von heterogenen Klassensituationen gerecht zu werden“ (Koop 2009, S.8), aufbauen. Diese motivational wirksame Kognition bezeichnet man als Selbstwirksamkeit. Sie und die Einstellung als eine Form der Werthaltung sind in Anlehnung an das Modell professioneller Handlungskompetenz von Lehrkräften (vgl. Baumert und Kunter 2006) wichtige Pfeiler der Entwicklung professioneller Kompetenz bezogen auf Inklusion.

Zur Implementierung inklusionsrelevanter Inhalte in die Lehrer/innenbildung empfiehlt die KMK (2015) eine curriculare Vernetzung der beteiligten Disziplinen und verschiedener Komponenten des Studiums. Schulpraktische Studien gewähren Lernmöglichkeiten, „die reflexive Auseinandersetzung mit dem professionellen Selbstkonzept angehender Lehrkräfte hinsichtlich der eigenen Haltung und Einstellung zum Thema schulischer Inklusion“ (ebd., S.4) zu fördern. Das in Bezug auf verschiedene Studieninhalte leisten zu können, ist eine verbreitete Erwartung an Portfolioarbeit (vgl. Wagner 2011), weshalb Portfolios häufig als „Instrument der Professionalisierung“ (Koch-Priewe 2013, S.47) eingesetzt werden.

Während bereits diskutiert wird, welchen „Beitrag zur Professionalisierung von Regellehramtsstudierenden innerhalb [...] allgemeiner Schulpädagogik und Inklusionspädagogik“ (Veber 2013, S.72) Schulpraktika im Studium leisten können, sind schulpraktische Anteile im Studiengang Lehramt an Förderschulen deutlich seltener im Fokus empirischer Untersuchungen oder konzeptioneller Überlegungen. Insbesondere weil Förderschullehrkräfte an der Umsetzung von Inklusion mit einem differenzierten Tätigkeitsprofil maßgeblich beteiligt sind (vgl. Melzer und Hillenbrand 2013), möchte die aktuelle Studie dieses Desiderat aufgreifen.

Im Rahmen der Erprobung des hessischen Praxissemesters absolvieren Studierende des Lehramts an Förderschulen der Justus-Liebig-Universität Gießen zwei aufeinander folgende Praxisphasen an je einer Förder- und einer Regelschule. Über dieses Praxissemester hinweg erstellen die Studierenden ein begleitendes Portfolio, das ihnen als Reflexionsinstrument dient.

In einem Studiendesign mit Prä-, Interim-, Post- und Follow-up-Befragung wird untersucht, wie die Bearbeitung des Portfolios sowie die studentische Einschätzung der Betreuung an Schule und Universität die Selbstwirksamkeit bezogen auf inklusive Unterrichtsgestaltung und die Einstellung zur Gestaltung und den Effekten inklusiven Unterrichts (KIESEL; vgl. Bosse und Spörer 2014) der Studierenden (N=123) beeinflusst.

Im Sinne des geforderten Einbezugs „lernprozessbezogene[r] Tätigkeiten, die Aufschluss über das implementierte Curriculum geben“ (König et al. 2014, S.6), orientiert sich das eingesetzte Instrument an den durch die Universität formulierten Intentionen des Portfolios. Es erfragt dessen Unterstützungspotential aus studentischer Sicht (z.B. "Die Bearbeitung meines Portfolios hat mich dabei unterstützt, zu Beginn des Praxissemesters persönliche Voraussetzungen für den Lehrberuf zu klären").

Die Faktorenladungen deuten auf zwei Dimensionen hin, die das Potenzial der Portfolioarbeit abbilden: Selbstreflexion sowie eine distanzierte Perspektive auf die Schulpraxis. Während die mentorale Betreuung der Studierenden an einer Förderschule im Zusammenhang mit der Ausprägung der Selbstwirksamkeit bezogen auf inklusive Unterrichtsgestaltung kurz nach der Praxisphase steht, korreliert der wahrgenommene Nutzen der Portfolioarbeit hinsichtlich eines distanzierten Blicks auf die Schulpraxis mit der Einstellung zu Inklusion.

Nachdem der Entwicklungsprozess der Studierenden differenziert dargestellt wird, soll anhand der Ergebnisse die Frage diskutiert werden, wie formalstrukturelle Aspekte einer Praxisphase wie das Verfassen eines Portfolios oder die schulische Betreuung zu gestalten sind, um einen Beitrag zum individuellen Professionalisierungsprozess der Studierenden leisten zu können.

### **Wirkungen von Kompetenzfremdeinschätzungen und deren Rückmeldung durch Praktikumsbetreuende auf die Kompetenzselbsteinschätzungen Lehramtsstudierender im Praxissemester**

**Franz Klingebiel, Marius Mähler, Anne Böhnert, Johannes Osterberg, Hans-Peter Kuhn**  
Universität Kassel

Schulische Praxisphasen universitärer Lehrerbildung verfolgen das Ziel, die Entwicklung professioneller Kompetenzen zu fördern. Professionelle Kompetenzen werden dabei nach Baumert und Kunter (2006; Kunter et al., 2011) umfassend verstanden. Dieser Beitrag fokussiert dabei auf die Facette „Unterrichtsführung“ aus dem Bereich des Professionswissens. Neben der Ermöglichung von relevanten – auch stellvertretenden – Erfahrungen in der Praxisphase spielt für die Kompetenzentwicklung Lehramtsstudierender auch die schulische sowie universitäre Betreuung eine Rolle (Gröschner, 2013; Hascher, 2006; Korthagen, Loughran, & Lunenberg, 2005; Brouwer & Korthagen, 2005). Dabei sind Feedbackgespräche ein wichtiges Instrument der Betreuung: In der Regel wird die Kompetenzentwicklung der Lehramtsstudierenden thematisiert, wobei das Potenzial dieser Feedbackgespräche in einem Abgleich der Perspektiven zwischen der Selbsteinschätzung der Studierenden und der Fremdeinschätzung der Betreuenden liegt (Horstmeyer, Appel, Ulrich, & Hansen, 2014).

Der Beitrag befasst sich mit der Frage, ob sich bei den Fremdeinschätzungen schulischer und universitärer Betreuender in Bezug auf die Kompetenzfacette „Unterrichtsführung“ Zusammenhänge und Einflüsse auf die Selbsteinschätzung zeigen lassen. Wir vermuten, dass diese Zusammenhänge und Einflüsse deutlicher sind, wenn es institutionell verankerte und strukturierte Feedbackgespräche gibt. Dazu vergleichen wir dieselben Studierenden in zwei unterschiedlichen Phasen ihres Praxissemesters: Im ersten Setting werden die Studierenden und ihre universitären Betreuenden anhand eines standardisierten Fragebogens etwa vier Wochen nach Beginn des Praktikums um eine Selbst- bzw. Fremdeinschätzung der Kompetenzfacette „Unterrichtsführung“ gebeten (Klingebiel, Neuber, Krawiec, & Böhnert, 2015). Im zweiten Setting werden die Studierenden am Ende des Praxissemesters (nach etwa fünf Monaten) wiederum um eine Selbsteinschätzung gebeten, die durch die Fremdeinschätzungen der schulischen und der universitären Betreuenden ergänzt wird. Diese drei Einschätzungen bilden die Grundlage für das Feedbackgespräch zwischen den universitären Betreuenden und den Studierenden. Etwa sechs Monate später geben die Studierenden eine weitere Selbsteinschätzung ab.

Zur Beantwortung der Forschungsfrage greifen wir auf die Daten der Evaluation des Praxissemesters in Hessen am Standort Kassel zurück. Im ersten Setting liegen Daten von N=112 Studierenden im Praxissemester vor, die Angaben zu Selbst- und universitären Fremdeinschätzungen enthalten. Im zweiten Setting nutzen wir Daten von N=241 Studierenden im Praxissemester, von denen Selbsteinschätzungen sowie je eine universitäre und eine schulische Fremdeinschätzung vorliegen. Zur Einschätzung der Zusammenhänge und des Einflusses der Fremdeinschätzungen auf die Selbsteinschätzung werden simultane und zeitverschobene lineare Regressionen berechnet.

Erste Ergebnisse zum ersten Setting zeigen, dass die zeitverschobenen Zusammenhänge zwischen Fremd- und Selbsteinschätzung im mittleren Bereich liegen und die simultanen Zusammenhänge übersteigen. Es wird diskutiert, ob es sich bei diesem Ergebnis um einen Sleeper-Effekt handeln könnte.

### **Der Einfluss von Berufskonzepten auf die berufliche Selbstwirksamkeitserwartung der Lehramtsstudierenden vor und nach fünf Wochen Praxisphase**

**Katja Knuth-Herzig, Katja Hartig, Immanuel Ulrich, Holger Horz**  
Goethe-Universität Frankfurt

Berufskonzepte, d.h. Vorstellungen über Berufe und deren antizipiertes Ansehen in der Gesellschaft, bestimmen die Berufswahl häufig mit: Berufe mit antizipiertem besserem Ansehen werden eher angestrebt (Eberhard, Krewerth & Ulrich, 2010). Bezogen auf den Beruf als Lehrkraft, können diese Berufskonzepte je nach Fach oder Fächerkombination noch einmal deutlich unterschiedlich ausgeprägt sein (siehe Kaub, Karbach, Biermann, Friedrich, Bedersdorfer, Spinath & Brünken, 2012). Da sich Berufskonzepte und berufliches Handeln gegenseitig bedingen (Leonard, Beauvais & Scholl, 1999), besteht die Vermutung, dass sich diese auch bereits im Studium auf die berufsbezogene Selbsteinschätzung von Studierenden auswirken. Berufskonzepte könnten einen Einfluss darauf haben, wie Studierende ihre Selbstwirksamkeit innerhalb des zukünftigen Handelns im Beruf erwarten. Die erste Praxiserfahrung im Lehramtsstudium in Form der Schulpraktischen Studien oder des Praxissemesters bringt allerdings zusätzlich stärker am Berufsalltag orientierte Erfahrungswerte mit ein (vgl. Bach, 2013). Vor der Praxisphase könnten somit eher die Berufskonzepte als Referenzrahmen verwendet werden, um die eigene Selbstwirksamkeitserwartung bezogen auf das Lehren zu quantifizieren da noch keine eigenen Erfahrungen vorliegen. Nach den ersten eigenen Praxiserfahrungen sollte sich weniger Einfluss von Berufskonzepten zeigen, da nun Bewertungsreferenzen aus dem eigenen praktischen Handeln und Erleben vorliegen.

Welchen Einfluss haben die ursprünglichen Berufskonzepte von Lehramtsstudierenden unterschiedlicher Fächer auf die Selbstwirksamkeitserwartung vor und nach fünf Wochen in der Praxisphase? Zeigt sich der erwartete Rückgang des Einflusses der Berufskonzepte durch erste eigene Praxiserfahrungen?

Zur Untersuchung dieser beiden Fragen wurden 310 Lehramtsstudierende an der Goethe-Universität Frankfurt mit den Fächern Englisch (N = 110), Mathe (N = 149) oder Sport (N = 51) bzgl. ihres Berufskonzeptes (als Lehrkraft mit ihren Fächern), vor Beginn ihrer (ersten) Praxisphase sowie zu der Einschätzung ihrer beruflichen Selbstwirksamkeitserwartung vor und nach fünf Wochen Praxisphase befragt. Zur Erfassung des Berufskonzeptes wurden die sechs Dimensionen Gebildet, Intelligent, Egoistisch, Anspruchlos, Angesehen und Fleißig verwendet. Die Studierenden sollten jeweils zwischen den Extrempolen auf einer sechsstufigen Likert-Skala (Beispiel: 1 = Ungebildet bis 6 = Gebildet) verorten, was die meisten Leute ihrer Vermutung nach über eine/n LehrerIn mit Ihrer Fächerkombination denken. Als Kontrollvariable wurde einbezogen, ob das Praktikum nach 5 Wochen bereits abgeschlossen war (Schulpraktische Studien) oder noch weitere 10 Wochen andauerte (Praxissemester).

Das Berufsbild der Studierenden für die drei Fächer Englisch, Mathematik, Sport unterschied sich auf vier (Gebildet, Intelligent, Angesehen, Fleißig) der sechs gemessenen Dimensionen signifikant. Post Hoc Tests zeigen, dass sich diese Unterschiede je nach Dimension zwischen unterschiedlichen Fächern finden.

Weitere Analysen ergaben zudem signifikante Korrelationen der Selbstwirksamkeitserwartung vor der Praxisphase mit den Dimensionen „Intelligent“ und „Fleißig“. Nimmt man diese schrittweise in das Regressionsmodell auf, zeigt sich, dass lediglich die Dimension „Intelligent“ signifikant Varianz aufklärt ( $R^2 = .03$ ). Zum Zeitpunkt fünf Wochen im Praktikum korreliert lediglich die Dimension „Intelligent“ signifikant mit dem Berufsbild vor der Praxisphase. Die Regression zeigt hier, dass diese Dimension nicht nur weiterhin signifikant Varianz aufklärt, sondern der Anteil sogar leicht ansteigt ( $R^2 = .04$ ). Je höher Lehramtsstudierende das Berufskonzept bezogen auf ihr Fach vor der Praxisphase auf der Dimension „Intelligent“ bewerten, desto selbstwirksamer schätzen sie sich sowohl vor, als auch nach fünf Wochen Praxisphase ein. Ein Einfluss der Art der Praxisphase (Schulpraktische Studien vs. Praxissemester) hierauf war nicht zu finden.

Die Ergebnisse deuten darauf hin, dass sich die erfassten Dimensionen der Berufskonzepte durch die Erfahrung in der Praxisphase unterschiedlich stabil in ihrem Einfluss auf die berufliche Selbstwirksamkeitserfahrung zeigen. Offen bleibt, ob sich auch das Berufskonzept in fünf Wochen Praxisphase, möglicherweise bezogen auf einzelnen Dimensionen, durch die Erfahrung (mit den Personen) in der Schule verändert oder stabil bleibt. Auch hier wäre der Blick auf Unterschiede zwischen den Fächern interessant.



## F12

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H119

### Lernverläufe im langfristigen Längsschnitt aus der Perspektive von Forschung und Anwendung

Chair(s): **Martin Tomasik** (Universität Zürich)

DiskutantIn(nen): **Elmar Souvignier** (Westfälische Wilhelms-Universität Münster)

Dieses Symposium beschäftigt sich mit der langfristigen längsschnittlichen Messung von Lernverläufen im schulischen Kontext, wobei sowohl eine forschungs- als auch eine anwendungsorientierte Perspektive eingenommen wird. Aus der forschungsorientierten Perspektive geht es zunächst darum, die Vielfalt der individuellen Lernverläufe zu organisieren und durch Prädiktoren auf der individuellen, der schulischen und der regionalen Ebene zu erklären. Aus der anwendungsorientierten Perspektive soll die Nutzung der gemessenen längsschnittlichen Lernverläufe im Rahmen eines formativen Feedbacksystems beleuchtet werden. Neben allfälligen Unterschieden ist beiden Perspektiven gemein, dass sie vor ähnliche methodische Herausforderungen – wie etwa die der vertikalen Skalierung – gestellt sind.

Das Symposium besteht aus zwei eher forschungsorientierten und einem eher anwendungsorientierten Beitrag. Der erste forschungsorientierte Beitrag verfolgt dabei einen typologischen Zugang und identifiziert fünf Verlaufsklassen, die sich durch eine gute inhaltliche Interpretierbarkeit, eine ausreichend grosse Klassengrösse und damit auch praktische Relevanz auszeichnen. Es wird gezeigt, dass sich die Zugehörigkeit zu einer der fünf Verlaufsklassen sowie der Schulerfolg innerhalb dieser Klassen durch individuelle und schulische Variablen vorhersagen lässt.

Der zweite forschungsorientierte Beitrag verwendet dieselbe Stichprobe, fokussiert jedoch auf die Rolle regionaler Variablen auf der Ebene von Schulbezirken. Es zeigt sich, dass die Lernzuwächse nicht nur von Merkmalen der Schülerinnen und Schüler oder der Schule, sondern auch vom weiteren Kontext abhängig sind. In beiden Beiträgen werden – im deutschsprachigen Raum einmalig – die Lernzuwächse über die gesamte obligatorische Schulzeit von der Einschulung bis zum Ende des 9. Schuljahrs betrachtet.

Im anschliessenden, anwendungsorientierten Beitrag werden erste Ergebnisse zum Aufbau eines computergestützten adaptiven Feedbacksystems vorgestellt, das zum Schuljahr 2017/18 einer Population von über 100'000 Schülerinnen und Schülern zur Verfügung gestellt wurde. Auch in diesem Beitrag wird eine langfristige längsschnittliche Perspektive eingenommen, weil das formative Feedback der individuellen Lernzuwächse auf einer gemeinsamen, die gesamte obligatorische Schulzeit umfassenden Skala abgebildet wird.

In allen drei Beiträgen werden die inhaltlichen und methodologischen Herausforderungen angesprochen, welche die langfristige längsschnittliche Messung von Lernverläufen mit sich bringt. Die anschliessende Diskussion beschäftigt sich mit der Frage, wie die Ergebnisse aus Leistungsmessung und Leistungsfeedback wechselseitig nutzbar gemacht werden können und liefert einen Ausblick auf die Möglichkeiten, die sich aus der langfristigen Nachverfolgung von Lernzuwächsen ergeben.

*Beiträge des Symposiums*

### Mischverteilung der Verlaufsmuster von Schulleistungen über die obligatorische Schulzeit

**Martin Tomasik, Urs Moser**  
Universität Zürich

In diesem Vortrag werden die Leistungszuwächse von Schülerinnen und Schülern während der obligatorischen Schulzeit betrachtet. Zwei Fragestellungen stehen im Vordergrund. Erstens wird untersucht, ob sich distinkte Typen von Leistungsverläufen identifizieren lassen und wenn ja, um wie viele Typen es sich handelt. Zweitens wird untersucht, ob sich die Zugehörigkeit zu einem Verlaufstyp durch Variablen auf der Ebene des Individuums, der Herkunftsfamilie und des Schulkontext vorhersagen lässt. Auf der individuellen Ebene wird das Geschlecht, das Alter und die psychometrische Intelligenz bei der Einschulung sowie das Arbeits- und Sozialverhalten in der 1. Klasse untersucht. Zusätzlich wird das Vorwissen in der 1. Klasse berücksichtigt. In Bezug auf den Herkunftskontext wird der höchste Bildungsabschluss der Eltern, die Anzahl der Bücher zu Hause, die zu Hause gesprochene Sprache sowie Freizeitaktivitäten mit den Eltern betrachtet. Auf der schulischen Ebene werden im Sinne von Kompositionseffekten (vgl. Dumont, Neumann, Maaz & Trautwein, 2013) alle eben genannten Variablen auf Schulklassenebene aggregiert untersucht.

Die Datengrundlage zur Untersuchung dieser Fragestellungen stellt die Zürcher Längsschnittstudie, in der der Leistungszuwachs in den Fächern Mathematik und Deutsch in einer Zufallsstichprobe von etwa 2'000 Schülerinnen und Schülern, die im Schuljahr 2003/04 im Kanton Zürich eingeschult worden sind, gemessen wurde. Die Schulleistungen wurden drei, sechs und neun Jahre nach der Einschulung mit standardisierten Tests auf Grundlage des Zürcher Lehrplans erfasst, nach dem Rasch-Modell skaliert und auf einer gemeinsamen metrischen Skala abgebildet, die es erlaubt, die Leistungen der Schülerinnen und Schüler über die gesamte Zeitspanne direkt miteinander vergleichen zu können. Pro Messzeitpunkt und Schulfach wurden WLE-Schätzer nach Warm (1989) generiert und für die weiteren Auswertungen verwendet. Durch die verpflichtende Teilnahme war der Stichprobenausfall sehr gering.

Die Auswertung erfolgte mithilfe von mischverteilten Wachstumskurvenmodellen (vgl. Wang & Bodner, 2007), wobei für die Identifikation der optimalen Anzahl von Typen der AIC, der BIC, die Entropie sowie die Log-Likelihood der einzelnen Modellen miteinander verglichen wurden. Es wurden fünf distinkte Typen von Schulleistungsverläufen identifiziert. Etwa ein Fünftel der Schülerinnen und Schüler konnte dem „durchschnittlichen Verlaufstyp“ zugeordnet werden. Ein weiteres knappes Drittel vom „nutznießendem Typ“ zeigte ein überdurchschnittliches Ausgangsniveau bei einem starken Leistungszuwachs. Der „leistungsstarke Typ“ umfasste etwa 16 Prozent der Stichprobe und zeichnete sich durch ein besonders hohes Ausgangsniveau aus, zeigte aber lediglich einen durchschnittlichen Leistungszuwachs. Der „leistungsschwache Typ“ zeichnete sich durch ein unterdurchschnittliches Ausgangsniveau und einen unterdurchschnittlichen Leistungszuwachs aus und umfasste auch etwa 16 Prozent der Stichprobe. Die restlichen neun Prozent konnten dem „nachhinkenden Typ“ zugeordnet werden, der sowohl beim Ausgangsniveau als auch beim Leistungszuwachs die geringsten Werte aufwies.

Die Vorhersage der Typenzugehörigkeit erfolgte in Mehrebenenmodellen (vgl. Windzio & Teltemann, 2013) und war am besten durch Individualvariablen möglich, wobei je nach Typ unterschiedliche Variablen unterschiedlich relevant waren. So waren beispielsweise die kognitiven Grundfähigkeiten ein zentraler Prädiktor für die Zugehörigkeit zum leistungsstarken Typ, waren aber für die Zugehörigkeit zum nutznießenden Typ weit weniger relevant. Ausserdem zeigte sich, dass die Lehrereinschätzungen zum Arbeits- und Sozialverhalten bereits zu Beginn der 1. Klasse prädiktiv für die Typenzugehörigkeit waren und hier vor allem zwischen dem leistungsstarken und dem leistungsschwachen Typ zu differenzieren vermochten. Bei

den herkunftsbezogenen Variablen war der Bildungsabschluss der Eltern hochgradig relevant, während das gemeinsame Freizeitverhalten nur eine marginale Rolle spielte. Die verschiedenen aggregierten Variablen des schulischen Kontexts, schliesslich, konnten zwar häufig die Zugehörigkeit zu einer Lernverlaufsklasse vorhersagen, büsst aber fast vollständig an Bedeutung ein, wenn man für die jeweilige Variable auf der individuellen Ebene statistisch kontrollierte. Die Ergebnisse deuten insgesamt darauf hin, dass die langfristige Prognostizierbarkeit von Schulleistungen vor allen Dingen durch Variablen auf der Ebene der Schülerinnen und Schüler gegeben ist, soziale Herkunftseffekte eine grosse Rolle spielen und die Kompositionseffekte zu Beginn der Schulzeit vernachlässigt werden können.

### **Langfristige Lernverläufe im Kontext regionaler sozialer Ungleichheiten: Mehrebenenanalysen zur Rolle von Herkunft und Erstsprache**

**Laura Helbling, Martin Tomasik, Urs Moser**  
Universität Zürich

Neben individuellen und sozialen Ressourcen, die sich auf die individuelle Leistungsentwicklung von Schülerinnen und Schülern auswirken, können Lernfortschritte je nach sozialem Kontext, in dem das Lernen stattfindet, variieren (Becker & Schulze, 2013). Hierbei werden in der Literatur verschiedene Mechanismen diskutiert. So wird zum Beispiel von „Ansteckungseffekten“ ausgegangen (Jencks & Mayer, 1990), wobei sich Schülerinnen und Schüler in ihren Lernbemühungen und Bildungsaspirationen im Rahmen nachbarschaftlicher und schulischer Kontexte an den durch die Mehrheit bestimmten Bezugsnormen orientieren. Auch wird angenommen, dass in Klassen mit lernförderlichem Klima, hoher durchschnittlicher Leistungen und wenig disziplinären Problemen, Unterricht auf höherem Anforderungsniveau stattfindet (Coleman, 1968), sowie in bildungsnahen schulischen Einzugsgebieten von elterlicher Seite her höhere Erwartungen an die Unterrichtsqualität gestellt werden (Jencks & Mayer, 1990). Zusätzlich können sich bildungsrelevante nachbarschaftliche Effekte über soziale Netzwerke und eine gemeinschaftliche Sozialisation (Galster, 2012) ergeben. Beispielsweise kann angenommen werden, dass bildungsfernere Eltern in sozial privilegierten Nachbarschaften über sozialen Austausch Zugang zu bildungsrelevanter Information erhalten und sich in ihren Bildungsaspirationen den normativen Standards der bildungsnahen Nachbarschaft angleichen.

Basierend auf einer repräsentativen Zürcher Klassenstichprobe von etwa 2'000 Schülerinnen und Schülern, deren Leistungen im Fach Deutsch und Mathematik im Rahmen standardisierter Leistungstests über die gesamte Primar- und Sekundarschulzeit erhoben wurden, wird in diesem Beitrag untersucht, inwiefern über den soziale Hintergrund und das Vorwissen der Schülerinnen und Schüler hinaus, langfristige Lernzuwächse nach sozialem Kontext der Schulen variieren. Damit untersucht dieser Beitrag primäre Herkunftseffekte (Boudon, 1974) vor dem Hintergrund der sich über die Schullaufbahn verschärfenden sozialen Ungleichheit im Kontext unterschiedlicher sozialer Lernumwelten.

Im Kanton Zürich erfolgt die Zuteilung der Kinder auf Primarschulen innerhalb von Schulgemeinden. Je nach sozio-ökonomischem Status der lokalen Wohnbevölkerung könnte dies zu ungünstigeren schulischen Lernumwelten für sozial benachteiligte Kinder führen, die sich zusätzlich negativ auf ihre Leistungsentwicklung auswirken können. Die „soziale Belastung“ verschiedener Schulgemeinden des Kantons Zürich wird über den Sozialindex erfasst, der über die Arbeitslosenquote, den Ausländeranteil, den Anteil von Einfamilienhäusern sowie der Sesshaftigkeitsquote operationalisiert wird und von praktischer Relevanz ist, weil darüber zusätzliche Personalmittel für sozial benachteiligte Schulgemeinden gesprochen werden.

In einem ersten Schritt werden latente Wachstumskurven für die Leistungsentwicklung im Fach Deutsch und Mathematik geschätzt um hierüber intra-individuelle Verläufe im Lernzuwachs über die Primar- und Sekundarschulzeit darzustellen. Latente Wachstumsanalysen aus der Familie der Strukturgleichungsmodelle erlauben die Modellierung von Leistungsveränderungen über die Zeit basierend auf längsschnittlichen Messreihen (Bollen & Curran, 2006) und sind daher geeignet, um Lernzuwächse basierend auf den zur Verfügung stehenden Rasch-skalierten und über die einzelnen Messzeitpunkte hinweg auf eine einzige, metrische Skala gelinkten (Kolen & Brennan, 2014) Schulleistungsdaten nachzuzeichnen. In einem nächsten Schritt wird versucht inter-individuelle und kontextabhängige Variation in den Lernverläufen über individuelle Hintergrundmerkmale sowie auf Schulgemeindeebene über den Sozialindex zu erklären. Hierfür werden die latenten Wachstumskurven der Leistungsentwicklung in den Kontext von Mehrebenenanalysen gestellt, welche eine Dekomposition von Effekten auf Individual- und Schulgemeindeebene erlauben (Hox, 2010).

Die Resultate zeigen nicht-lineare Leistungszuwächse in den Fächern Deutsch und Mathematik, wobei Lernzuwächse über die Primarschulzeit am grössten sind und auf der Sekundarstufe I abflachen. Der soziale Status des Elternhauses wirkt sich, kontrolliert für unterschiedliche Lernvoraussetzungen durch unterschiedliches Vorwissen beim Schuleinstieg und der Muttersprache, positiv auf die Leistungsergebnisse in der 3. Primarschulklasse der Schülerinnen und Schüler sowie deren darauffolgende Lernzuwächse aus. Somit zeichnet sich in den Ergebnissen ab, dass soziale Ungleichheiten im Kompetenzerwerb über die Schulzeit hinweg verschärft werden. Kontrolliert für individuelle Einflussfaktoren zeigt sich zudem Variation in den mittleren Leistungsergebnissen zwischen den Zürcher Schulgemeinden, die zum Teil über den Sozialindex der Zürcher Schulgemeinden erklärt werden kann.

### **Konzept und Kalibrierung eines adaptiven formativen Feedbacksystems für den Unterricht**

**Stéphanie Berger<sup>1</sup>, Angela Verschoor<sup>2</sup>, Theo Eggen<sup>2</sup>, Urs Moser<sup>1</sup>**  
<sup>1</sup>Universität Zürich, <sup>2</sup>Cito

In diesem Vortrag werden das Konzept sowie erste Ergebnisse zum Aufbau eines adaptiven formativen Feedbacksystems vorgestellt, das zurzeit für Schulen in der Nordwestschweiz für eine Population von mehr als 100'000 Schülerinnen und Schüler entwickelt und seit dem Schuljahr 2017/18 erprobt wird. Ziel dieses Feedbacksystems ist es, Lehrpersonen im Sinne von „data based decision making“ (Van der Kleij, Vermeulen, Schildkamp, & Eggen, 2015) dabei zu unterstützen, den Lernstand ihrer Schülerinnen und Schüler von der dritten bis zur neunten Klasse in den Fächern Mathematik, Deutsch, Englisch und Französisch nach Bedarf wiederholt zu überprüfen. Das Feedback zeigt den Lehrpersonen einerseits auf, wie gut ihre Schülerinnen und Schüler die Ziele des Lehrplans erreichen und in welchen Bereichen sie besondere Stärken oder Schwächen aufweisen. Andererseits können die Lehrpersonen das System nutzen, um den Leistungszuwachs ihrer Schülerinnen und Schüler im Längsschnitt zu verfolgen und damit auch die Wirkung von Fördermassnahmen zu überprüfen. Hattie und Timperley (2007; s. a. Hattie, 2009) konnten zeigen, dass sowohl die Identifikation von Stärken und Schwächen als auch die Reflexion der Effektivität des eigenen Unterrichts durch die Lehrperson einen positiven Effekt auf den Leistungszuwachs der Schülerinnen und Schüler haben.

Die Grundvoraussetzung, um kompetenzorientiertes Feedback zur Verfügung zu stellen und den Leistungszuwachs der individuellen Schülerinnen und Schüler über mehrere Schuljahre hinweg vergleichen zu können, stellt die Entwicklung einer vertikalen Messskala dar. Diese steht im Mittelpunkt des Vortrags, wobei drei Themenkomplexe behandelt werden. Erstens wird aufgezeigt, ob der im Lehrplan beschriebene Kompetenzaufbau innerhalb der verschiedenen Fächer anhand von eindimensionalen, klassenstufenübergreifenden Skalen abgebildet werden kann. Zweitens wird untersucht, ob verschiedene Methoden der vertikalen Skalierung zu unterschiedlichen Ergebnissen führen. Drittens wird auf schulfachspezifische Besonderheiten bei der Skalierung eingegangen.

Als Grundlage für die Skalen wurden Testaufgaben auf der Basis des kompetenzorientierten Lehrplans für die Deutschschweiz („Lehrplan 21“; s. Deutschschweizer Erziehungsdirektoren-Konferenz, 2014) entwickelt. Der Lehrplan definiert innerhalb eines Kompetenzbereichs (z.B. Deutsch Lesen) mehrere Kompetenzen mit Kompetenzstufen aufsteigender Schwierigkeit, die von den Schülerinnen und Schülern kontinuierlich über die obligatorische Schulzeit hinweg erworben werden sollen. Die Beschreibung dieser Kompetenzstufen diente als Referenzrahmen für die Aufgabenentwicklung. Für die Kalibrierung der Testaufgaben wurden im Frühling 2017 Kalibrierungstests mit einer Zufallsstichprobe von rund 5'400 Schülerinnen und Schülern aus sieben Klassenstufen durchgeführt. Den Schülerinnen und Schülern wurden mehrere hundert im Lehrplan verortete Testaufgaben in Form von computerbasierten Tests vorgelegt. Die Aufgaben wurden anhand eines komplexen Multimatrixdesigns auf mehrere klassenspezifische Testversionen verteilt. Die erhobenen Daten wurden einerseits pro Klassenstufe nach dem Rasch-Modell skaliert. Anschliessend wurde anhand der Linkingmethode von Stocking und Lord (1983) für jeden Kompetenzbereich eine vertikale, klassenübergreifende Skala gebildet (Kolen & Brennan, 2014). Andererseits wurden alle Testaufgaben gleichzeitig skaliert.

Erste Ergebnisse der Kalibrierung zeigen je nach Fach- und Kompetenzbereich mittlere bis starke Zusammenhänge zwischen den theoretischen Kompetenzstufen des Lehrplans und den empirischen Schwierigkeitsparametern der Aufgaben. Die beiden untersuchten Methoden der Rasch-Skalierung führten dabei zu vergleichbaren Ergebnissen. Die für die Kalibrierungstests beobachteten Leistungszuwächse waren konsistent mit jenen Verläufen, die in früheren Untersuchungen gefunden wurden (vgl. Moser, Oostlander & Tomasik, 2017), was für die Validität der neuen Skalierung spricht.

Durch diese Kalibrierung wurde die Basis gelegt für die Abbildung von Lernzuwächsen auf einer klassenstufenübergreifenden Skala. Die empirischen Schwierigkeitsparameter der Testaufgaben ermöglichen zudem eine gezielte, auf die aktuellen Fähigkeiten der Schülerinnen und Schüler abgestimmte Auswahl der Aufgaben im Sinne von computerbasiertem adaptivem Testen (z.B. Van der Linden & Glas, 2010). Abschliessend werden mögliche Einschränkungen der Skalierung aufgrund von fachspezifischen Unterschieden diskutiert. Als Ausblick wird aufgezeigt, inwiefern das adaptive formative Feedbacksystem auch als Forschungsinstrument zur Untersuchung von langfristigen, längsschnittlichen Lernverlaufsanalysen im formativen Bereich genutzt werden kann.

## F13

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H118

### Strategien zur Förderung fehlerbasierten Lernens in verschiedenen Domänen: Was wirkt wie und was nicht?

Chair(s): **Martin Klein** (Fachrichtung Bildungswissenschaften, Universität des Saarlandes), **Nicole Heitzmann** (LMU München, Department of Psychology und Klinikum der LMU München, Institut für Didaktik und Ausbildungsforschung in der Medizin)

DiskutantIn(nen): **Fritz Oser** (Université de Fribourg)

Auch in der Arbeitswelt und der Professionalisierungsforschung gewinnen eine positive Fehlerkultur und das damit verbundene Bekenntnis zum offenen Umgang mit Fehlern und damit auch dem Lernen aus Fehlern in den verschiedensten Disziplinen an Bedeutung (Bauer & Mulder, 2008; Frese & Keith, 2015; Harteis, Bauer & Gruber, 2008; Hayes, Jackson, Davidson & Power, 2015). Dieses Symposium befasst sich daher aus verschiedenen Perspektiven und in verschiedenen Domänen mit dem Thema "Lernen aus Fehlern" sowie dem dadurch erworbenen "negativen Wissen" (Oser, 2015). Dabei werden bewusst, ganz im Sinne des Lernens aus Fehlern, auch Studien präsentiert, in denen die Interventionen nicht immer den gewünschten Erfolg erbrachten. Ziel des Symposiums ist ein Überblick über Fördermaßnahmen, die sich zur Optimierung fehlerbasierten Lernens bzw. der Förderung negativen Wissens in verschiedenen Domänen als lernwirksam erwiesen haben, sowie die Erörterung von Erklärungen für fehlende Nachweise der Wirksamkeit theoretisch gut begründeter Interventionen.

Einzelbeiträge:

Hartmann et al. präsentieren eine Studie zum sog. Productive Failure (Kapur, 2014). Dieser Ansatz betont die Wichtigkeit generativer Prozesse beim fehlerbasierten Lernen. Lernende generieren in einer Problemlösephase fehlerhafte Lösungen, bevor in einer Instruktionsphase lösungsrelevantes konzeptuelles Wissen vermittelt wird. Der Vergleich dieses Ansatzes mit dem Nachvollzug fremder Fehler erbrachte bisher Vorteile eigener Fehler. Hartmann et al. reichern daher fremde Fehler mit zusätzlichen Informationen über die Generierung an, um einen direkteren Vergleich der Ansätze mit Fokus auf die Bedeutung der Generierungsprozesse zu ermöglichen.

Klein et al. stellen eine Interventionsstudie zum Lernen aus Fehlern zur Förderung der Diagnosekompetenz bei Medizinstudierenden vor. Erklärungen der Fehler in einer beispielhaften fehlerhaften Diagnose werden mit zusätzlichen abstrakten Informationen über gemeinsame Fehlerursachen angereichert und unterstützen so den Erwerb negativen Wissens insbesondere im Hinblick auf den Transfer. Im Beitrag werden verschiedene Abstraktionsformen auf ihre Lernwirksamkeit untersucht und Implikationen für die Forschung zum fehlerbasierten Lernen sowie unmittelbar ableitbare didaktische Implikationen diskutiert.

Der Beitrag von Heitzmann et al. vergleicht die Effekte von Fehleranalyseprompts auf den Diagnosekompetenzerwerb im Kontext von Lösungsbeispielen mit integrierten Fehlern in zwei Domänen. Zusätzlich wurde der Einfluss von Lernvoraussetzungen und des Diagnoseprozesses domänenübergreifend untersucht. Dabei zeigte sich ein domänenübergreifender negativer Effekt der Fehleranalyseprompts. Unterschiede zwischen den Domänen wurden hinsichtlich des Lernprozesses deutlich. Es wird diskutiert inwiefern Domäne als wichtiger Faktor für die Wirkungsweise von Instruktion in Betracht gezogen werden sollte.

Wagner et al. präsentieren die Ergebnisse einer Studie zum kooperativen Lernen aus Fehlern in der Lehrerbildung mit Fokus auf die Interaktion der Lernenden. Untersucht werden die Effekte verschiedener Strukturierungen der Transaktivität, also des Ausmaßes, in dem sich Lernpartner auf die Beiträge des Anderen beziehen und diese als Ressource nutzen. Es werden die Effekte unterschiedlicher Diskursformen auf den Erwerb verschiedener Wissensarten in der Nutzung bildungswissenschaftlicher Evidenz thematisiert.

Die Diskussion führt mit Prof. Fritz Oser (Université de Fribourg) ein Experte auf dem Gebiet der Fehlerforschung. Die wachsende Zahl von Symposien und Arbeitsgruppen zum Lernen aus Fehlern auf Fachtagungen wie der EARLI 2017 oder der PAEPSY 2017 zeigt die wachsende Bedeutung des Themas. Das hier vorgestellte Symposium mit seinen unterschiedlichen Perspektiven und Domänen soll insbesondere zur Forschung über praxisrelevante Fördermöglichkeiten fehlerbasierten Lernens und negativen Wissens beitragen.

*Beiträge des Symposiums*

### Lernen aus fremden Fehlern? Eine experimentelle Studie zur Rolle der eigenen Lösungsgenerierung im Productive Failure Ansatz

**Christian Hartmann<sup>1</sup>, Nikol Rummel<sup>1</sup>, Tamara Van Gog<sup>2</sup>**  
<sup>1</sup>Ruhr-Universität Bochum, <sup>2</sup>Utrecht University, Niederlande

Der Productive Failure (PF)-Ansatz beinhaltet die Annahme, dass Schülerinnen und Schüler (SuS) mehr konzeptuelles Wissen erwerben, wenn sie zunächst ein Problem bearbeiten, bevor ihnen das dahinterstehende konzeptuelle Wissen vermittelt wird (Kapur, 2010). Die Wirksamkeit des PF-Ansatzes wird durch die eigenständige Lösungsgenerierung während der Problemlösephase erklärt. Es wird angenommen, dass SuS in der Problemlösephase Vorwissen aktivieren, ihre – oftmals falschen – Lösungsversuche analysieren und sich so über eigene Wissenslücken bewusst werden (Loibl, Roll, & Rummel, 2016). Dieses Bewusstsein kann anschließend für die konzeptuelle Wissensvermittlung genutzt werden, indem typische Lösungsversuche mit dem Zielkonzept kontrastiert werden (Loibl, & Rummel, 2014). Kapur (2014a, 2014b) konnte zeigen, dass SuS mehr konzeptuelles Wissen erwarben, wenn sie innerhalb der Problemlösephase eigenständige Lösungsversuche generierten (PF), als wenn sie ausschließlich Lösungsversuche anderer SuS evaluierten (Vicarious Failure, VF), was auf die Notwendigkeit der eigenständigen sowie fehlerhaften Lösungsgenerierung während der PF-Problemlösephase hindeutet. Jedoch hatten die VF-SuS bei Kapur (2014a, 2014b) keinen Zugang zum Prozess der Lösungsgenerierung, sondern erhielten nur die fehlerhaften Lösungen, also das Produkt der Generierung. Die vorliegende Studie untersucht daher die Effekte einer VF-Bedingung, in der auch die Generierungsprozesse der fehlerhaften Lösungsansätze anderer SuS zur Verfügung stehen, um mehr über die Notwendigkeit der eigenständigen Lösungsgenerierung während der Problemlösephase zu erfahren. Es wird vermutet, dass eine eigenständige Lösungsgenerierung nicht erforderlich ist, wenn sich SuS mit dem Prozess der Lösungsgenerierung anderer auseinandersetzen können.

Der vorliegende Beitrag umfasst ein Feldexperiment mit N = 180 SuS (Alter = 16.06; SD = .77). Alle Probanden bearbeiteten zunächst ein für sie neues mathematisches Problem. Im Anschluss folgte die Wissensvermittlung durch Instruktion. Die experimentellen Bedingungen unterschieden sich hinsichtlich der Problemlösephase: SuS in der PF-Bedingung generierten

eigenständig Lösungsversuche. SuS in der VF-Produkt-Bedingung bekamen Lösungsversuche anderer SuS vorgelegt. SuS in der VF-Prozess-Bedingung erhielten die gleichen Lösungen, bekamen jedoch zusätzlich den Generierungsprozess in Form von Video- und Audiodaten vorgespielt. Nach der Instruktionsphase wurde der konzeptuelle Wissenserwerb erfasst. Es wurde angenommen, dass sich die SuS der VF-Prozess-Bedingung hinsichtlich des konzeptuellen Wissenserwerbs nicht von SuS der PF-Bedingung unterscheiden. Für die VF-Produkt-Bedingung wurde erwartet, dass SuS in dieser Bedingung weniger konzeptuelles Wissen erwerben, als SuS in den beiden anderen Bedingungen.

Um die Äquivalenzhypothese zu testen (VF-Prozess = PF), wurde u.a. ein Äquivalenztest durchgeführt, mit dem geprüft wurde, ob das Konfidenzintervall der Mittelwertdifferenzen beider Bedingungen innerhalb eines zuvor festgelegten Äquivalenzbereichs liegt (siehe Lakens, 2017). Der Äquivalenzbereich wurde auf Grundlage von Studien festgelegt, die den PF-Effekt für dasselbe Problem replizieren konnten (Kapur, 2010; Kapur, & Bielaczyc, 2011; Loibl, & Rummel, 2014a; Loibl, & Rummel, 2014b; Kapur, 2014a, 2014b). Durchschnittlich zeigten die Studien einen starken Effekt ( $d = 1.02$ ; 95% CI [.54, 1.5]) zugunsten von Bedingungen mit einer generativen Problemlösephase (PF). Auf Basis der unteren Grenze des Konfidenzintervalls wurde angenommen, dass Bedingungsunterschiede kleiner als  $d = .54$  nicht als äquivalent zum empirisch aufgezeigten PF-Effekt gelten.

Ein der Äquivalenzanalyse vorgeschalteter T-Tests zeigte, dass SuS der VF-Prozess-Bedingung ( $M = 2.99$ ;  $SD = 1.48$ ) mehr konzeptuelles Wissen erwarben, als SuS der PF-Bedingung ( $M = 2.22$ ;  $SD = 1.32$ ;  $t(113) = -2.93$ ,  $p = .004$ ;  $d = .546$ ). Folglich bildete der Äquivalenztest zwischen beiden Bedingungen keine Signifikanz ab ( $t(113) = -.039$ ,  $p = .516$ ). Der Äquivalenztest zwischen PF und VF-Produkt ( $M = 2.57$ ;  $SD = 1.80$ ) wurde signifikant ( $t(120) = 1.75$ ,  $p = .04$ ). Der Unterschied zwischen VF-Prozess und VF-Produkt war nicht signifikant ( $t(120) = 1.54$ ,  $p = .062$ ).

Die Befunde zeigen, dass die eigenständige Lösungsgenerierung während der Problemlösephase den PF-Effekt nicht erklären kann. Die Überlegenheit der VF-Prozess- gegenüber der PF-Bedingung lässt vielmehr vermuten, dass Prozesse der aktiven Auseinandersetzung mit (fehlerhaften) Lösungsversuchen die anschließende Wissensvermittlung vorbereiten.

### **Fehlererklärungs-prompts bei Lösungsbeispielen mit integrierten Fehlern: ein Domänenvergleich**

**Nicole Heitzmann<sup>1</sup>, Frank Fischer<sup>2</sup>, Martin R. Fischer<sup>3</sup>**

<sup>1</sup>LMU München, Department of Psychology und Klinikum der LMU München, Institut für Didaktik und Ausbildungsforschung in der Medizin, <sup>2</sup>LMU München, Department of Psychology, <sup>3</sup>Klinikum der LMU München, Institut für Didaktik und Ausbildungsforschung in der Medizin

Selbsterklärungs-prompts gelten als wertvolle instruktionale Unterstützungsmethode beim Lernen aus Fehlern (Wylie & Chi, 2014). Selbsterklärungs-prompts haben die Funktion, die Aufmerksamkeit des Lerners auf spezifische Aspekte des Lernmaterials wie beispielsweise auf integrierte Fehler zu lenken (Quintana et al., 2004). Selbsterklärungs-prompts können Lerner beim Erkennen von generellen Muster unterstützen (Williams, Lombrozo, & Rehder, 2013) jedoch besteht die Gefahr, dass dadurch weniger Details beachtet werden (Legare & Lombrozo, 2014). Die Befundlage zum Einsatz von Selbsterklärungs-prompts ist uneindeutig. So gab es positive Effekte bei weniger komplexen Aufgaben (Chi, et al., 1989; Chi et al., 1994) und auch bei komplexen Aufgaben (Schworm & Renkl, 2007; Chamberland et al., 2011, 2013). Es zeigten sich jedoch auch ausbleibende Effekte (Heijltjes et al., 2015; Mwangi & Sweller, 1998; Wylie et al., 2009) und negative Effekte (Kuhn & Katz, 2009).

Drei möglichen Erklärungen für diese uneindeutigen Befunde soll im Folgenden nachgegangen werden. (1) Die Domäne in der die Selbstklärungs-prompts eingesetzt wurden. (2) Charakteristika des Lernens wie beispielsweise Vorwissen oder metakognitive Kompetenz. (3) Selbsterklärungs-prompts könnten den Lernprozess ungünstig beeinflussen indem sie Cognitive Load erhöhen oder Motivation verringern. Aktuell gibt es nicht genügend ähnliche Studien um diese Fragen metaanalytisch zu beantworten. Darum schlagen wir Cross-Domain Validity Replikationen (CDV) Studien vor. CDV Studien haben das Ziel unnötige Varianz zu verringern indem möglichst vergleichbare Materialien in unterschiedlichen Domänen eingesetzt werden. Zusätzlich sollten für Lernvoraussetzungen und Lernprozesse möglichst gleiche Instrumente verwendet werden.

In zwei Domänen (Lehramtsausbildung und Pflegeausbildung) wurde untersucht, ob Lerner mit Selbsterklärungs-prompts im Kontext von Lösungsbeispielen mit integrierten Fehlern mehr Diagnosekompetenz erwarben und inwiefern sich Unterschiede zwischen den Domänen, den Lernvoraussetzungen und dem Lernprozess auf die Wirksamkeit von Selbsterklärungs-prompts auswirken.

In beiden Studien bearbeiteten Lerner Lösungsbeispiele mit integrierten Fehlern in einer Online-Lernumgebung. Bei den Lösungsbeispielen wurden die Teilnehmer aufgefordert sich in die Rolle eines fiktiven Praktikanten in der Schule bzw. im Krankenhaus zu versetzen und anschließend schrittweise eine Diagnose zu stellen. Der fiktive Praktikant macht während jedes Diagnoseschritts Fehler. In der Bedingung mit Selbsterklärungs-prompts wurden die Lernenden dazu aufgefordert diese Fehler zu analysieren. Diagnosekompetenz wurde mittels unterschiedlicher Tests (MC-Test, Key-Feature- und Problemlöseaufgaben) erfasst. An Studie 1 nahmen  $N=108$  Studierende des Lehramts und an Studie 2  $N=152$  Krankenpflegeschüler teil.

In beiden Studien zeigte sich ein negativer Effekt der Selbsterklärungs-prompts auf den Diagnosekompetenzerwerb (Pflegeausbildung:  $F(1,148)=17.44$ ,  $p<.01$ , partial  $\eta^2=.11$ ; Lehramtsausbildung:  $F(1,104) =5.04$ ,  $p <.05$ , partial  $\eta^2=.05$ ). Lokale Moderationsanalysen mit der Johnson-Newman Technique (Johnson & Neyman, 1936) zeigten, dass der negative Effekt der Selbsterklärungs-prompts stärker war für Lerner mit mittlerem bis hohem Vorwissen, metakognitiver Kompetenz und Interesse. Bei Motivation und Cognitive Load zeigten sich Unterschiede zwischen den beiden Domänen. Während sich in der Pflegeausbildung der negative Effekt der Selbsterklärungs-prompts für einen breiten Bereich von Motivation und Cognitive Load zeigte, war der Bereich in der Lehramtsausbildung für Motivation schmaler bzw. für Cognitive Load nicht vorhanden.

Insgesamt zeigte sich in beiden Studien ein negativer Effekt der Selbstklärungs-prompts auf den Diagnosekompetenzerwerb. Möglicherweise lenken die Prompts die Aufmerksamkeit der Lernenden auf die Fehler und somit vom Gesamtprozess ab. Bezüglich der Lernvoraussetzungen zeigte sich, dass insbesondere Lerner mit guten Lernvoraussetzungen von den Selbsterklärungs-prompts behindert wurden. Dies steht im Widerspruch zu Große und Renkl (2007). Hier war ausreichend Vorwissen nötig um von Selbsterklärungs-prompts zu profitieren. Ein möglicherweise bedeutsamer Unterschied zu der Studie von Große und Renkl (2007) ist jedoch, dass die Lernaufgabe dort innerhalb der gut definierten Domäne Wahrscheinlichkeitsrechnung angesiedelt war. Währenddessen die beiden CDV Replikationen im schlecht strukturieren Bereich des Diagnostizierens stattfand. In zukünftigen Studien sollten auch gut definierte Domänen einbezogen werden. In der Pflegeausbildung war der Lernprozess von größerer Bedeutung. Die zugrundeliegenden Mechanismen für die negativen Effekte der Selbstklärungs-prompts sind möglicherweise domänenabhängig.

## **Förderung der Diagnosekompetenz durch Lernen aus Fehlern: Effekte von Abstraktion in Fehlererklärungen auf die Transferleistung bei Medizinstudenten**

**Martin Klein<sup>1</sup>, Robin Stark<sup>1</sup>, Bärbel Otto<sup>2</sup>, Martin R. Fischer<sup>2</sup>**

<sup>1</sup>Fachrichtung Bildungswissenschaften, Universität des Saarlandes, <sup>2</sup>Klinikum der LMU München, Institut für Didaktik und Ausbildungsforschung in der Medizin

Trotz technischer Fortschritte und zahlreicher Ansätze zur Verbesserung des Diagnoseprozesses sind 10-15% aller klinischen Diagnosen fehlerhaft (Berner & Graber, 2008; Croskerry, 2013). Bis zu 80% dieser diagnostischen Fehler entstehen durch kognitive Fehlleistungen (Thammasitboon & Cutrer, 2013; Lawson & Daniel, 2009), die sich der Vorbeugung durch Regeln und standardisierte Prozeduren entziehen. Die vorliegende Studie setzt daher derartige Fehler als Lerngelegenheiten ein, indem häufige Fehler im Diagnoseprozess dargestellt, erklärt und mit korrekten Lösungen kontrastiert werden (vgl. advokatorisches Lernen aus Fehlern; Oser, 2015). Dadurch erworbenes negatives Wissen über fehlerhafte Entscheidungen, Handlungen oder Strategien bietet eine wichtige Schutzfunktion gegenüber zukünftigen Fehlern (ebd.). Die Anreicherung des negativen Wissens mit einer zusätzlichen Abstraktionsebene in der Erklärung der Fehler in dieser Studie sollte die Transferleistung Medizinstudierender in Tests zur Diagnosekompetenz verbessern.

Positive Effekte von Abstraktion auf Transfer sind in der Pädagogik und Medizindidaktik gut belegt (Bransford & Schwartz, 1999; Kulasegaram et al., 2017). Schalk, Saalbach und Stern (2016) unterscheiden drei Formen der Abstraktion: Die aktive Konstruktion abstrakten Wissens durch Lernende anhand des Vergleichs multipler Beispielfälle, die reine Präsentation eines abstrakten Prinzips oder abstrakte Information mit Fallbezug (Instanzierung). Bisher zeigt sich jedoch in Studien kein klarer Vorteil einer Methode, zudem wurden sie noch nicht im Zusammenhang mit fehlerbasiertem Lernen untersucht (ebd.). Daher wurde in dieser Studie systematisch untersucht, inwiefern unterschiedliche Umsetzungen des Abstraktionsprinzips beim fehlerbasierten Lernen die Diagnosekompetenz Medizinstudierender und insbesondere die Transferleistung fördern.

N=91 Medizinstudierenden wurden in einer Online-Lernumgebung 16 klinische Fallbeispiele zur arteriellen Hypertonie in Form ausgearbeiteter Lösungsbeispiele präsentiert. Jeder Fall beinhaltete eine Patientenbeschreibung, eine fehlerhafte Diagnose durch einen Famulanten und eine Expertenrückmeldung durch einen erfahrenen Arzt, der den Fehler erklärt und eine korrekte Diagnose anbietet. In der ungestützten Beispielbedingung (BB; n = 28), die als Kontrollgruppe fungierte, obliegt die Abstraktion den Lernenden; Fehler werden nur mit konkretem Fallbezug ohne Beziehungen oder Parallelen zu anderen Fällen erklärt. In der Abstraktionsbedingung (AB; n = 30) werden zusätzlich abstrakte Fehlerdefinitionen einschließlich gemeinsamer kognitiver Fehlerursachen oberflächlich unterschiedlicher Fehler über Fälle hinweg präsentiert. Dadurch wurden die Lernenden bei der Erkennung des abstrakten Fehlerprinzips im konkret beschriebenen Fehler unterstützt. In der Instanzierungs-Bedingung (IB; n = 33) werden diese abstrakten Fehlerdefinitionen mit Instanzierungshilfen angereichert, die verdeutlichen, wie sich umgekehrt das abstrakte Prinzip im konkreten Fall widerspiegelt.

Diagnosekompetenz wurde als Konglomerat aus konzeptuellem biomedizinischem Wissen (KBW), strategischem und konditionalem diagnostischem Wissen sowie negativem Wissen über mögliche diagnostische Fehler und ihre Ursachen erfasst. Alle Tests mit Ausnahme von KBW unterschieden zwischen nahem (Thema: arterielle Hypertonie) und weitem Transfer (Thema: Hyperthyreose). Das Vorwissen wurde in allen statistischen Analysen als Kovariate miteinbezogen.

Erwartungsgemäß zeigte sich kein Effekt der abstrakten Fehlerdefinitionen auf KBW ( $F(2,87) = 1.45$ ;  $p = .240$ ,  $\text{partial } \eta^2 = .032$ ). Beim nahen Transfer trat kein globaler Effekt auf ( $F(2,87) = 2.83$ ;  $p = .065$ ,  $\text{partial } \eta^2 = .061$ ), jedoch ließen direkte Kontraste einen Vorteil der IB gegenüber der BB erkennen ( $t(87) = 5.52$ ;  $p = .023$ ,  $d = .62$ ). Beim weiten Transfer hingegen zeigte sich ein mittlerer Effekt zugunsten der Experimentalbedingungen ( $F(2,87) = 7.00$ ;  $p = .002$ ,  $\text{partial } \eta^2 = .139$ ), hier wurde die BB durch beide unterstützten Bedingungen signifikant übertroffen (AB:  $t(87) = -4.72$ ;  $p = .011$ ;  $d = .51$ ; IB:  $t(87) = 6.51$ ;  $p < .001$ ;  $d = .99$ ).

Die Ergebnisse der vorliegenden Studie belegen positive Effekte der Abstraktion auf nahen und weiten Transfer beim Lernen aus Fehlern und tragen zum Verständnis des fehlerbasierten Lernens und dessen Förderung in der Medizindidaktik bei. Aus einer didaktischen Perspektive lassen sich unmittelbare praktische pädagogische Implikationen für die Gestaltung fehlerbasierter klinischer Lernfälle ableiten: Fehlerhafte Fallbeispiele sollten Lernenden nicht nur Erklärungen zu konkreten Fehlern liefern, sondern auch zusätzliche, abstrakte Informationen, die Lernenden dabei unterstützen, konzeptuelle Strukturmerkmale von Fehlern über multiple Kontexte hinweg zu identifizieren und zu verstehen.

## **Kooperatives Lernen aus Fehlern im Lehramtsstudium: Effekte transaktiver Diskurse auf den Erwerb anwendbaren bildungswissenschaftlichen Wissens, den subjektiven Lernerfolg und die Akzeptanz der Lerneinheit**

**Kai Wagner<sup>1</sup>, Maria Bergner<sup>2</sup>, Ulrike-Marie Krause<sup>2</sup>, Robin Stark<sup>1</sup>**

<sup>1</sup>Fachrichtung Bildungswissenschaften, Universität des Saarlandes, <sup>2</sup>Institut für Pädagogik Fachrichtung Bildungs- und Sozialwissenschaften Carl von Ossietzky Universität Oldenburg

Die Anwendung (bildungs-)wissenschaftlicher Theorien zum evidenzorientierten Erklären pädagogischer Problemsituationen und damit die wissenschaftliche Orientierung der Reflexion pädagogischer Prozesse sind Schlüsselqualifikationen angehender Lehrkräfte (KMK, 2014; Meier, 2006). Die vorliegende Studie untersucht, wie hierzu notwendiges Wissen anhand kooperativen fehlerbasierten Lernens im Kontext einer integrierten Lerneinheit (Reinmann & Mandl, 2006) gefördert werden kann.

Kooperatives fehlerbasiertes Lernen hat sich hinsichtlich des Erwerbs anwendbaren Wissens als lernwirksam erwiesen (Wagner, Klein, Klopp & Stark, 2016). Beim kooperativen Lernen spiegelt die Transaktivität das Maß der Interdependenz der Beiträge von Lernpartnern im Diskurs wieder (z.B. Stegmann, Weinberger & Fischer, 2011). Integrierte Lerneinheiten, die Elemente konstruktivistischer und kognitivistischer Lerntheorien systematisch verbinden, haben sich bei der Förderung der beschriebenen Schlüsselqualifikationen bewährt (z.B. Wagner, Klein, Klopp & Stark, 2014). Hier sind motivational relevante Aspekte wie der subjektive Lernerfolg und die Akzeptanz der Lerneinheit von Bedeutung (Stark, Herzmann & Krause, 2010).

Es wurde eine integrierte Lerneinheit konzipiert, die kooperatives Lernen in einen Ansatz fehlerbasierten Lernens einbindet. Bei der Bearbeitung authentischer Fallvignetten wurden Phasen des Lernens aus advokatorischen, von einem Modell stellvertretend für die Lernenden exemplifizierten Fehlern (Phase A) und eigenen Fehlern der Lernenden (Phase B) sequenziert. Anhand der Kontrastierung von Richtigem mit Falschem sollte hier negatives Wissen (z.B. „Wie funktioniert etwas nicht?“) erworben werden, welches die Auftretenswahrscheinlichkeit zukünftiger Fehler mindert (z. B. Oser, 2015). Die Bearbeitung erfolgte dyadisch.

Zentral war die Frage, inwieweit die Transaktivität der Diskurse in den Dyaden den individuellen Wissenserwerb hinsichtlich wissenschaftlichen Erklärens schulischer Probleme beeinflusst. Darüber hinaus waren Effekte auf den subjektiven Lernerfolg und die Akzeptanz der Lerneinheit von Interesse.

Im Rahmen einer experimentellen Interventionsstudie ( $n = 98$  Lehramtsstudierende, 69w) wurden in einem 1x3-faktoriellen Design mit Messwiederholung drei Arten der Transaktivität dyadischer Diskurse induziert: schnelle Konsensbildung sowie konflikt- und integrationsorientierte Konsensbildung. Die Induktion der Transaktivität wurde anhand dreier Kooperationskripts in Phase B variiert. Die Studie erstreckte sich über vier Termine bildungswissenschaftlicher Seminare (je 90 Minuten).

Erhoben wurden konzeptuelles Wissen bezüglich Unterschieden alltäglicher und evidenzorientierter Erklärungen sowie Generierungs- und Analysewissen, das zum Erstellen evidenzorientierter Erklärungen notwendig ist. Der subjektive Lernerfolg und die Akzeptanz der Lerneinheit wurden mittels Ratingskalen erfasst.

Hinsichtlich der Förderung konzeptuellen Erklärungswissens war die schnelle Konsensbildung lernwirksamer als die integrationsorientierte ( $F(1,94) = 9.06$ ,  $p = .003$ ,  $\text{partial } \eta^2 = .09$ ). Der Vergleich mit konfliktorientierter Konsensbildung war nicht signifikant ( $F(1,94) = 1.78$ ,  $p = .19$ ). Bezüglich des Generierungswissens war die konfliktorientierte Konsensbildung effektiver als die schnelle Konsensbildung ( $F(1,94) = 12.3$ ,  $p = .001$ ,  $\text{partial } \eta^2 = .12$ ) und die integrationsorientierte Konsensbildung ( $F(1,94) = 5.54$ ,  $p = .02$ ,  $\text{partial } \eta^2 = .06$ ). Im Analysewissen zeigten sich keine Gruppenunterschiede ( $F(2,98) = 0.21$ ,  $p = .81$ ). Hinsichtlich der subjektiven Dimensionen gab es keine Gruppenunterschiede. Lediglich das Generierungswissen war positiv mit beiden subjektiven Dimensionen assoziiert (subjektiver Lernerfolg und Akzeptanz der Lernumgebung je  $r = .25$ ,  $p < .05$ ), das Analysewissen korrelierte positiv mit dem subjektiven Lernerfolg ( $r = .21$ ,  $p < .05$ ).

Bezüglich des Erwerbs konzeptuellen Erklärungswissens scheint eine wenig ressourcenintensive, organisatorische Kritik (schnelle Konsensbildung) ausreichend zu sein. Positionen der Lernpartner müssen hierzu im Diskurs nicht ausgehandelt werden. Die Generierung evidenzorientierter Erklärungen wird durch kritisches Hinterfragen der eigenen Position bzw. derjenigen des Lernpartners und der damit einhergehenden Kompilierung des hierfür notwendigen Wissens (z.B. Anderson, 1983) eher unterstützt (konfliktorientierte und integrationsorientierte Konsensbildung). Hinsichtlich des Erwerbs von Analysewissen scheint Phase A (s.o.) entscheidender zu sein. D.h. der positive Effekt der Variation in Phase B wirkte sich auf das Analysewissen nicht oder zumindest nicht stark genug aus.

Dieses komplexe Befundmuster bedarf der Replikation, bevor Konsequenzen für die instruktionale Praxis gezogen werden. Anhand der Triangulation quantitativer und qualitativer Daten können in zukünftigen Studien Prozessdaten aus der Lernphase Wirkmechanismen fehlerbasierten dyadischen Lernens weiter erklären.

## F14

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H117

### Auf den Spuren des Bildungserfolgs: Differenzielle Effekte von Persönlichkeitseigenschaften auf Schulleistungen

Chair(s): **Clemens Lechner** (GESIS, Mannheim)

DiskutantIn(nen): **Marion Spengler** (Hector-Institut für empirische Bildungsforschung)

Wer ist in der Schule erfolgreich? In der bildungswissenschaftlichen Forschung hat sich eine Vielzahl an Studien zu Determinanten des Bildungserfolgs vor allem auf kognitive Fähigkeiten, den sozioökonomischen Hintergrund eines Lernenden sowie auf spezifische motivationale Konstrukte (z. B. Wertkomponenten) fokussiert. Einer weit verbreiteten Meinung folgend wird dabei den kognitiven Fähigkeiten ein entscheidender Einfluss auf die akademische Leistungen zugesprochen. Im letzten Jahrzehnt hat sich jedoch zunehmend Evidenz dazu kumuliert, dass auch nicht kognitive Faktoren einen bedeutsamen Einfluss auf den Bildungserfolg haben. Unter dem Sammelbegriff der „nicht-kognitiven Fähigkeiten“ haben Untersuchungen die Rolle verschiedener Persönlichkeitseigenschaften adressiert. Darunter zählen die Big-Five Persönlichkeitseigenschaften zu den am meisten untersuchten Faktoren. Vor allem die Dimensionen Gewissenhaftigkeit, Offenheit und auch Verträglichkeit weisen robuste Zusammenhänge mit Schulleistung auf (sowohl über Noten als auch Testleistungen erfasst). Dies gilt wohlgerne auch über kognitive Fähigkeiten sowie den sozioökonomischen Hintergrund hinaus (inkrementelle Effekte).

Auch wenn die Relevanz nicht-kognitiver Fähigkeiten für den Bildungserfolg bereits als gut fundiert angesehen werden kann, bestehen bedeutsame Forschungslücken. Insbesondere erfassten die meisten bisherigen Studien vor allem globale Leistungsmaße (wie die Durchschnittsnote), was Fragen nach möglichen spezifischen Zusammenhängen der Big-Five Persönlichkeitseigenschaften mit Leistungen in verschiedenen Schulfächern und verschiedenen Leistungserfassungsmaßen unbeantwortet lässt. Des Weiteren adressierte die bisherige Forschung vor allem „einfache Effekte“ der Persönlichkeitseigenschaften auf den Bildungserfolg, wodurch potentiell konditionale (oder „interaktive“) Effekte übersehen werden könnten.

Unser Symposium setzt an dieser Forschungslücke an. Es umfasst vier Studien, die sich mit differentiellen Effekten der Big Five hinsichtlich des Schulfachs, der Schulart, des Alters bzw. des Klassenjahrgangs, des sozialen Hintergrunds, der kognitiven Fähigkeiten und weiteren Faktoren beschäftigen. Die vier Studien decken dabei diverse Bildungskontexte wie Grund- und Sekundarschulen ab und adressieren dabei neben subjektiven Leistungserfassungsmaßen wie Noten auch Zusammenhänge zu standardisierten Leistungsmaßen. Alle beruhen auf kürzlich erhobenen large-scale Datensätzen von hoher Qualität.

Die erste Studie untersucht den Zusammenhang der Big Five mit den Leistungen in Mathematik und Deutsch und beleuchtet dabei differentielle Zusammenhänge sowohl über die verschiedenen Schulfächer als auch über verschiedene Schulformen (Gymnasium, Realschule, Hauptschule). Die zweite Studie betrachtet differentielle Zusammenhänge sowohl über die verschiedenen Schulfächer Mathematik und Englisch als auch über unterschiedliche Leistungsmaße (Noten, Abschlussnote, Leistungstests). Die dritte Studie erweitert den Blickwinkel um die Hinzunahme einer Entwicklungsperspektive, indem Persönlichkeits-Leistungs-Assoziationen in akademischen und nicht-akademischen Schulformen von der Grundschule bis zur Sekundarschule begleitet werden. Die vierte Studie inkludiert ebenfalls die Entwicklungsperspektive von Schülerinnen und Schülern im Verlauf der Sekundarstufe, berücksichtigt jedoch ferner auch die außerschulische Umwelt der Schülerinnen und Schüler, indem die Rolle der Eltern miteinbezogen wird.

Zusammengefasst liefern diese vier Studien robuste Hinweise dafür, dass die Zusammenhänge zwischen Persönlichkeitseigenschaften und Bildungserfolg komplexer sind als bislang angenommen. Sie zeigen, dass differentielle Zusammenhänge (je nach Schulfach, Schulform, oder auch sozialem Hintergrund) eher die Regel als die Ausnahme darstellen. Die Befunde verdeutlichen die Notwendigkeit, über einfache Effekte der Beeinflussung des Bildungserfolgs durch Persönlichkeitsmerkmale hinaus zu denken und diese durch eine differentielle Perspektive zu erweitern. Im Hinblick auf mögliche Interventionen beleuchten die Studien den Bedarf nach einer noch feineren theoretischen und empirischen Fundierung, um die Mechanismen, die der Beeinflussung von Bildungserfolg durch Persönlichkeitseigenschaften unterliegen, besser zu verstehen.

*Beiträge des Symposiums*

### Differenzielle Effekte von kognitiven und nicht-kognitiven Fähigkeiten über Schulfächer und Lernumwelten hinweg

**Naemi Brandt<sup>1</sup>, Clemens Lechner<sup>2</sup>, Julia Tetzner<sup>1</sup>, Beatrice Rammstedt<sup>2</sup>**

<sup>1</sup>DIPF Berlin, Deutschland, <sup>2</sup>GESIS, Mannheim

#### Hintergrund und Ziele

Welche Merkmale unterscheiden einen erfolgreichen von einem weniger erfolgreichen Schüler/in? Traditionell wurde vielfach die Bedeutung kognitiver Fähigkeiten für Lernen und Leistung betont (Deary, Strand, Smith, & Fernandes, 2007; Kuncel, Hezlett, & Ones, 2004; Roth et al., 2015). Dennoch werden akademische Leistungen und Lernen nicht allein durch kognitive Faktoren beeinflusst; auch nicht-kognitive Faktoren spielen eine wichtige Rolle. Neben motivationalen Konstrukten haben sich in jüngerer Forschung vor allem die Big-Five Persönlichkeitseigenschaften als relevante nicht-kognitive Determinanten der akademischen Leistung und des Bildungserfolgs erwiesen. Insbesondere Gewissenhaftigkeit, Offenheit, und Verträglichkeit erwiesen sich als leistungsrelevant (z. B. Lechner, Danner, Rammstedt, 2017; Poropat, 2009).

Doch wie universell ist die Rolle von Persönlichkeitseigenschaften für Lernen und Leistung? Diese Frage lässt die bestehende Forschung weitgehend offen. Es ist jedoch gut bekannt, dass Lernumwelten die akademische Leistung eines Schülers beeinflussen können (Baumert, Stanat, & Watermann, 2006). Unterschiedliche akademische Ziele, didaktische Prinzipien sowie das Verhalten von Lehrkräften und Mitschüler/innen (Weinert, 2001b) formen differentielle Lernumwelten, in denen sich unterschiedliche Möglichkeiten der Einflussnahme kognitiver als auch nicht-kognitiver Faktoren ergeben. In Deutschland stellen beispielsweise Schulformen solche differenziellen Lernumwelten dar (z. B. Baumert et al., 2006). Zudem geben theoretische Überlegungen Anlass zu der Annahme, dass die Einflüsse der Big-Five Persönlichkeitsmerkmale auch je nach untersuchtem Unterrichtsfach variieren, da sich Lernprozesse auch aufgrund unterschiedlicher Fachinhalte unterscheiden. Während beispielsweise in Mathe der Umgang mit Zahlen, Gleichungen oder geometrischen Aufgaben vor allem die Anwendung fester mathematischer Regeln erfordert, bieten die Textanalyse und das Verfassen von Texten in Deutsch auch kreative Anteile.



Dennoch beschäftigen sich bisher kaum Studien mit möglichen differentiellen Effekten kognitiver und nicht-kognitiver Faktoren in verschiedenen Lernumwelten oder verschiedenen Schulfächern.

#### Methode

Das Anliegen der vorliegenden Arbeit war daher, den relativen Einfluss kognitiver (fluider Intelligenz) und nicht-kognitiver (Big Five Traits) auf die Schulleistung getrennt für (1) verschiedene Schulformen (Gymnasium, Realschule, Hauptschule) als differenzielle Lernumwelten sowie (2) für verschiedene Schulfächer (Deutsch versus Mathe) als differenzielle Lerninhalte zu bestimmen. Eine Stichprobe von 12.915 Schüler/innen der 9. Klasse aus dem Nationalen Bildungspanel (NEPS) bildete die Datengrundlage. Die Schüler beantworteten Items eines kurzen Big-Five Inventars sowie zu fluider Intelligenz Items eines Matrizen-tests. Zudem wurden Halbjahresnoten und standardisierte Leistungstests in Deutsch und Mathe erhoben. Latente Mehrgruppenmodelle wurden zur Beantwortung der Hypothesen eingesetzt. Die vier Leistungsindikatoren wurden darin auf die Big-Five Persönlichkeitseigenschaften und fluide Intelligenz regrediert, um einerseits den Gesamtbeitrag (im Sinne von erklärter Varianz,  $R^2$ ) von Persönlichkeitseigenschaften und fluider Intelligenz zur Schulleistung und andererseits ihre spezifischen Effekte in den verschiedenen Schulformen und Schulfächern zu untersuchen.

#### Ergebnisse

Die Ergebnisse zeigten, dass sowohl fluide Intelligenz als auch Persönlichkeitseigenschaften Zusammenhänge zur akademischen Leistung der Schüler/innen aufwiesen. Ihr jeweiliger Anteil als auch das spezifische Muster unterschied sich jedoch deutlich je nach Schulform und Schulfach. Während Persönlichkeitseigenschaften einen inkrementellen Varianzanteil über fluide Intelligenz in der Deutsch- und Mathenote erklärten (und dies vor allem am Gymnasium), erklärten die Persönlichkeitseigenschaften nur einen kleinen (inkrementellen) Anteil in Leistungstests (der jedoch kaum über Schulformen und Fächer variierte).

#### Diskussion

Zusammengefasst verdeutlichen die Ergebnisse, dass sich die Zusammenhänge von fluider Intelligenz und Persönlichkeitseigenschaften mit akademischer Leistung keineswegs einheitlich über Schulformen und Fächer gestalten. Dies gilt insbesondere für Noten. Offenbar lassen unterschiedliche Schulformen und -fächer je unterschiedlichen Raum zur Entfaltung spezifischer kognitiver und nicht-kognitiver Fähigkeiten. Untersuchungen solcher differentieller Effekten bieten die Chance, eine stärker kontextualisierte Perspektive auf die Einflüsse kognitiver und nicht-kognitiver Faktoren auf Lernen und Schulleistung zu entwickeln – und damit womöglich neue Wege für Interventionen und Unterrichtsgestaltung zu bahnen.

### **Persönlichkeit und domänenspezifische Schulleistungen in der Sekundarstufe II: Differenzielle Zusammenhänge mit Noten, Tests und Prüfungen**

**Jennifer Meyer<sup>1</sup>, Johanna Fleckenstein<sup>1</sup>, Jan Retelsdorf<sup>2</sup>, Olaf Köller<sup>1</sup>**

<sup>1</sup>IPN, Kiel, <sup>2</sup>Universität Hamburg

#### Hintergrund und Ziele der Untersuchung

Persönlichkeitseigenschaften als nicht-kognitive Merkmale erklären über kognitive Fähigkeiten hinaus inkrementelle Varianz in Schulleistungen (Dumfart & Neubauer, 2016). Insbesondere die Dimensionen des Fünf-Faktoren-Modells (Costa & McCrae, 1992) Gewissenhaftigkeit und Offenheit haben sich als bedeutsam für die Vorhersage von Schulerfolg erwiesen (Poropat, 2009). Ein Großteil der bisherigen Forschung verwendete über verschiedene Domänen bzw. Schulfächer hinweg aggregierte Leistungsmaße (z. B. Abiturgesamtnote). Bei detaillierterer Betrachtung zeigen sich jedoch domänenspezifische Zusammenhänge (Spengler et al., 2013): Während Offenheit mit Deutschleistung korreliert, aber nicht mit Mathematikleistung, erweist Gewissenhaftigkeit sich als bedeutsam für beide Domänen. Ähnliche Ergebnisse berichten Lüdtko et al. (2004), die Zusammenhänge von Offenheit mit der Englisch-, aber nicht mit der Mathematikleistung finden. Diese Befunde lassen auf einen differenziellen Zusammenhang von Persönlichkeitseigenschaften mit Sprach- im Vergleich zu Mathematikleistung schließen, der möglicherweise durch die unterschiedliche Neigung zu Freizeitaktivitäten, die inzidentelles Sprachlernen fördern, mediiert wird.

Weiterhin zeigen sich differenzielle Befunde bei der Berücksichtigung unterschiedlicher Leistungsmaße: Gewissenhaftigkeit erweist sich als wichtigster Prädiktor für die Vorhersage von Noten, während die Ergebnisse standardisierter Leistungstests besser durch Offenheit vorhergesagt werden können (Spengler et al., 2013). Dennoch sind bisher nur wenige Studien verfügbar, die mit sowohl Domänen als auch Leistungsmaßen beide Aspekte berücksichtigen. An diese Forschungslücke knüpfen wir mit diesem Beitrag an. Wir ergänzen den Forschungsstand zusätzlich um ein weiteres Leistungsmaß, indem wir die Ergebnisse der schriftlichen Abiturprüfungen in beiden Fächern hinzuziehen. Diese Prüfungen beinhalten als schulbasiertes high-stakes Leistungsmaß sowohl Merkmale von Noten als auch standardisierten Tests. Die Berücksichtigung dieser drei unterschiedlichen Leistungsmaße kann zu einem detaillierteren Verständnis des Zusammenhangs von Persönlichkeitseigenschaften und Schulleistungen beitragen, da unterschiedliche Aspekte von Schulleistung abgebildet werden.

Ziel dieser Studie war somit die Untersuchung der prädiktiven Vorhersagekraft von Gewissenhaftigkeit und Offenheit über kognitive Fähigkeiten hinaus, wobei als Leistungsmaße neben Noten auch standardisierte Testergebnisse und Prüfungsleistungen in zwei Fachdomänen (Englisch als Fremdsprache und Mathematik) einbezogen wurden.

#### Methode

Die Stichprobe bestand aus  $N = 3.705$  Gymnasiastinnen und Gymnasiasten (13. Klasse, 52,5 % weiblich,  $M = 19,5$  Jahre) in Schleswig-Holstein. Zur Messung der fünf Persönlichkeitsdimensionen wurde die Kurzversion des Big-Five Inventory (BFI-K; Rammstedt & John, 2005) eingesetzt. Die Englischleistungen wurden mit den Aufgaben des Ländervergleichs (Köller et al., 2010) erfasst, Mathematikleistungen mit den Aufgaben des Nationalen Bildungspanel (NEPS; Neumann et al. 2013). Zur Gewinnung reliabler Leistungswerte verwendeten wir fünf Plausible Values, die zu einer Gesamtschätzung kombiniert wurden (Rubin, 1987).

Die Datenanalyse mittels Strukturgleichungsmodellen erfolgte mit einer latenten multiplen Regression in einem Multi-Gruppen-Ansatz (allgemeines und berufliches Gymnasium; Modellfit: CFI= .995; TLI= .980; RMSEA=.022; SRMR=.019). Kognitive Fähigkeiten wurden mit verbalem und figuralem Untertest des KFT4-12R (Heller & Perleth, 2000) erhoben und als Kovariate zusammen mit Geschlecht und sozioökonomischem Status in das Modell aufgenommen.

#### Ergebnisse

Die Ergebnisse zeigten positive Zusammenhänge von Gewissenhaftigkeit in beiden Domänen mit sowohl Noten (Englisch  $\beta = .16$ ,  $p \leq .001$ ; Mathe  $\beta = .30$ ,  $p \leq .001$ ) als auch Prüfungsergebnissen (Englisch  $\beta = .08$ ,  $p \leq .001$ ; Mathe  $\beta = .24$ ,  $p \leq .001$ ). Des

Weiteren fanden wir einen positiven Zusammenhang von Offenheit mit der Englischnote ( $\beta = .08, p = .039$ ), sowie einen negativen Zusammenhang von Offenheit mit der Mathenote ( $\beta = -.12, p \leq .001$ ) und den Mathe-Prüfungsleistungen ( $\beta = -.13, p \leq .001$ ).

Für Testergebnisse ergaben sich domänenspezifische Zusammenhänge: Offenheit sagte die Testleistung in Englisch vorher ( $\beta = .14, p \leq .001$ ), Gewissenhaftigkeit dagegen die Testleistung in Mathe ( $\beta = .14, p \leq .001$ ). Die prädiktive Vorhersagekraft durch Intelligenz lag zwischen  $\beta = .20$  (Englischnote) und  $\beta = .49$  (Mathe-Testscore).

Diskussion

Die Studie bietet klare Hinweise auf differenzielle Zusammenhänge von Persönlichkeitseigenschaften und Leistungsmaßen in unterschiedlichen Domänen. Weitere Studien sind notwendig, um die zugrundeliegenden Prozesse (z. B. Lern- und Freizeitverhalten) zu untersuchen und praktische Implikationen abzuleiten.

## **Persönlichkeit und Schulleistung im Jugendalter: Differentielle Effekte nach Klassenstufen und Schulformen**

**Julia Tetzner<sup>1</sup>, Michael Becker<sup>2</sup>, Naemi Brandt<sup>1</sup>**  
<sup>1</sup>DIPF Berlin, <sup>2</sup>DIPF Berlin, IPN Kiel

Hintergrund und Ziele der Untersuchung

Die Persönlichkeit von Schülerinnen und Schülern ist einer der wichtigsten Prädiktoren zur Vorhersage ihres Bildungserfolgs. Immer mehr empirische Studien bestätigen einen bedeutsamen Zusammenhang zwischen den Big Five-Persönlichkeitsfaktoren (Extraversion, Neurotizismus, Offenheit, Verträglichkeit, Gewissenhaftigkeit) und unterschiedlichen Leistungsindikatoren. Ein besonders deutlicher Einfluss wird dabei den Faktoren Gewissenhaftigkeit und Offenheit zugeschrieben. Bislang ist jedoch wenig darüber bekannt, ob diese Zusammenhänge über verschiedene Klassenstufen und Altersgruppen hinweg konstant bleiben oder ob sie sich verändern. Unter der Annahme kumulativer Lernzuwächse könnte beispielsweise erwartet werden, dass auch Zusammenhänge zwischen Persönlichkeit und Leistung (vor allem für Gewissenhaftigkeit und Offenheit) über die Zeit ansteigen. Konträr zu dieser Annahme suggerieren bisherige Befunde jedoch abnehmende Zusammenhänge von der Grund- zur Sekundarstufe (Poropat, 2009). Um diesen Aspekt genauer und umfassender zu untersuchen, scheint es zudem notwendig, auch die Aufteilung der Schülerinnen und Schüler in unterschiedliche Sekundarschulformen zu beachten, da sich diese in ihrer akademischen und sozialen Ausrichtung unterscheiden und somit unterschiedliche Möglichkeiten für den Einfluss der Persönlichkeit der Schülerinnen und Schüler bereitstellen (cf. Köller et al., 2003).

Die vorliegende Untersuchung nutzte Daten von Berliner Schülerinnen und Schülern, um die Entwicklung des Zusammenhangs zwischen Persönlichkeit und Schulleistungen von der Primar- zur Sekundarstufe zu untersuchen. Um den potentiellen Einfluss differentieller Sekundarschulformen mit einzubeziehen, wurden dabei Schülerinnen und Schüler, die in der Sekundarstufe ein Gymnasium besuchten, mit jenen, die eine nichtgymnasiale Schulform besuchten, verglichen.

Methode

Es wurden Daten von zwei repräsentativen Zufallsstichproben genutzt:  $N = 3658$  Schülerinnen und Schüler der 6. Klasse in ihrem letzten Grundschuljahr und  $N = 2129$  Schülerinnen und Schüler der 9. Klasse, die unterschiedliche Sekundarschulformen besuchten ( $n = 566$  Gymnasiasten;  $n = 1563$  Nichtgymnasiasten). In beiden Stichproben wurde Persönlichkeit mittels einer Kurzversion des Big Five Inventory (BFI-S; Gerlitz & Schupp, 2005) erfasst. Zur Messung der Schulleistungen wurden sowohl standardisierte Schulleistungstests als auch Schulnoten in Deutsch, Mathematik und der ersten Fremdsprache (Englisch) betrachtet. Mittels Mehrgruppenmodellen wurden Zusammenhänge zwischen Leistungsindikatoren und allen Persönlichkeitsfaktoren – auch unter Kontrolle kognitiver Fähigkeiten – getestet und diese zwischen 6. und 9. Klässlern verglichen.

Ergebnisse

Die Befunde der Untersuchung belegten positive Zusammenhänge zwischen den Schulleistungen der Schülerinnen und Schüler und ihrer Gewissenhaftigkeit, Offenheit, Extraversion und Verträglichkeit. Zudem zeigten sie negative Zusammenhänge zwischen Neurotizismus und Schulleistungen. Gewissenhaftigkeit zeigte dabei stärkere Zusammenhänge zu Schulnoten als zu standardisierten Leistungstests. Diese Zusammenhänge reduzierten sich von der Primar- zur Sekundarstufe für alle Persönlichkeitsfaktoren (mit Ausnahme von Offenheit) in der Gesamtstichprobe, zeigten jedoch differentielle Entwicklungen zwischen Schulformen: Während die Zusammenhänge zwischen Schulleistungen und Offenheit sowie Gewissenhaftigkeit für Gymnasiastinnen und Gymnasiasten anstiegen, zeigten Verträglichkeit und Extraversion ausschließlich in nichtgymnasialen Schulformen positive Zusammenhänge zu Schulleistungen.

Diskussion

Die Ergebnisse dieser Studie unterstreichen einerseits die bedeutsame Rolle der Schülerpersönlichkeit für den Bildungserfolg im Jugendalter. Darüber hinaus zeigen sie, dass sich der Einfluss der Persönlichkeit über die Bildungskarriere hinweg und in Abhängigkeit der Lernumwelt verändert. Die Befunde sprechen daher für die Notwendigkeit, Zusammenhänge zwischen Persönlichkeit und Bildungsergebnissen aus einer entwicklungs sensitiven Perspektive und unter Einbezug kontextueller Bedingungen zu betrachten.

## **Das Zusammenspiel von Persönlichkeits- und Leistungsentwicklung in der Adoleszenz und die Bedeutung des Familienzusammenhalts**

**Anne Israel, Oliver Lüdtke, Jenny Wagner**  
IPN, Kiel

Hintergrund und Ziele der Untersuchung

Die bisherige Forschung konnte einen bedeutsamen Zusammenhang zwischen Persönlichkeit und akademischer Leistung bei jungen Erwachsenen aufzeigen: Eine reifere Persönlichkeit ist assoziiert mit besseren Leistungen. Es gibt jedoch kaum längsschnittliche Studien, die die Beziehung zwischen Persönlichkeits- und Leistungsveränderung während der Adoleszenz näher beleuchten; weiterhin existieren nur wenige Forschungsarbeiten, die mögliche Prädiktoren für diese Veränderungen betrachten. Die Adoleszenz ist eine Entwicklungsphase, die durch biologische Reifungsprozesse (Casey, Jones, & Hare, 2008), Identitätsbildung (Erikson, 1959) und neue soziale Rollenerwartungen (Hogan & Roberts, 2004) gekennzeichnet ist. Dieser Entwicklungsprozess kann als dynamische, kontinuierliche und reziproke Interaktion zwischen Individuen und ihrer

Umwelt verstanden werden (Magnusson, 1990). Dementsprechend ist die Berücksichtigung von möglichen Entwicklungskontexten bei der Betrachtung von Persönlichkeits- und Leistungsentwicklung von Schülerinnen und Schülern entscheidend (Caspi, Roberts, & Shiner, 2005). Die vorliegende Studie widmet sich diesem Thema und legt den Fokus auf zwei Forschungsfragen: (1) Zeigen sich zwischen der Persönlichkeit (Big Five) von Schülerinnen und Schülern und verschiedenen akademischen Leistungsindikatoren (Noten sowie Kompetenztests in Deutsch und Mathematik) reziproke Effekte im Längsschnitt? (2) Kann Familienzusammenhalt, als ein bedeutendes Merkmal des familiären Umfeldes, aus Schüler- und Elternperspektive Persönlichkeits- und Leistungsveränderungen der Schülerinnen und Schüler über einen Zeitraum von zwei Jahren vorhersagen?

#### Methode

Auf Grundlage der längsschnittlichen Daten des Nationalen Bildungspanels (NEPS; N = 4716, Alter MT1 = 12.9 Jahre, 49% weiblich) zu Kompetenzentwicklungen und Bildungsprozessen (Blossfeld, Roßbach, & von Maurice, 2011) wurden mithilfe von cross-lagged Modellen die Daten der fünf Persönlichkeitstrait, der vier unterschiedlichen Leistungsindikatoren und des Familienzusammenhalts aus der Schüler- und Elternperspektive analysiert. Die vorliegende Studie verwendet die Daten der Schülerinnen und Schüler aus der siebten (2012/2013) und neunten (2014/2015) Klasse sowie die Daten der Eltern aus der Erhebung in 2012/2013.

#### Ergebnisse

Die durchgeführten Analysen ergeben zum einen differenzielle reziproke Effekte zwischen Persönlichkeit und akademischer Leistung: Alle Persönlichkeitstrait (bis auf Verträglichkeit) sind mit mindestens einem Leistungsindikator sowie alle Leistungsindikatoren mit den Persönlichkeitstrait zwei Jahre später assoziiert. Gewissenhaftigkeit erweist sich als stärkster Prädiktor und zeigt konsistente Effekte auf alle gemessenen Leistungsindikatoren. Bezüglich der Effekte der Leistung auf die Persönlichkeit sagt vor allem Leistung im Mathematikbereich die Persönlichkeitsveränderung zwei Jahre später vorher. Des Weiteren ist die Veränderung in Offenheit positiv assoziiert mit der Veränderung in Deutschleistungen (Deutschnote und Orthografiekompetenztest), wohingegen die Veränderung in Gewissenhaftigkeit positiv mit Veränderungen in den Noten (Deutsch- und Mathenote) assoziiert ist. Zum anderen zeigt der wahrgenommene Familienzusammenhalt aus Schüler- und Elternsicht in der siebten Klasse differenzielle Ratereffekte auf die Persönlichkeit und Leistung zwei Jahre später. Familienzusammenhalt aus Schülersicht ist assoziiert mit höherer Extraversion, höherer Verträglichkeit, höherer Gewissenhaftigkeit und einer besseren Deutschnote. Familienzusammenhalt aus Elternsicht zeigt hingegen ausschließlich Effekte auf die Persönlichkeitsveränderung der Schülerinnen und Schüler, nicht aber auf die Leistungsveränderungen. So geht ein höherer Familienzusammenhalt aus Elternsicht mit weniger Neurotizismus und ebenfalls einer höheren Verträglichkeit der Schülerinnen und Schüler zwei Jahre später einher.

#### Diskussion

Die vorliegende Studie betrachtet als eine der ersten das dynamische Zusammenspiel von Persönlichkeit und Leistung in der Adoleszenz anhand längsschnittlicher Daten und bezieht in diesem Zusammenhang das Familienumfeld als weiteren wichtigen Entwicklungskontext in die Analysen mit ein. Die Ergebnisse zeigen die Bedeutung von Schule und Familie für die Persönlichkeitsentwicklung sowie die Rolle von Persönlichkeit und Familie für die Leistungsentwicklung auf und legen nahe, dass alle drei Faktoren gemeinsam betrachtet werden sollten, um die Dynamiken von Stabilität und Veränderung von Persönlichkeits- und Leistungsfaktoren in der Adoleszenz besser verstehen zu können.

## F15

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H116

### Three fits all? Zur Notwendigkeit einer Erweiterung der drei Basisdimensionen

Chair(s): **Anna-Katharina Praetorius** (Universität Zürich), **Thilo Kleickmann** (Universität Kiel)

DiskutantIn(nen): **Eckhard Klieme** (DIPF)

In der deutschsprachigen quantitativen Unterrichtsforschung dominiert zur Konzeptualisierung von Unterrichtsqualität das Modell der drei Basisdimensionen: Klassenführung, konstruktive Unterstützung sowie kognitive Aktivierung (zuerst bei Klieme, Schümer & Knoll, 2001). Klassenführung dient durch den präventiven und intervenierenden Umgang mit Unterbrechungen und Disziplinproblemen dazu, Schüler\*innen möglichst viel Zeit zum Lernen zur Verfügung zu stellen. Konstruktive Unterstützung fördert deren Erleben von Autonomie, Kompetenz und sozialer Eingebundenheit. Kognitive Aktivierung bezieht sich auf die Anregung kognitiv anspruchsvoller Prozesse des Problemlösens und des Verstehens. Mittlerweile deuten einige Studien darauf hin, dass die unmittelbar eingängige Dreiteilung nicht die Einordnung aller zentralen Aspekte von Unterrichtsqualität ermöglicht. Besonders auffallend mit Blick auf bisherige Publikationen ist dies für den Bereich der inhaltlichen Strukturierung und Klarheit, der in manchen Studien gar nicht berücksichtigt wird (z.B. Fauth et al., 2014; Gabriel, 2013), während in anderen Veröffentlichungen zugehörige Aspekte der Klassenführung zugeordnet werden (z.B. Klieme et al., 2001; Taut & Rakoczy, 2016) oder aber weitere Dimensionen zusätzlich zu den drei Basisdimensionen ergänzt bzw. diese weiter ausdifferenziert werden, um entsprechende Aspekte abdecken zu können (z.B. Kunter & Voss, 2011). Auch in theoretischen Arbeiten finden sich Forderungen nach einer stärkeren Berücksichtigung eines entsprechenden Faktors (Einsiedler & Hardy, 2010; Kleickmann, 2012). Zudem wird vorgeschlagen, expliziter fachliche Aspekte von Unterrichtsqualität in einem Modell der Unterrichtsqualität aufzunehmen (z.B. fachliche Akkuratheit, siehe Hill et al., 2008). Dass eine solche Ergänzung unter Umständen sinnvoll erscheint, bildet sich auch in Modellen zu Unterrichtsqualität ab, die zur Erfassung von Unterrichtsqualität in spezifischen Fächern entwickelt wurden (z.B. Mathematical Quality of Instruction, Learning Mathematics for Teaching, 2011; Mathematics-Scan, Walkowiak et al. 2014), the Reformed Teaching Observation Protocol, Sawada et al., 2002).

Im Rahmen des Symposiums werden drei Ansätze für eine Ergänzung der drei Basisdimensionen vorgestellt und Befunde zur prädiktiven Validität dieser Ergänzungen im Vergleich zu den Basisdimensionen vorgestellt. Der erste Beitrag schlägt eine Ergänzung um den Bereich der kognitiven Unterstützung durch inhaltliche Strukturierung und Klarheit vor (Kleickmann, Praetorius & Steffensky). Der zweite und dritte Beitrag schlagen jeweils Ergänzungen der Basisdimensionen um explizit fachspezifische Merkmale der Unterrichtsqualität vor (mathematikdidaktische Qualität, siehe Lipowsky, Drollinger-Vetter, Pauli & Reusser; stoffbezogene mathematikdidaktische sowie unterrichtsbezogene mathematikdidaktische Qualität, siehe Kaiser, Schlesinger, Jentsch, Suhl, König & Blömeke). Allen Beiträgen ist gemein, dass die Unterrichtsqualität beobachterbasiert erfasst wurde, in einem der Beiträge (Kleickmann et al.) wurden zudem zusätzlich Einschätzungen von Schüler\*innen eingesetzt. Die Beiträge decken neben Mathematik (Lipowsky et al., Kaiser et al.) auch die Naturwissenschaften ab (Kleickmann et al.) und weisen auch in Bezug auf die einbezogenen Schularten Variabilität auf (Grundschule, Kleickmann et al.; Sekundarschule, Lipowsky et al., Kaiser et al.). Die Beiträge werden unter anderem hinsichtlich der potentiellen Notwendigkeit einer Erweiterung der Basisdimensionen wie auch der Frage der Fachspezifität von Dimensionen der Unterrichtsqualität diskutiert (Klieme).

*Beiträge des Symposiums*

### Qualität naturwissenschaftlichen Unterrichts in der Grundschule: Mehr als drei Basisdimensionen?

**Thilo Kleickmann<sup>1</sup>, Anna-Katharina Praetorius<sup>2</sup>, Mirjam Steffensky<sup>3</sup>**

<sup>1</sup>Universität Kiel, <sup>2</sup>Universität Zürich, <sup>3</sup>IPN

Theoretischer Hintergrund: Unterrichtsqualität wird im deutschsprachigen Raum oft anhand von drei Basisdimensionen beschrieben (kognitive Aktivierung, konstruktive Unterstützung und Klassenführung; Klieme et al., 2001), mittlerweile wird aber kritisiert, dass bestimmte Aspekte nicht hinreichend berücksichtigt werden. Theoretische und empirische Arbeiten verweisen vor dem Hintergrund sozio-konstruktivistischer Lerntheorien und von Ansätzen zum Scaffolding darauf, dass insbesondere kognitive Unterstützung eine weitere wichtige Dimension von Unterrichtsqualität darstellt, wenn ein tiefes fachliches Verständnis bei den Schüler\*innen gefördert werden soll (Hardy et al., 2006; Puntambekar & Hübscher, 2005). Kognitive Unterstützung umfasst inhaltliche Strukturierung (z.B. inhaltliche Kohärenz und Feedback) und Klarheit (z.B. Zielklarheit und verständliche Darstellungsformen) (Einsiedler & Hardy, 2010; van de Pol et al., 2010).

Eine gewisse Berücksichtigung kognitiver Unterstützung findet sich zwar in einigen Studien, die die Basisdimensionen als Rahmenmodell nutzen (Lipowsky et al., 2009; Kunter & Voss, 2011). Doch werden die Aspekte trotz der anzunehmenden unterschiedlichen Wirkmechanismen hier nur gemeinsam mit emotionaler Unterstützung (z.B. wertschätzender Umgang mit den Schüler\*innen) in der Dimension konstruktive Unterstützung erfasst. In anderen Studien werden Aspekte der kognitiven Unterstützung auch in der Dimension Klassenführung oder insgesamt nur randständig erfasst (z.B. Fauth et al., 2014; Taut & Rakoczy, 2016).

Fragestellung: Vor diesem Hintergrund wurde in der vorliegenden Studie kognitive Unterstützung als separate Unterrichtsdimension neben kognitiver Aktivierung, emotionaler Unterstützung und Klassenführung erfasst, und es wurde der Frage nach der prädiktiven Validität dieser vier Dimensionen für das fachliche Verständnis und das unterrichtsbezogene Interesse der Schüler\*innen nachgegangen. Wir erwarteten, dass neben kognitiver Aktivierung und Klassenführung auch kognitive Unterstützung für das fachliche Verständnis der Schüler\*innen prädiktiv ist, während emotionale Unterstützung und Klassenführung für das unterrichtsbezogene Interesse prädiktiv sein sollten.

Methode: Die Fragestellung wurde anhand der Daten von 1.326 Viertklässler\*innen (60 Klassen) im Rahmen einer Unterrichtsreihe zum Thema Aggregatzustände und ihre Übergänge („Wasserkreislauf“) untersucht. Die jeweils erste Doppelstunde wurde videobasiert von drei geschulten Beobachter\*innen im Hinblick auf die o.g. vier Dimensionen anhand eines standardisierten Manuals beurteilt. Die Beobachterübereinstimmung war zufriedenstellend (relativer Generalisierbarkeitskoeffizient > .85).

Die Schüler\*innen beurteilten am Ende der Unterrichtsreihe die Qualität des Unterrichts anhand von Fragebogenskalen zu den o.g. vier Dimensionen. Die Reliabilität der Unterrichtseinschätzungen auf Klassenebene war zufriedenstellend (ICC2 > .80).

Das Verständnis der Schüler\*innen von Aggregatzuständen und ihren Übergängen wurde vor und nach der Unterrichtsreihe mit einem Test erfasst (26 Aufgaben; EAP/PV-Reliabilität = .74 bzw. .82), das unterrichtsbezogene Interesse der Schüler\*innen mit einer Skala aus sechs Items (Cronbachs Alpha = .80 bzw. .82).

Die Daten wurden mit Zwei-Ebenen-Regressionsanalysen analysiert. Auf Ebene 1 wurden das Verständnis von Aggregatzuständen bzw. das Interesse vor dem Unterricht, das Geschlecht, allgemeine kognitive Fähigkeiten sowie der sozio-ökonomische Status als Kovariaten berücksichtigt.

Ergebnisse und Diskussion: Die Ergebnisse zur prädiktiven Validität der Einschätzungen der Schüler\*innen (Ebene 2) spiegelten die theoretischen Erwartungen nur teilweise wider. Kognitive Unterstützung (Beta = .41,  $p < .01$ ) und Klassenführung (Beta = .37,  $p < .05$ ) waren für das naturwissenschaftliche Verständnis der Schüler\*innen prädiktiv. Kognitive Aktivierung (Beta = .69,  $p < .01$ ), kognitive Unterstützung (Beta = .59,  $p < .01$ ) und emotionale Unterstützung (Beta = .77,  $p < .01$ ) waren für das Interesse prädiktiv.

Die Ergebnisse zur prädiktiven Validität der Urteile der Beobachter\*innen zeigten durchweg das erwartete Muster. Kognitive Aktivierung (Beta = .33,  $p < .01$ ), kognitive Unterstützung (Beta = .26,  $p < .05$ ) und Klassenführung (Beta = .30,  $p < .01$ ) waren für das naturwissenschaftliche Verständnis der Schüler\*innen prädiktiv. Emotionale Unterstützung (Beta = .37,  $p < .01$ ) und auch Klassenführung (Beta = .22,  $p < .05$ ) waren für das Interesse prädiktiv.

Es könnte also lohnend sein, kognitive Unterstützung in künftigen Studien separat zu erfassen und die Faktorstruktur mit größeren Stichproben zu untersuchen. Im Beitrag werden weitere Desiderata und Implikationen für die Erfassung von Unterrichtsqualität diskutiert.

### **Kognitive Aktivierung und fachdidaktische Unterrichtsqualität – nicht die gleiche Seite der Medaille**

**Frank Lipowsky<sup>1</sup>, Barbara Drollinger-Vetter<sup>2</sup>, Christine Pauli<sup>3</sup>, Kurt Reusser<sup>4</sup>**  
<sup>1</sup>Universität Kassel, <sup>2</sup>PH Zürich, <sup>3</sup>Universität Freiburg, <sup>4</sup>Universität Zürich

Theoretischer Hintergrund: In der Unterrichtsforschung mehren sich die Stimmen, die eine Erweiterung der sogenannten „Basisdimensionen“ von Unterrichtsqualität vorschlagen, um fachbezogene Merkmale von Unterrichtsqualität differenzierter zu erfassen.

Bezogen auf das Fach Mathematik wird vorgeschlagen, unter anderem auch auf den mathematischen Gehalt, die Relevanz und Verknüpfungen der ausgewählten Inhalte, die fachliche Angemessenheit und Korrektheit zu achten (vgl. zusammenfassend Schlesinger & Jentsch, 2016). Beispielsweise verstehen Hiebert und Grouws (2007) unter „attending to concepts“ die explizite und kohärente Behandlung und Verknüpfung mathematischer Kernideen. Ebenso nimmt die Gruppe um Hill und Ball die Bedeutung von zur Verfügung gestellten fachlichen Lerngelegenheiten auf (Hill et al., 2008). Mit dem MQI-Instrument werden die Dichte und das Niveau der mathematischen Fachsprache erhoben und auch, inwieweit unterschiedliche Repräsentationsformen genutzt und aufeinander bezogen werden. Auch Schoenfeld (2013) akzentuiert in seinem System TRU Math die Auseinandersetzung mit mathematisch relevanten Inhalten. In der deutsch-schweizerischen Studie „Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen“ (Klieme, Pauli & Reusser, 2009) wurden neben den Basisdimensionen von Unterrichtsqualität auch Facetten einer stärker fachbezogenen Unterrichtsqualität erfasst. Ebenso wie die genannten amerikanischen Ansätze wurde hierbei ermittelt, inwieweit die Lernenden Gelegenheit hatten, sich mit inhaltspezifischen Kernelementen (sog. Verstehenselementen) auseinanderzusetzen, wie kohärent und strukturiert diese Elemente im Unterricht behandelt und auf welchem Qualitätsniveau sie repräsentiert wurden (Drollinger-Vetter, 2011; Drollinger-Vetter & Lipowsky, 2006).

Aus konzeptionell-theoretischer Sicht stellt sich die Frage, ob und inwieweit solche fachbezogenen Aspekte in den drei generischen „Basisdimensionen“ bereits berücksichtigt werden oder eine zusätzliche Qualitätsdimension darstellen (vgl. auch Schlesinger & Jentsch, 2016). Diese Frage betrifft insbesondere die kognitive Aktivierung, deren Konzeptualisierung in verschiedenen Studien hinsichtlich der Berücksichtigung fachdidaktischer Aspekte stark variiert.

Fragestellung: Im Beitrag wird der Frage nachgegangen, inwieweit die fachdidaktische Qualität des Mathematikunterrichts und die kognitive Aktivierung zusammenhängen und damit dieselbe Seite der gleichen Medaille repräsentieren oder aber unterschiedliche Dimensionen von Unterrichtsqualität abbilden.

Methode: Diese Fragestellung wird anhand der Stichprobe der o.g. deutsch-schweizerischen Pythagorasstudie untersucht. Deren Stichprobe besteht aus 39 deutschen und schweizerischen Klassen mit ca. 1.000 Schüler/-innen.

Im Rahmen der Studie wurde in Anlehnung an Aebli (1980/1981) der Begriff der sogenannten Verstehenselemente eingeführt. Hierbei handelt es sich um jene Teilinhalte des Konzepts „Satz von Pythagoras“, die im Unterricht behandelt werden sollten, um ein tragfähiges Konzept vom Satz des Pythagoras aufzubauen. Entsprechend wurde die fach- und themenbezogene Unterrichtsqualität im Rahmen der drei Lektionen zum Satz des Pythagoras anhand der drei fachdidaktischen Dimensionen (1) Vorkommen der Verstehenselemente, gebildet über einen Summenscore, (2) Qualität der Repräsentationen (Skala bestehend aus 4 Items;  $\alpha=.73$ ) sowie (3) strukturelle Klarheit (Skala bestehend aus 4 Items;  $\alpha=.88$ ) erfasst. Diese drei Dimensionen erwiesen sich als eng zusammenhängend (von  $r=.72$  bis  $r=.82$ ) und wurden zur Skala „Fachdidaktische Qualität“ zusammengefasst.

Das Ausmaß der „kognitiven Aktivierung“ wurde hochinferent über die Merkmale (1) herausfordernde Aufgaben und Probleme, (2) Exploration des Vorwissens, (3) Exploration der Denkweisen, (4) evolutionärer Umgang mit Schülerbeiträgen und (5) rezeptives Lernverständnis (invertiert) erfasst ( $\alpha=.80$ ).

Um der o.g. Fragestellung nachzugehen, wird zum einen der Zusammenhang der fachdidaktischen Unterrichtsqualität mit der kognitiven Aktivierung und deren Korrelaten (konstruktivistische Lehr- und Lernüberzeugungen, ko-konstruktive Gesprächsqualität) geprüft. Zum anderen wird ermittelt, wie beide Qualitätsdimensionen (fachdidaktische Unterrichtsqualität; kognitive Aktivierung) mit verständnisbezogenen Schülerwahrnehmungen zusammenhängen und welche Effekte sie auf die Schülerleistungen haben.

Ergebnisse: Zunächst zeigt sich, dass die fachdidaktische Unterrichtsqualität weder mit der kognitiven Aktivierung noch mit ihren Korrelaten zusammenhängt. Mit den Schülerwahrnehmungen kovariert die fachdidaktische Unterrichtsqualität stärker als die kognitive Aktivierung. Die fachdidaktische Unterrichtsqualität vermag zudem die mathematischen Leistungen stärker vorherzusagen als die kognitive Aktivierung. Insgesamt stellen kognitive Aktivierung und fachdidaktische Unterrichtsqualität damit zwei verschiedene Seiten einer Medaille dar.

## Welcher Zusammenhang besteht zwischen fachspezifischer Unterrichtsqualität und den Mathematikleistungen Hamburger Schülerinnen und Schüler?

Gabriele Kaiser<sup>1</sup>, Lena Schlesinger<sup>2</sup>, Armin Jentsch<sup>2</sup>, Ute Suhl<sup>3</sup>, Johannes König<sup>4</sup>, Sigrid Blömeke<sup>5</sup>

<sup>1</sup>Universität Hamburg/Australian Catholic University, <sup>2</sup>Universität Hamburg, <sup>3</sup>Humboldt-Universität Berlin/Universität zu Köln, <sup>4</sup>Universität zu Köln, <sup>5</sup>Universität Oslo

Einflüsse des Unterrichtsangebotes, das Lehrpersonen bereitstellen, werden als wichtige Einflussgröße für Leistungsfortschritte von Schülerinnen und Schülern angesehen (z.B. Baumert et al., 2010). Zur Konzeptualisierung und Messung von Lernangeboten wurden in den letzten Jahren sowohl fachübergreifende Ansätze entwickelt, als auch solche, die für den Mathematikunterricht spezifisch sind (u.a. Pianta & Hamre, 2009; Hill et al., 2008; Schoenfeld, 2013). Um eine Vergleichbarkeit dieses Unterrichtsangebots auf fachübergreifender Ebene zu ermöglichen, hat sich im deutschsprachigen Raum in den letzten zehn Jahren der Ansatz von Klieme und anderen (2006) etabliert, wonach Unterrichtsqualität mittels der drei Basisdimensionen Klassenführung, konstruktive Unterstützung der Lernenden und kognitive Aktivierung beschrieben wird. Der Ansatz von Klieme und anderen berücksichtigt den fachlichen Gehalt des Unterrichts (Blum, 2006) jedoch nicht explizit als eigenständige Dimension. Hierunter fallen etwa ein angemessener Umgang mit mathematischer Rigorosität und geeignete Repräsentationsformen (Hill et al., 2008).

In diesem Beitrag wird deshalb der Frage nachgegangen, welcher Zusammenhang zwischen fachspezifischer Unterrichtsqualität und den Mathematikleistungen von Hamburger Schülerinnen und Schülern besteht im Vergleich zu den Basisdimensionen und es werden erste Analyseergebnisse vorgestellt.

Zur Bearbeitung der Fragestellung wurden bei einer Teilstichprobe der Studie TEDS-Unterricht (n = 38) Unterrichtsbeobachtungen im Mathematikunterricht in den Klassenstufen 7 bis 10 an Stadtteilschulen und Gymnasien durchgeführt. TEDS-Unterricht ist ein Nachfolgeprojekt der internationalen Vergleichsuntersuchung Teacher Education and Development Study: Learning to Teach Mathematics (TEDS-M, Blömeke, Kaiser & Lehmann, 2010) und wurde in Hamburg durchgeführt. Die Teilnahme der Lehrpersonen erfolgte auf freiwilliger Basis. Die Unterrichtsbeobachtungen wurden von jeweils zwei umfangreich geschulten Ratern durchgeführt. Die Rater hatten mindestens ein Bachelorstudium für das Lehramt in Mathematik inne. Die Unterrichtsstunden im Rahmen der Studie wurden in vivo mit einem Beobachtungsinstrument bewertet, das unter Adaption etablierter Instrumente für TEDS-Unterricht entwickelt wurde. Dazu wurden in einer systematischen Literaturrecherche (Schlesinger & Jentsch, 2016) und in Anlehnung an Überlegungen zur Natur mathematikdidaktischen Wissens (Buchholtz, Kaiser & Blömeke, 2014) eine stoffdidaktische und eine unterrichtsbezogene Perspektive auf fachspezifische Unterrichtsqualität neben den Basisdimensionen herausgearbeitet. Die erste Dimension, die stoffbezogene mathematikdidaktische Dimension, umfasst dabei Aspekte, die sich auf mathematische Korrektheit, Präzision und Tiefe der Unterrichtsinhalte beziehen. Die zweite unterrichtsbezogene, mathematik-didaktische Dimension beinhaltet demgegenüber Aspekte, die sich auf unterrichtsbezogene Entscheidungen beziehen, wie die Verwendung verschiedener Repräsentationsformen und angemessener Beispiele und Unterrichtsarrangements wie intelligente Übungsphasen.

Innerhalb einer Doppelstunde (90 Minuten) wurde die Unterrichtsqualität nach vier annähernd gleichen Zeitintervallen von beiden Ratern unabhängig eingeschätzt. Die Bewertung des Aspekts erfolgte auf einer vierstufigen Skala (1 = "trifft überhaupt nicht zu" bis 4 = "trifft vollständig zu"). Die Interraterreliabilität lag nach Ausschluss eines Items bei  $.80 < ICC \leq .96$  und kann somit als gut beurteilt werden. Hinweise auf die faktorielle Validität des Instruments konnten durch eine Zwei-Ebenen-Faktorenanalyse mit akzeptablen Fitwerten gefunden werden ( $\chi^2/df = 1.52$ , CFI = .94, RMSEA = .04). Die resultierenden fünf Faktoren weisen dabei latente Interkorrelationen in einem moderaten bis hohen Bereich ( $.39 < r \leq .89$ ) und zufriedenstellende interne Konsistenzen auf (Cronbachs  $\alpha > .67$ ).

Als Maß für die Leistungsfortschritte von Hamburger Schülerinnen und Schülern im Unterrichtsfach Mathematik wurde auf standardisierte Leistungstests („KERMIT“) zurückgegriffen, die in Hamburg jährlich vom Institut für Bildungsmonitoring und Qualitätsentwicklung durchgeführt werden. Die mittels Tests erfassten Rohdaten wurden dabei nach dem 1PL-Modell skaliert. Anschließend wurden Fähigkeitsparameter für alle Schülerinnen und Schüler geschätzt. Durch die Verwendung von Ankeritems ist damit die Beschreibung von Leistungsfortschritten möglich. Dazu wurde der Differenzwert aus dem aktuellen KERMIT-Testergebnis und jenem aus dem Vorjahr gebildet.

Es zeigen sich signifikante und mindestens moderate Korrelationen zwischen dem auf diese Weise gebildeten Differenzwert und den Dimensionen „kognitive Aktivierung“ sowie den fachspezifischen Qualitätsdimensionen, was als Hinweis auf die prädiktive Validität des Instruments verstanden werden kann.

## F16

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H115

### Heterogenität als Rahmenbedingung der Bildung von Jugendlichen mit Flucht- und Migrationshintergrund – Herausforderungen für SchülerInnen und Lehrpersonen

Chair(s): **Elisabeth Maué** (Universität Konstanz, Deutschland), **Stephan Schumann** (Universität Konstanz, Deutschland), **Claudia Diehl** (Universität Konstanz, Deutschland)

DiskutantIn(nen): **Axinja Hachfeld** (Universität Konstanz, Deutschland)

Etwa ein Drittel der seit 2015 nach Deutschland gekommenen Flüchtlinge und Asylsuchenden ist 16 bis 25 Jahre alt (BAMF, 2017), weshalb deren strukturelle Integration in (Aus-)Bildung und Arbeitsmarkt von großer Bedeutung ist (Esser, 2001).

Über die Bildungsbiographien dieser Jugendlichen im Herkunftsland und in Deutschland ist bislang wenig bekannt. Weiterhin ist die Übertragbarkeit bisheriger Forschungsbefunde zur Bildung Jugendlicher mit Migrationshintergrund auf die jüngst zugewanderten Jugendlichen unklar.

Diesem Defizit begegnet das interdisziplinäre Symposium, das eine soziologische und eine bildungswissenschaftliche Perspektive vereint. Vor dem Hintergrund der Annahme einer großen Heterogenität der Bildungsbiographien werden Aspekte und Rahmenbedingungen der Bildung Jugendlicher mit Flucht- und Migrationshintergrund in den Blick genommen. Dabei werden Sichtweisen und Fähigkeiten unterschiedlicher Akteure – SchülerInnen und Lehrpersonen – an allgemeinbildenden und beruflichen Schulen ebenso berücksichtigt wie verschiedene methodische Ansätze.

Der erste Beitrag fokussiert mittels einer Tagebuchstudie über einen Zeitraum von fünf Wochen die Wahrnehmungen geflüchteter Jugendlicher (N=34; 16-22 Jahre) ihrer betrieblichen Praktika u.a. hinsichtlich der Berufsorientierung, des Erlebens am Arbeitsplatz, der Passung zu den Basic Needs gemäß der Selbstbestimmungstheorie der Motivation und des Selbstwirksamkeitserleben. Quantitative wie qualitative Auswertungen zeigen, dass Betriebspraktika die Berufsorientierung unter der Voraussetzung einer Passung zu den Basic Needs fördern kann.

Für die Teilhabe an Bildung, Arbeit und Gesellschaft sind grundlegende Kenntnisse der deutschen Sprache essentiell. Diese hängen von Gelegenheiten der Verwendung der Sprache als Bedingung des Spracherwerbs ab. Dies fokussiert der zweite Beitrag mit der Untersuchung des Einflusses der ethnischen Klassenkomposition auf die Sprachkompetenzen Jugendlicher mit Migrationshintergrund in ihren Herkunftssprachen Türkisch und Russisch sowie der Unterrichtssprache Deutsch. Mehrebenenanalysen mit Daten aus der NEPS-Startkohorte 4 (Klasse 9: N=14981 SchülerInnen; N=1006 Klassen) ergeben bei SchülerInnen mit nicht-deutscher Erstsprache einen Zusammenhang von ethno-lingualer Klassenkomposition mit Leseverstehen (negativ) sowie mit herkunftssprachlichen Kompetenzen (positiv).

Den Effekt der ethnischen Klassenkomposition auf Deutschkompetenzen Jugendlicher mit Migrationshintergrund greift der dritte Beitrag auf und untersucht diese Fragestellung mit Blick auf geflüchtete Jugendliche in beruflichen Schulen (N=675; Klassen: N=60; Alter: 15-22 Jahre). Zudem ist die Messung der Deutsch-Fähigkeiten und deren Kongruenz mit den Fähigkeitsselbstschätzungen der Lernenden von Interesse. Die Deutsch-Fähigkeiten differieren je nach Herkunftsland und stehen in Zusammenhang mit der ethnischen Klassenkomposition sowie mit Fähigkeitsselbstschätzungen. Die Ergebnisse liefern Hinweise auf die Übertragbarkeit von Befunden unterschiedlicher Migrationsgruppen.

Der vierte Beitrag untersucht mittels halbstrukturierter Interviews den Umgang von Lehrpersonen (N=16) mit leistungsbezogener, sprachlicher und kultureller Heterogenität in Klassen mit geflüchteten Jugendlichen. Es zeigt sich, dass alle Lehrpersonen Binnendifferenzierungsmethoden nutzen; dabei vorrangig bezüglich der Aufgabenstellung, Lehrerhilfe und der Sozialform. Allerdings verweisen sie auch auf Grenzen und benennen einen Professionalisierungsbedarf.

Die Einzelbeiträge beleuchten im Zusammenspiel den Umgang mit unterschiedlichen Heterogenitätsdimensionen sowie deren Auswirkungen auf Persönlichkeitsmerkmale und Sprachfähigkeiten Jugendlicher mit Flucht- und Migrationshintergrund. Zudem liefern sie Hinweise zur Frage der Übertragbarkeit der Befunde zur Bildung Jugendlicher mit Migrationshintergrund auf die jüngst Zugewanderten.

Die sprachliche, leistungsbezogene, kulturelle und religiöse Heterogenität stellt die Lehrpersonen vor besondere Herausforderungen. Zudem erschweren die mit Blick auf die Zielgruppe limitierten Möglichkeiten der Diagnostik ein passgenaues schulisches Angebot. (Heinrichs et al., 2016) Professionelles Handeln vollzieht sich demnach vorrangig im Kontext von Unsicherheit und mangelnder Handlungsroutine, was auf notwendige Professionalisierungsmaßnahmen verweist.

*Beiträge des Symposiums*

### Individuelle Erlebensqualitäten junger Asylsuchender und Flüchtlinge in betrieblichen Praktika

**Hannes Reinke<sup>1</sup>, Tobias Kärner<sup>2</sup>, Karin Heinrichs<sup>1</sup>**

<sup>1</sup>Universität Bamberg, Deutschland, <sup>2</sup>Universität Konstanz, Deutschland

Mit Blick auf die Flucht- und Migrationsbewegungen der letzten Jahre ergeben sich Herausforderungen hinsichtlich der Integration derjenigen Personen, die längerfristig in Deutschland leben werden. Fast ein Drittel der im Jahr 2015 zugewanderten Personen sind Jugendliche oder junge Erwachsene im Alter zwischen 16 und 25 Jahren. Entgegen der Erwartung, gut ausgebildete und qualifizierte Menschen aufzunehmen, offenbarte sich eine starke Heterogenität hinsichtlich der Bildungsgrade der Zugewanderten. Hieraus resultiert wiederum ein erheblicher Qualifikationsbedarf, wenn man davon ausgeht, dass die eingewanderten Personen gesellschaftlich integriert und dem Arbeitsmarkt zugeführt werden sollen (OECD, 2016). Eine Chance für junge Asylsuchende und Flüchtlinge (jAuF) stellt die Teilnahme an zielgruppenspezifischen Bildungsgängen, wie den zweijährigen Berufsintegrationsklassen (BIK) in Bayern, dar. Inhaltlich zählen dazu neben der Sprachförderung in besonderem Maße auch die berufliche Grundbildung und Orientierung.

Berufsorientierung kann verstanden werden als ein lebenslanger Prozess des Abgleichs zweier Seiten. Auf der einen stehen individuelle Interessen, Fähigkeiten, aber auch soziale Rollenvorstellungen und auf der anderen gesellschaftliche und wirtschaftliche Ressourcen und Angebote (Butz, 2008; Gottfredson, 2005; Driesel-Lange et al., 2013). In Anlehnung an die sozial-kognitive Laufbahntheorie von Lent, Brown und Hackett (1994) bedingen abgesehen von soziodemografischen und biografischen Faktoren auch Selbstwirksamkeitserwartungen (SWE) die Berufswahl. Zudem vermögen das Erleben lern- und entwicklungsförderlicher Tätigkeitsmerkmale, wie Berufsfeldbezug, Anforderungsvielfalt und -niveau, sowie die Befriedigung der Basic Needs, Autonomieerleben, Erleben sozialer Eingebundenheit, Kompetenzerleben (Deci & Ryan, 1993), die Entwicklung beruflicher Interessen sowie von Berufswahlentscheidungen zu beeinflussen (Rausch, 2011).

Eines der am häufigsten genutzten Instrumente der schulischen Berufsorientierung stellt das Betriebspraktikum dar, welches auch Teil der BIK ist (ISB, 2016). Jedoch kann dessen berufsorientierender Charakter bei mangelnder Konzeption und Planung schnell verfehlt werden. Daher geht der Beitrag der Frage nach, welche Gestaltungsmerkmale betrieblicher Praktika aus Sicht der JAuF geeignet sind, um die individuelle Berufsorientierung im Sinne von Selbst- und Berufserkundung sowie Berufswahlklarheit zu unterstützen.

Um der Frage nachzugehen, wird auf die Ergebnisse einer Tagebuchstudie bei JAuF zurückgegriffen. In einer Eingangserhebung wurden 34 Lernende ( $m = 33$ ; 16–22 Jahre,  $M = 17.76$ ,  $SD = 1.3$ ) zu ihren beruflichen Vorstellungen in ihrem Heimatland und in Deutschland befragt und gebeten, diese hinsichtlich Geschlechtstyp, Prestige und Anforderungsniveau zu bewerten. In der sich anschließenden Tagebuchstudie, die sich über einen Zeitraum von fünf Wochen zu Beginn eines einjährigen Praktikums erstreckte, gaben die Befragten wöchentlich ihr Erleben am Arbeitsplatz hinsichtlich der Basic Needs und des Selbstwirksamkeitserlebens sowie Informationen zu lern- und entwicklungsfördernden Tätigkeitseigenschaften (in offenen und geschlossenen Antwortformaten) an. Zudem wurden Berufswünsche und deren Begründung regelmäßig erfragt.

Die Ergebnisse verdeutlichen, dass individuelle Unterschiede in der Anpassung beruflicher Vorstellungen und Ziele über die Aufenthalts- und Schulzeit in Deutschland bestehen. So haben sich Berufswünsche und deren Bewertungen bei zwei Dritteln der Schüler über die Zeit verändert. Bezogen auf die Förderung der Berufsorientierung durch betriebliche Praktika zeigten sich signifikant positive Rangkorrelationen zwischen der Höhe (i. S. d. Niveaus) des Erlebens sozialer Eingebundenheit während der Praktikumszeit, des Kompetenzerlebens, des Autonomieerlebens sowie der berufsbezogenen SWE. Auch zeichnen sich positive Zusammenhänge dieser Erlebensaspekte mit der Zufriedenheit mit dem Praktikum sowie dem Wunsch, im Praktikumsberuf eine Ausbildung absolvieren zu wollen, ab.

Weiterhin legen die korrelativen Analysen nahe, dass intraindividuell zeitlich stabile Ausprägungen (i. S. einer geringeren Streuung) der berufsbezogenen SWE, des Autonomieerlebens sowie des Kompetenzerlebens tendenziell mit dem Wunsch einhergehen, im Praktikumsberuf eine Ausbildung absolvieren zu wollen. Ergänzend zu den Auswertungen der standardisierten Skalen ergeben die Analysen der qualitativen Tagebuchdaten, dass eine Unterstützung der Berufsorientierung vor allem mit als positiv wahrgenommenen Aspekten der ausgeführten Praktikumsaktivitäten sowie mit dem Kompetenzerleben und dem Erleben sozialer Eingebundenheit einhergehen, was die Ergebnisse der quantitativen Analysen stützt.

Die Ergebnisse der Studie weisen darauf hin, dass Betriebspraktika die Berufsorientierung von JAuF unterstützen können, sofern sie psychologische Grundbedürfnisse berücksichtigen.

## **Heterogenität der Deutsch-Fähigkeiten junger Geflüchteter in der Berufsbildung**

**Elisabeth Maué, Stephan Schumann, Claudia Diehl**  
Universität Konstanz, Deutschland

### Theoretischer Hintergrund

Da etwa 30% der jüngst zugewanderten Flüchtlinge zwischen 16 und 25 Jahre alt sind (BAMF, 2017), spielt das System der beruflichen Bildung eine entscheidende Rolle für deren gesellschaftliche Integration in Deutschland. In Deutschland haben die Bundesländer auf den Zuzug von Asylsuchenden und Flüchtlingen mit der Einrichtung von spezifischen ein- oder zweijährigen Bildungsgängen reagiert, die strukturell der Berufsvorbereitung zugerechnet werden können (Baethge & Seeber, 2016). Jenseits diverser Ausgestaltungsunterschiede ist eine intensive Sprachförderung ein Hauptanliegen dieser zumeist vollzeitschulischen Angebote, da die Förderung der Fähigkeiten in Deutsch zweifellos eine wesentliche Säule der Integration darstellt (Esser, 2001).

### Fragestellung

Bislang liegen zu dieser Zielgruppe und insbesondere aus dem System der beruflichen Bildung kaum empirische Daten vor. Vor diesem Hintergrund ist es Ziel des vorliegenden Beitrags, die Deutsch-Fähigkeiten geflüchteter Jugendlicher und den Einfluss der ethnischen Klassenkomposition auf die Deutsch-Fähigkeiten in den Blick zu nehmen. Dabei interessiert auch, inwieweit die Selbsteinschätzungen der Lernenden mit objektiv gemessenen Leistungsdaten korrespondieren und welche subgruppenspezifischen Unterschiede sich zeigen.

### Methode

Die Datengrundlage entstammt der interdisziplinären Studie „RISE“, welche als Primärdatenerhebung zwei Messzeitpunkte in zwei aufeinander folgenden Schuljahren umfasst (t1: Sommer 2017; t2: Winter 2017/18). Im Vortrag werden Ergebnisse des ersten Erhebungszeitpunkts präsentiert. Als Instrumente wurden ein in sieben Sprachen verfügbarer Online-Fragebogen sowie Kurztests der kognitiven Grundfähigkeit (CFT20; Weiß, 1978) und des Sprachstands in Deutsch (Eigenentwicklung) eingesetzt. Die Untersuchung erfolgte mittels externer TestleiterInnen im jeweiligen Klassenverband. Das Sample beinhaltet 675 SchülerInnen (79% männlich) in 60 berufsvorbereitenden, spezifisch für geflüchtete Jugendliche angebotenen Klassen an 22 beruflichen Schulen im Regierungsbezirk Freiburg. Der Deutschttest umfasst 30 Items und adressiert in inhaltlicher Sicht basales und auf Alltagssituationen ausge-richtetes Begriffs- und Grammatikwissen sowie Leseverstehen. Der Test wurde auf Basis existierender Einstufungstests aus dem Bereich „Deutsch als Zweitsprache“ entwickelt und im Vorfeld der Untersuchung anhand einer vergleichbaren Zielgruppe pilotiert (EAP-Reliabilität=0.93). Die Selbsteinschätzung der Deutsch-Fähigkeiten erfolgte mittels vier Items (6-stufige Likertskala), die auf die Fähigkeiten in den Bereichen Verstehen, Sprechen, Lesen und Schreiben fokussieren. Die vier Items laden auf einem gemeinsamen Faktor und wurden daher für die bivariaten Analysen zu einer Skala zusammengefasst (Cronbachs Alpha=0.86). Als Indikator für die ethnische Klassenkomposition wurde pro Klasse die Anzahl der unterschiedlichen Geburtsländer durch die Anzahl der SchülerInnen dividiert (Wert nahe 0=große Homogenität; 1=völlige Heterogenität).

### Ergebnisse

Ein großer Anteil der Jugendlichen schätzt seine Deutsch-Fähigkeiten als zumindest gut, wenn nicht gar gut oder sehr gut ein. Weniger als 10% sind diesbezüglich kritischer. Dieser für beide Geschlechter gleichermaßen beobachtbare Befund kann möglicherweise durch Referenzgruppeneffekte erklärt werden. Eine gewisse Validität der Aussagen kann jedoch angesichts einer substanziellen bivariaten Korrelation der Selbsteinschätzung mit dem im Leistungstest erzielten Wert angenommen werden ( $r=0.34$ ; unter Kontrolle der kognitiven Fähigkeiten  $r=0.32$ ). Im Schnitt erzielten die Jugendlichen 19.07 Punkte von maximal 30 Punkten ( $SD=6.88$ ). Auch hier können keine Geschlechtseffekte identifiziert werden. Markant ist, dass Jugendliche, die laut Selbstauskunft das Sprachniveau A2 gemäß Gemeinsamen Europäischen Referenzrahmen erreicht haben, deutlich besser abschneiden als Lernende, die dies noch nicht erreicht haben ( $M=21.73$  vs.  $M=13.24$ ; Cohens  $d=1.57$ ). Differenziert nach häufigsten Herkunftsländern zeigen sich unter Kontrolle der kognitiven Grundfähigkeiten leicht überdurchschnittliche



Werte für Jugendliche aus Syrien (M=20.45; SD=6.20), während Lernende aus Afghanistan (M=18.54; SD=7.25, Cohens d=0.29; Referenz Syrien), dem Irak (M=17.63; SD=7.10, Cohens d=0.44) und Gambia (M=15.84; SD=7.59, d=0.70) vergleichsweise schlechter abschnitten. In Klassen mit größerer Heterogenität erzielen SchülerInnen bessere Ergebnisse im Deutschtest (r=.17; unter Kontrolle der kognitiven Fähigkeiten r=.18). Dieser Befund kann als empirischer Hinweis darauf gedeutet werden, dass eine ethnische Mischung sich möglicherweise positiv auf den Erwerb der deutschen Sprache auswirkt, da die Lernenden dadurch gezwungen sind, sich untereinander auf Deutsch zu unterhalten. Weitere Auswertung, insbesondere die mehrbenenanalytische Überprüfung des Effekts der Klassenkomposition, werden an der GEBF präsentiert.

### **Ethnische Klassenzusammensetzung und Sprachkompetenzen: Erstsprachgebrauch von Jugendlichen mit Migrationshintergrund im schulischen Umfeld**

**Julian Seuring<sup>1</sup>, Camilla Rjosk<sup>2</sup>, Petra Stanat<sup>2</sup>**

<sup>1</sup>Universität Bamberg, Deutschland, <sup>2</sup>Institut für Qualitätsentwicklung im Bildungswesen (IQB), Humboldt-Universität zu Berlin, Deutschland

Zu Einflüssen der ethnischen Klassenzusammensetzung auf den schulischen Leistungserwerb wird häufig die Verwendung der Herkunftssprache unter Jugendlichen mit Migrationshintergrund als vermittelnde Einflussgröße diskutiert. Es wird angenommen, dass Klassen mit einem hohen Anteil an SchülerInnen mit Migrationshintergrund erhöhte Opportunitäts- und Anreizstrukturen bieten die Erst- bzw. Herkunftssprache (L1) anstelle der Unterrichtssprache (L2) zu verwenden (Esser, 2006). Dies kann sich ungünstig auf den Kompetenzerwerb in der Unterrichtssprache auswirken, da ein häufiger Gebrauch der Herkunftssprache gleichzeitig die Gelegenheiten begrenzt, die L2 zu sprechen und dabei entsprechende Kenntnisse auszubauen (Driessen, 2002). „Language-competetion“ konnte für den L1-Gebrauch im familiären Kontext mehrfach beobachtet werden (z.B. Baumert & Schümer, 2001; Stanat et al., 2010; Strobel & Seuring, 2016; Walter, 2009). Hingegen ist unklar, ob sich vergleichbare Muster auch für das schulische Umfeld bestätigen lassen und inwiefern der L1-Gebrauch mit der zuwanderungsbezogenen Klassenkomposition zusammenhängt.

Der vorliegende Beitrag greift diese Fragestellung auf und untersucht, ob sich die zuwanderungsbezogene Klassenkomposition, vermittelt über den L1-Gebrauch im schulischen Umfeld, auf das sprachliche Kompetenzniveau auswirkt. Von den theoretischen Überlegungen des durch den Sprachgebrauch bedingten Zugangs zur L2 lassen sich spezifische Erwartungen mit Blick auf Einflüsse der zuwanderungsbezogenen Klassenkomposition ableiten. Zunächst setzt der L1-Gebrauch im schulischen Umfeld voraus, dass mehrere MitschülerInnen dieselbe L1 sprechen. Demzufolge sollte weniger die ethnische Zusammensetzung der Klasse als vielmehr die sprachliche bzw. ethno-linguale Komposition ausschlaggebend für das in der Unterrichtssprache erreichte Kompetenzniveau sein. Zudem kann erwartet werden, dass die ethno-linguale Klassenkomposition vorrangig mit den sprachlichen Leistungen von SchülerInnen mit nicht-deutscher Erstsprache in Zusammenhang steht, da SchülerInnen, die ausschließlich die Unterrichtssprache sprechen, auch nur diese im schulischen Umfeld verwenden können. Weiterhin lässt sich die Wirkungsweise des L1-Gebrauchs analog auf den Erwerb bzw. Erhalt der Erstsprachkompetenzen übertragen. Wenn also Einflüsse der ethno-lingualen Klassenkomposition maßgeblich über einen zunehmenden L1-Gebrauch im schulischen Umfeld vermittelt werden, dann sollte sich dies ebenfalls in höheren Erstsprachkompetenzen widerspiegeln.

Für die empirische Analyse wurden Daten der Startkohorte 4 (Klasse 9) des Nationalen Bildungspanels (NEPS; Blossfeld et al., 2011) herangezogen. Die Analytestichprobe umfasst N=14981 RegelschülerInnen (N=4330 mit Migrationshintergrund; N=3123 mit nicht-deutscher Erstsprache) in N=1006 Schulklassen. Basierend auf den individuellen Angaben zum Geburtsland der SchülerInnen, der Eltern und Großeltern sowie zu der in der Familie erlernten Erstsprache wurde der Anteil SchülerInnen mit Migrationshintergrund bzw. mit nicht-deutscher Erstsprache auf Klassenebene aggregiert und so transformiert, dass eine Steigungseinheit zehn Prozentpunkten entspricht. Die sprachlichen Leistungen wurden anhand eines Leseverstehenstests in Deutsch (Gehrer et al., 2012) sowie zwei Hörverstehenstests in den Herkunftssprachen Russisch und Türkisch (Edele et al., 2012) gemessen. Zusätzlich enthalten die Daten einen Indikator zum individuellen Sprachgebrauch im schulischen Umfeld, womit sich die angenommene Mediationshypothese direkt testen lässt. Die empirische Prüfung der theoretischen Annahmen erfolgte über die Schätzung linearer Mehrebenenmodelle, unter zusätzlicher Berücksichtigung leistungsbezogener und sozio-ökonomischer Hintergrundmerkmale auf Individual- sowie Klassen-Ebene.

Zunächst wurden Einflüsse der ethnischen Kompositionen denen der ethno-lingualen Komposition gegenübergestellt. Bei gleichzeitiger Berücksichtigung beider Merkmale konnte nur zwischen der ethno-lingualen Komposition und dem deutschsprachigen Leseverstehen ein signifikant negativer Zusammenhang festgestellt werden ( $\beta=-0,01$ ;  $p<0,01$ ). Weiterhin scheinen ausschließlich SchülerInnen mit einer nicht-deutschen Erstsprache von den Effekten der ethno-lingualen Klassenkomposition betroffen zu sein. Gleichzeitig zeigte sich für diese SchülerInnen ein positiver Zusammenhang zwischen der ethno-lingualen Klassenkomposition mit dem herkunftssprachlichen Hörverstehen (Russisch:  $\beta=0,10$ ; n.s. / Türkisch:  $\beta=0,08$ ;  $p<0,05$ ). Die Effekte der ethno-lingualen Klassenkomposition auf die Kompetenzen in L2 und L1 verringerten sich deutlich, wenn in den Modellen zusätzlich der Sprachgebrauch im Schulumfeld berücksichtigt wurde. Insgesamt stützen die Befundmuster die theoretische Annahme, dass Einflüsse der zuwanderungsbezogenen Klassenzusammensetzung auf den sprachlichen Leistungserwerb über den L1-Gebrauch im Schulumfeld vermittelt werden. Allerdings zeigt sich ebenfalls, dass Effekte der zuwanderungsbezogenen Klassenkomposition auf die Leistungen in der Unterrichtssprache sehr schwach ausfallen und demnach nur geringfügig zur Erklärung von Leistungsunterschieden beitragen.

### **Umgang mit Heterogenität: Binnendifferenzierung im Unterricht mit jugendlichen Asylsuchenden und Flüchtlingen aus der Sicht von Lehrpersonen**

**Tobias Kärner<sup>1</sup>, Karin Heinrichs<sup>2</sup>, Hannes Reinke<sup>2</sup>**

<sup>1</sup>Universität Konstanz, Deutschland, <sup>2</sup>Universität Bamberg, Deutschland

Im Zuge des anhaltenden Flüchtlingsstroms finden immer mehr Kinder und Jugendliche den Weg nach Deutschland und somit in das deutsche Bildungssystem (Berthold, 2014). Im Sinne einer erfolgreichen Integration gilt es daher auch bildungspolitisch, den damit einhergehenden Herausforderungen zu begegnen. Hierbei ist zu berücksichtigen, dass sich die betreffende Schülerschaft hinsichtlich ihrer Zusammensetzung durch einen hohen Grad an Heterogenität auszeichnet, da sich neben interindividuellen Unterschieden in Herkunftsländern und Konfessionen v.a. große Differenzen in den jeweiligen schulischen Vorkenntnissen und im Sprachniveau zeigen. Der Umgang mit leistungsbezogener, kultureller und sprachlicher Heterogenität im Klassenzimmer wird von den Lehrpersonen als große Herausforderung angesehen (Kärner et al., 2016). Im Rahmen der Heterogenitätsdebatte wird daher verstärkt der Einsatz von Binnendifferenzierungsmethoden (BD) gefordert (Scholz, 2010) und angesichts der zu bewältigenden Herausforderungen im Zusammenhang mit der Beschulung jugendlicher Asylsuchender und Flüchtlinge (jAuF) als gewinnbringend erachtet. Durch entsprechende Maßnahmen soll den individuellen Bedürfnissen eines/r

jeden Lernenden möglichst gut nachgekommen und somit zu einer gelingenden beruflichen und gesellschaftlichen Integration beigetragen werden (Heinrichs et al., 2016).

Welche konkreten BD im Unterricht mit jAuF jedoch als praktikabel erscheinen und inwiefern aufseiten der Lehrenden adäquate Kompetenzen, entsprechende Maßnahmen umzusetzen, in ausreichendem Maß vorhanden sind, erscheint bislang nicht ausreichend erforscht zu sein. In unserem explorativen Beitrag wird daher den Fragen nachgegangen, (1) auf welche methodisch-didaktischen Variationsmöglichkeiten die im Rahmen der Beschulung von jAuF eingesetzten BD abzielen und welche Schüler/innenmerkmale als Differenzierungskriterien herangezogen werden, (2) welche Grenzen Lehrkräfte bei der Umsetzung von BD in Flüchtlingsklassen sehen und (3) inwiefern sich die Lehrkräfte für fähig bzw. ausreichend ausgebildet halten, BD im Flüchtlingsunterricht adäquat anzuwenden.

Unsere Stichprobe setzt sich aus 16 Lehrpersonen der Sekundarstufe (Mittelschule, Berufsschule) zusammen (8 weiblich; Durchschnittsalter=41.7 Jahre), welche alle in der Flüchtlingsbeschulung tätig sind und im Durchschnitt 7.4 Jahre Erfahrung im Unterrichten von Flüchtlingen bzw. Migranten haben. Die Lehrpersonen wurden einzeln und persönlich in halbstrukturierten Interviews befragt. Im Anschluss an die Datenerhebung wurden die Daten mittels strukturierender Inhaltsanalyse nach Mayring (2015) auf Grundlage eines deduktiv-induktiven Kategorienschemas ausgewertet.

In den Ergebnissen zeigt sich folgendes Bild: Bezogen auf die Frage, welche Schüler/innenmerkmale von den Lehrpersonen als Differenzierungskriterien herangezogen werden, orientieren sich alle Befragten am Leistungsvermögen der Lernenden. Die Hälfte der Teilnehmer/innen nannte den Sprachstand als weiteres Differenzierungsmerkmal, gefolgt von soziodemografischen Merkmalen. Hinsichtlich der Frage, auf welche methodisch-didaktischen Variationsmöglichkeiten die eingesetzten BD abzielen, finden sich die häufigsten Nennungen (15) zur Differenzierung anhand der Aufgabenstellung (z.B. Variation des Schwierigkeitsgrads). Jeweils 11 Lehrpersonen nannten Lehrerhilfe und Sozialform als weitere Facetten, welche hinsichtlich einer möglichst individuellen Förderung Variationsspielraum zulassen. Auf die Frage nach wahrgenommenen Grenzen des Einsatzes von BD wurde eine ganze Bandbreite an möglichen Problemfeldern genannt, welche u.a. den (Zeit-)Aufwand für deren Planung und Einsatz, den Mangel an Lehrpersonal sowie die eigene unzureichende Ausbildung in diesem Bereich beinhalten. Auf die Frage nach notwendigen Lehrer/innenkompetenzen, die für die Realisierung von BD im Unterricht mit jAuF notwendig erscheinen, nannten 12 von 16 Lehrpersonen interkulturelle und soziale Kompetenzen, gefolgt von didaktischen (9) und diagnostischen (8) Kompetenzen.

Die gefundenen Ergebnisse zeichnen ein Bild, nach dem alle der 16 befragten Lehrkräfte den Einsatz von BD im Unterricht mit jAuF für notwendig erachten. Auch werden bereits unterschiedliche Methoden in den Unterricht integriert, was jedoch nach Aussagen der Befragten in der schulischen Praxis an unterschiedlichen Stellen an Grenzen stößt. Inhaltliche Implikationen unserer Befunde ergeben sich für die erste und zweite Phase der Lehrer/innenausbildung und sie zeigen Anforderungen an die Weiterentwicklung fachdidaktischer Konzepte bezüglich des Unterrichts in sehr heterogenen Klassen auf. Da es sich jedoch um eine explorativ-qualitative Studie mit einer relativ geringen Stichprobe handelt, wären die gefundenen Ergebnisse in weiterführenden, darauf aufbauenden quantifizierenden Studien mit größeren Stichproben zu validieren.

## F17

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H114

### **Schulisches Engagement und Schulentfremdung: Ergebnisse empirischer Forschung aus der Schweiz, Deutschland und Luxemburg**

Chair(s): **Elena Makarova** (Pädagogische Hochschule FHNW, Schweiz)

DiskutantIn(nen): **Robert Grassinger** (Pädagogische Hochschule Weingarten, Deutschland)

Frühzeitiger Schulabbruch oder School Dropout ist eine der zentralen Herausforderungen, die sich im Bereich der Chancengerechtigkeit in den Bildungswesen der OECD Länder stellen (OECD, 2012). In der jüngsten empirischen Schul- und Bildungsforschung wird die Ansicht vertreten, dass es sich beim Schulabbruch um den Kulminationspunkt eines langfristigen Prozesses des geringeren schulischen Engagements und der zunehmenden Entfremdung von der Schule handelt (Archambault et al., 2009; Fall & Roberts, 2012; Henry et al., 2012; Rumberger & Rotermund, 2013). Folglich ist es von hohen Bedeutung, den Forschungsfokus nicht auf den eigentlichen Schulabbruch, sondern auf dessen Vorlauf zu richten. Neben den kognitiven und verhaltensbezogenen Aspekten, werden sowohl die motivationalen und emotionalen Determinanten des Lern- und Bildungserfolges von Kindern und Jugendlichen als auch die Dimensionen der sozialen Ungleichheit verstärkt untersucht (Göbel et al., 2016; Hadjar et al., 2015; Hascher & Hagenauer, 2010; Makarova & Herzog, 2013).

Vor diesem Hintergrund fokussiert das Symposium die individuums- und schulbezogenen Modalitäten und Determinanten schulischen Engagements und der Schulentfremdung und strebt an, die Zusammenhänge zwischen den beiden Konstrukten aufzudecken. Es werden Ergebnisse aus zwei Studien, die im Rahmen internationaler Kooperationsprojekte entstanden sind, referiert. Es handelt sich um Studien aus der Schweiz und Deutschland, die als Teil der internationalen Studie «School dropout among immigrant students: Types of dropout and predictors» durchgeführt wurden und um eine bi-nationale Studie aus der Schweiz und Luxemburg mit dem Schwerpunkt „School Alienation in Switzerland and Luxembourg“.

Der erste Beitrag thematisiert schulisches Engagement von Jugendlichen mit Migrationshintergrund und einheimischen Jugendlichen und geht insbesondere auf die Aspekte der sozialen Integration in Zusammenhang mit dem Ausmass an verhaltensbezogenem, emotionalen und kognitiven Engagement in Schule und Unterricht von Jugendlichen der 7. bis 9. Schulstufe in der Schweiz ein.

Im Zentrum des zweiten Beitrags steht die Frage, welche Dimensionen der wahrgenommenen Beziehung zur Klassenlehrerin bzw. zum Klassenlehrer und deren Kompetenzen als Erklärung für das schulische Engagement von Schülerinnen und Schülern der 7. Jahrgangsstufe verschiedener Schulformen in Deutschland (NRW) relevant sind.

Ausgehend von der Annahme, dass es sich zwischen Schulentfremdung und der Mitarbeit von Schülerinnen und Schülern im Unterricht um eine reziproke Beziehung handelt, werden im dritten Beitrag anhand der Längsschnittdaten von Jugendlichen der 7.- 8. Schulstufe im Kanton Bern mehrere Kausalzusammenhänge zwischen Entfremdung von Mitschülerinnen und Mitschülern, von Lehrpersonen bzw. vom schulischen Lernen und der aktiven Mitarbeit im Unterricht analysiert.

Der vierte Beitrag verfolgt einen domänenspezifischen Ansatz in der Analyse von Schulentfremdungsprofilen entlang der Achsen sozialer Ungleichheit unter den Luxemburgischen Schülerinnen und Schülern der 7.-8. Klasse. Der Beitrag zeigt, anhand welcher Charakteristika von Schülerinnen und Schülern sich Gruppen innerhalb der Schulentfremdungsprofile unterscheiden und welche Einflüsse domänenspezifischer Schulentfremdung sich auf Schuldevianz sowie Schulleistungen nachweisen lassen.

Gesamthaft leistet das Symposium einen wertvollen Beitrag zum Tagungsthema, da darin präsentierte Ergebnisse der empirischen Forschung zum schulischen (Dis-)Engagement und zur Schulentfremdung von Schülerinnen und Schülern für das pädagogische Handeln von Lehrpersonen höchst relevant sind. Zum einen können sie zur frühzeitigen Identifikation und gezielten Förderung von Jugendlichen, die ein erhöhtes Schulabbruchrisiko aufweisen, genutzt werden. Zum anderen können daraus Möglichkeiten zur Prävention und Intervention des frühzeitigen Schulabbruchs abgeleitet werden.

#### *Beiträge des Symposiums*

### **Schulisches Engagement von Schülerinnen und Schülern mit und ohne Migrationshintergrund: Eine empirische Studie unter Jugendlichen der 7. bis 9. Schulstufe in der Schweiz**

**Elena Makarova<sup>1</sup>, Selina Teuscher<sup>1</sup>, Marina Kipfer<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule FHNW, Schweiz, <sup>2</sup>Pädagogische Hochschule Bern, Schweiz

#### Theoretischer Hintergrund

In der empirischen Forschung, die sich mit Schulabbruch (School Dropout) auseinandersetzt, wird die Ansicht vertreten, dass es sich dabei nicht um eine „Kurzschlussreaktion“ oder um ein „plötzliches Ereignis“ (Stamm, 2012, p. 25, p.137) handelt, sondern um den Höhepunkt eines langfristigen Prozesses der Entfremdung von der Schule (Alexander, Entwisle, & Horsey, 1997; Makarova & Herzog, 2013). Diese Annahme wird durch Befunde von Längsschnittstudien bekräftigt, die belegen, dass mangelndes schulisches Engagement auf einen späteren Schulabbruch hindeutet. Folglich wird der Fokus der gegenwärtigen Forschung nicht auf den eigentlichen Schulabbruch, sondern vermehrt auf dessen Vorlauf gelegt (Lamb & Markussen, 2011).

Die Forschung zum schulischen Engagement ist relativ jung und nimmt ihren Anfang in den 1980er Jahren (Appleton, Christenson, & Furlong, 2008). In der zeitgenössischen Forschung finden sich mehrere Konzeptionen schulischen Engagements. Einigung besteht dahingehend, dass es sich dabei um ein multidimensionales Konstrukt handelt, wobei sich die Anzahl der Teilaspekte, die schulisches Engagement konstituieren, variiert (Furlong et al., 2003).

#### Fragestellung

Die Schweizer Studie versteht schulisches Engagement als ein multidimensionales Konstrukt und analysiert individuums- und schulbezogenen Determinanten schulischen Engagements unter Jugendlichen mit und ohne Migrationshintergrund. Die Studie geht folgenden Fragen nach:

- Gibt es Unterschiede im Ausmass schulischen Engagements von Jugendlichen mit Migrationshintergrund und von einheimischen Jugendlichen?
- Wie wirken sich wahrgenommene Diskriminierung und die Qualität von Peer-Beziehungen im schulischen Kontext auf schulisches Engagement aus?

#### Methode

Die Schweizer Studie wurde im Rahmen der internationalen „Studie School dropout among immigrant students: Types of dropout and predictors“ durchgeführt. Diese Studie umfasste Einzelstudien in Deutschland, den Niederlanden, Israel, Türkei und der Schweiz und verfolgte das Ziel, länderspezifische und -übergreifende Determinanten von schulischem Engagement in einem ländervergleichenden Design zu analysieren.

In der Schweizer Studie partizipierten 220 Schülerinnen und Schüler, die im Durchschnitt 15 Jahre alt waren. Zwei Drittel der Schülerinnen und Schüler (71,5%) hatten einen Migrationshintergrund. Die Studie differenzierte zwischen drei Herkunftsgruppen: Jugendliche mit Migrationshintergrund waren entweder jene Jugendlichen, deren Eltern beide im Ausland geboren wurden (Kategorie Migrant/in), oder jene aus Familien, in denen entweder die Mutter oder der Vater aus dem Ausland stammte (Kategorie Bi-national) die dritte Herkunftsgruppe bildeten einheimische Jugendliche, deren Eltern aus der Schweiz stammen. Die Daten wurden mittels eines standardisierten Fragebogens erhoben. Das Konstrukt schulisches Engagement wurde in drei Dimensionen operationalisiert: verhaltensbezogenes Engagement, emotionales Engagement und kognitives Engagement (Fredricks, Blumenfeld, Friedel, & Paris, 2005). Die Daten wurden mit multivariaten Verfahren analysiert.

#### Ergebnisse

Zusammenfassend leisten die Ergebnisse der Studie einen wertvollen Beitrag zum internationalen Forschungsstand. Zwar zeigen die Ergebnisse, dass männliche Jugendliche im Vergleich zu weiblichen Jugendlichen ein höheres Risiko haben, von schulischem Disengagement betroffen zu sein. Dieser Befund konnte jedoch lediglich für die verhaltensbezogene Dimension festgestellt werden. Im Weiteren zeigte sich, dass bi-nationale Jugendliche am stärksten von schulischem Disengagement betroffen sind, während Jugendliche der Kategorie Migrant/in sich in keiner Dimension signifikant von einheimischen Jugendlichen unterscheiden. Da Migrantinnen und Migranten trotz einem verhältnismässig hohen schulischen Engagement leistungsmässig weniger erfolgreich sind, liefert unsere Studie einen Hinweis auf die Existenz des „Engagement-Achievement Paradox“ (Shernoff & Schmidt, 2008, p. 564) im Schweizerischen Bildungssystem.

Übereinstimmend mit dem bisherigen Forschungsstand (Juvonen, Espinoza, & Knifsend, 2013), zeigt die Studie, dass die soziale Integration von zentraler Bedeutung für schulisches Engagement der Schülerinnen und Schüler mit Migrationshintergrund ist. Zugleich deuten die Ergebnisse unserer Studie darauf hin, dass sich der Effekt der sozialen Integration nicht generell für schulisches Engagement postulieren lässt, sondern nur für verhaltensbezogenes und emotionales Engagement, nicht jedoch für kognitives Engagement von Jugendlichen. Die Ergebnisse sind für pädagogisches Handeln von Lehrpersonen von hoher Bedeutung, da sie für die Erarbeitung von Präventionsmöglichkeiten eines frühzeitigen Schulabbruchs genutzt werden können.

Keywords: Schulisches Engagement, soziale Integration, Jugendliche mit Migrationshintergrund, standardisierte Befragung

## **Der Einfluss der Lehrperson auf die emotionalen, verhaltensbezogenen und kognitiven schulischen Einstellungen von Schülerinnen und Schülern**

**Kerstin Göbel, Zuzanna M. Lewandowska**  
Universität Duisburg-Essen, Deutschland

#### Theoretischer Hintergrund

Ein zentraler Faktor, welcher insbesondere bei sozioökonomisch benachteiligten Schülerinnen und Schülern eine wesentliche Rolle im Kontext von Bildungserfolg spielen kann, sind deren schulbezogene Einstellungen bzw. ihr schulbezogenes Wohlbefinden (School Engagement) (Fredericks, Blumenfeld & Paris, 2004; Appleton, Christensen & Furlong, 2008). Ihren Ursprung fand die Forschung zu dem sog. emotionalen, kognitiven und verhaltensbezogenen School Engagement von Kindern und Jugendlichen in dem Versuch, Erklärungen für schulischen Absentismus bzw. Schulabbruch zu finden (Mosher & MacGowan, 1985; Stamm, 2007; Makarova & Herzog, 2013; Wang & Fredericks, 2014). Seit dem grundlegenden Artikel von Mosher & MacGowan im Jahre 1985 ist das multidimensionale Konzept School Engagement vor allem im US-amerikanischen Raum zu einem zentralen Gegenstand der Bildungsforschung geworden (Mahdikhani, 2015). Im Rahmen verschiedener empirischer Untersuchungen zeichnete sich ab, dass vor allem die Beziehung zu einflussreichen Erwachsenen eine relevante Bedeutung für Schülerinnen und Schüler für ihre schulischen Einstellungen hat (Skinner, Wellborn, & Connell, 1990; Woolley & Bowen, 2007). Wesentlich für eine positive Einstellung zur Schule scheinen die Lehrerinnen und Lehrer, insbesondere die Klassenlehrerin bzw. der Klassenlehrer, zu sein (Graziano, Reavis, Keane, & Calkins, 2007; Roorda, Koomen, Split & Oort, 2011; Göbel, Makarova & Filsecker, 2016; Vollet, Kindermann & Skinner, 2017). Im deutschsprachigen Raum gibt es bislang wenig Forschung zu den kognitiven, verhaltensbezogenen und emotionalen Aspekten von School Engagement der Schülerinnen und Schüler. Besonders zum spezifischen Zusammenhang zwischen der Unterstützung, Motivation und Kompetenz der Lehrenden und dem School Engagement der Lernenden gibt es aktuell kaum Forschungsarbeiten.

#### Fragestellung

Das Ziel der vorliegenden Studie ist es, herauszuarbeiten, welche Faktoren im Zusammenhang mit niedrigem bzw. hohem School Engagement bei in Deutschland beschulten Lernenden stehen. Im Zentrum steht die Frage, in welcher Weise die wahrgenommene Beziehung zur Klassenlehrerin bzw. zum Klassenlehrer und deren Kompetenzen als Erklärung für School Engagement der Lernenden herangezogen werden können.

#### Methode

Der vorliegende Beitrag präsentiert Daten einer Untersuchung von Schülerinnen und Schülern der 7. Jahrgangsstufe verschiedener Schulformen in Nordrhein-Westfalen (N=708), welche im Kontext der Studie School Engagement der Universität Duisburg-Essen befragt wurden. Als Prädiktoren für das kognitive, verhaltensbezogene und emotionale School Engagement von Schülerinnen und Schülern wurden die wahrgenommene Unterstützung der Lehrperson, die Kompetenz des Erklärens von Unterrichtsinhalten der Lehrperson, die Ermutigung durch die Lehrperson, das Engagement der Lehrperson, die Beziehung der Schülerin bzw. des Schülers zur Lehrperson sowie soziodemografische Informationen über die Lernenden mithilfe eines Fragebogens erhoben.

#### Ergebnisse

Unter statistischer Kontrolle von Geschlecht, Migrationshintergrund, sozioökonomischem Status und kulturellem Kapital konnten im Rahmen von Regressionsanalysen, allem die Beziehung der Schülerin bzw. des Schülers zur Klassenlehrerin bzw. zum Klassenlehrer als relevanter Prädiktor für School Engagement identifiziert werden. Die wahrgenommene Qualität der Beziehung zur Klassenlehrerin bzw. zum Klassenlehrer erklärt auch einen deutlich höheren Anteil an Varianz von School Engagement als dies für die Individualmerkmale der Fall ist. Die Ergebnisse der Analysen weisen darauf hin, dass Lehrkräfte eine entscheidende Rolle im Hinblick auf die emotionalen, kognitiven und verhaltensbezogenen Einstellungen von

Schülerinnen und Schülern einnehmen. Vor dem Hintergrund der Befunde werden im Beitrag Vorschläge für die universitäre Lehrerbildung sowie für schulische Interventionsmöglichkeiten diskutiert.

Keywords: School Engagement; emotionale, verhaltensbezogene, kognitive schulbezogene Einstellungen; schulisches Wohlbefinden; Unterstützung von Lehrpersonen; schulischer Absentismus

## **Schulentfremdung und Mitarbeit im Unterricht**

**Tina Hascher, Julia Morinaj, Kaja Marcin**  
Universität Bern, Schweiz

### Theoretischer Hintergrund

Das Phänomen der Schulentfremdung rückt zunehmend in den Blickpunkt der Bildungsforschung, da Schulentfremdung den Lernerfolg der Schülerinnen und Schüler beeinträchtigt. Kinder und Jugendliche können sich von ihren Mitschüler/-innen, den Lehrpersonen sowie dem schulischen Lernprozess entfremden (Hascher & Hadjar, 2017). Schulentfremdung wird von einer Vielzahl negativer Faktoren begleitet, wie beispielsweise einer Depression, emotionalem Rückzug oder deviantem Verhalten (Brown, Higgins & Paulsen, 2003, Hascher & Hagenauer, 2010, Tarquin & Cook-Cottone 2008). Bisherige Forschung weist zudem darauf hin, dass Schulentfremdung die Mitarbeit der Schülerinnen und Schüler im Unterricht reduziert (e.g., Altenbauch, Engel, & Martin, 1995; Barber et al., 2005; Carlson, 1995; Hascher & Hagenauer, 2010). Die aktive Teilnahme am Unterricht ist jedoch eine notwendige, minimale Voraussetzung für erfolgreiches Lernen (Finn, 1989) und spielt eine wichtige Rolle beim Erreichen der Bildungsziele (Barber, Stone, & Eccles, 2005). Umgekehrt fördert eine aktive Teilnahme am Unterricht das Gefühl der Zugehörigkeit zur Schule und den schulischen Akteuren sowie die Wertschätzung von schulischem Lernen. Gelingt es Lehrpersonen die Schülerinnen und Schüler aktiv im Unterricht zu involvieren, kann dies folglich der Schulentfremdung entgegenwirken.

Forschungen zu schulischer Mitarbeit und Schulentfremdung beschränkten sich bisher auf Querschnittsuntersuchungen, welche die Richtung des Zusammenhangs zwischen den beiden Konstrukten offenlassen. Da auch eine reziproke Beziehung zwischen Schulentfremdung und der Mitarbeit im Unterricht denkbar ist, wurden in der vorliegenden Längsschnittstudie mehrere Kausalzusammenhänge untersucht: Dies bedeutet, es wurden alle möglichen Wirkrichtungen zwischen Entfremdung von Mitschüler/-innen, von Lehrpersonen bzw. vom schulischen Lernen und der aktiven Mitarbeit im Unterricht analysiert.

### Forschungsfrage

Hauptziel dieser explorativen Studie war es, die Richtung des Zusammenhangs zwischen Schulentfremdung und Schülerbeteiligung an den Unterrichtsaktivitäten über zwei Messzeitpunkte hinweg mithilfe von cross-lagged Panel-Analysen zu untersuchen. Es wurden die Korrelationen zu jedem einzelnen Messzeitpunkt sowie die cross-lagged Effekte zwischen Schulentfremdung und Mitarbeit im Unterricht über die zwei Messzeitpunkte hinweg berechnet. Somit wird getestet, wie sich Entfremdung von Mitschüler/-innen, Lehrpersonen und dem schulischen Lernen auf die Mitarbeit im Unterricht auswirkt, und zugleich erforscht, welchen Einfluss die Mitarbeit im Unterricht auf die Entwicklung von Schulentfremdung in den drei o.g. Bereichen ausübt. Aufgrund des mangelnden Befundlage werden diesbezüglich keine Hypothesen spezifiziert.

### Methode

Die Daten basieren auf den ersten zwei Erhebungswelle (2016 und 2017) des binationalen Forschungsprojekts „School Alienation in Switzerland and Luxembourg“ (SASAL, 2015-2018). Die Stichprobe umfasste zum ersten Erhebungszeitpunkt 550 Lernende der 7. Klassenstufe (45.2% männlich; Malter= 13.0 Jahre [SD = .55]), zum Erhebungszeitpunkt 540 Lernende der 8. Klassenstufe (45.6% männlich; Malter= 14.0 Jahre [SD = .57]). Die für die Analysen eingesetzten Daten basieren auf den Antworten von 508 Schülerinnen und Schüler der Sekundarstufe des Kantons Bern, welche an beiden Erhebungswellen teilnahmen. Schulentfremdung wurde mithilfe des „School Alienation Scales“ (SALS; Morinaj et al., 2017) erfasst, welcher die Entfremdung in den drei Bereichen „Mitschüler/-innen“, „Lehrpersonen“ und „schulisches Lernen“ erfasst. Die Mitarbeit im Unterricht wurde mit Hilfe von 5 Items erfasst (Eder, 1995).

### Ergebnisse

Erste Analysen weisen auf signifikante Effekte zwischen Entfremdung von Lehrpersonen und Entfremdung vom schulischen Lernen in der 7. Klassenstufe auf die Mitarbeit im Unterricht in der 8. Klassenstufe hin ( $p < .05$ , bzw.  $p < .05$ ). Die Ergebnisse implizieren einen Kausalzusammenhang zwischen Entfremdung von Lehrpersonen bzw. dem schulischen Lernen und der Mitarbeit der Schülerinnen und Schüler im Unterricht. Bezüglich der Entfremdung von Mitschüler/-innen konnte kein signifikanter Effekt auf die Unterrichtsmitarbeit festgestellt werden. Auch umgekehrte oder wechselseitige Wirkrichtungen liessen sich nicht nachweisen.

Keywords: Schulentfremdung, Mitarbeit im Unterricht, Sekundarstufe, Längsschnittdesign, cross-lagged Panel Analysen

## **Schulentfremdung im Luxemburgischen Sekundarbereich: Eine domänenspezifische Betrachtung ausgewählter Ursachen und Wirkungen**

**Andreas Hadjar, Jan Scharf, Alyssa Greu**  
Universität Luxemburg, Luxemburg

### Theoretischer Hintergrund

Der Prozess der über die Schulzeit abnehmenden Bindung von Schülerinnen und Schüler an Bildung und Schule wird als zunehmende Schulentfremdung konzeptualisiert. Unter Schulentfremdung werden dabei negative Einstellungen gegenüber bestimmten schulischen Domänen – Lernen, Lehrpersonen und Mitschüler/-innen – verstanden, die emotionale und kognitive Komponenten haben. Schulentfremdung ist Teil der Mechanismen hinter Bildungsungleichheiten, da sie zum einen im Sinne eines Ressourcendefizits eng mit Unterschieden in Schulleistungen (primären Effekte) verbunden ist, zum anderen aber auch Einflüsse auf Bildungsentscheidungen (sekundäre Effekte; Boudon 1974) haben kann. Der bisherige Forschungsstand zeigt, dass Jungen und Arbeiterkinder eine ausgeprägtere Schulentfremdung aufweisen, während Schüler/-innen mit Migrationshintergründen tendenziell weniger schulentfremdet sind. Schulentfremdung hat Wirkungen auf der Verhaltensebene, indem sie Handlungsalternativen dahingehend strukturiert, dass tendenziell Handlungsmuster – etwa Schulsekretariat – in Betracht gezogen werden, die diesen negativen Einstellungen entsprechen, während etwa schulerfolgspförderliche Handlungen nicht in das Set an Alternativen aufgenommen werden oder dort eine geringere Rolle spielen. Entsprechend geht Schulentfremdung mit einem geringeren Engagement in der Schule, einer höheren Schuldevianz und einem höheren Risiko, die Schullaufbahn früh zu beenden, einher (vgl. Archambault et al. 2009; Hascher/Hagenauer 2010; Hadjar et al. 2015). Die Ursachen für Schulentfremdung sind sowohl sozialisationstheoretisch als auch situativ zu fassen: Einerseits wird

Schulentfremdung durch Einstellungen gegenüber Schule und Bildung im Elternhaus oder Peer-Umwelten geprägt, zum anderen ist sie abhängig von situativen Erlebnissen der Schüler/innen in der Schule (Hascher/Hagenauer 2010; Hadjar et al. 2015). Eine Schulumwelt, die den Bedürfnissen der Schüler/innen gerecht wird (Eccles/Midgley 1989) und wenig schulentfremdende Sozialisationseinflüsse (Stichwort: Schulkomposition) bietet, kann die Entwicklung von Schulentfremdung verhindern.

#### Fragestellung

Entsprechend dieses theoretischen Hintergrundes sollen im Rahmen des Beitrags folgende Forschungsfragen behandelt werden:

- a) Welche Profile lassen sich hinsichtlich der Entwicklung der domänenspezifischen Schulentfremdung identifizieren?
- b) Wie unterscheiden sich verschiedene Schüler/innen-Gruppen – charakterisiert entlang der Achsen sozialer Ungleichheit soziale Herkunft, Geschlecht und Migrationshintergrund – innerhalb dieser Profile von Schulentfremdung?
- c) Welche Einflüsse domänenspezifischer Schulentfremdung zeigen sich auf Schuldevianz sowie Schulleistungen?

#### Methode

Die Analysen werden auf Basis quantitativer längsschnittlicher Daten des internationalen mixed-method Forschungsprojekts SASAL – Schulentfremdung in der Schweiz und Luxemburg, einer Kooperation zwischen der Universität Bern und der Universität Luxemburg (2015-2019), vorgenommen. Die Paneldaten wurden in zwei Wellen 2016 und 2017 bei Sekundarschüler/innen in der 7. und 8. Klasse erhoben. Hinsichtlich der verschiedenen Fragestellungen werden Ergebnisse von Clusteranalysen (Domänen der Schulentfremdung) und OLS-Regressionsmodellen (Ungleichheiten, Zusammenhänge mit Schuldevianz und Schulleistungen) präsentiert.

#### Ergebnisse

Die Analysebefunde unterstreichen die Notwendigkeit, spezifische Domänen von Schulentfremdung in den Blick zu nehmen. Vier wesentliche Profile zeigen sich für Luxemburgische Sekundarschüler/innen: Ein in allen Domänen nicht-entfremdetes Profil (ca. 35 Prozent), ein vom Lernen gering entfremdetes Profil (ca. 33 Prozent) sowie domänen-übergreifende Profile. Die letzteren umfassen solche Schüler/innen, die im mittleren Ausmaß vom Lernen und von Lehrpersonen entfremdet sind (ca. 19 Prozent) sowie diejenigen, welche im geringen Ausmaß von Mitschüler/-innen und Lehrpersonen (ca. 13 Prozent) entfremdet sind. Hinsichtlich der Risikogruppen, erweisen sich Jungen als am stärksten von Schulentfremdung hinsichtlich verschiedener Domänen betroffen. Während Entfremdung vom Lernen und Lehrpersonen mit erhöhter Schuldevianz einhergeht, reduziert eine Entfremdung von den Mitschüler/-innen tendenziell das Auftreten devianter Verhaltensweisen. Ähnliche differentielle Wirkungen zeigen sich auch für die Schulleistungen, wobei der Entfremdung vom Lernen wiederum eine besondere Rolle zukommt.

Keywords: Schulentfremdung, Schuldevianz, Schulerfolg, Bildungsungleichheiten

## F18

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H102

### **Transitionen von der Erstausbildung ins Erwerbsleben – Längsschnittanalysen zu Übergängen in Schule, Ausbildung und Arbeitsmarkt vom Schul- bis ins Erwachsenenalter**

Chair(s): **Katja Scharenberg** (Pädagogische Hochschule Freiburg, Deutschland)

DiskutantIn(nen): **Jürgen Baumert** (Max-Planck-Institut für Bildungsforschung)

Übergänge im Bildungssystem (Transitionen) markieren strukturelle Gelenkstellen, an denen sich individuelle Bildungsverläufe verzweigen (Cortina, Baumert, Leschinsky, Mayer & Trommer, 2008). Neben dem Übergang von der Grundschule in die weiterführende Schule (z.B. Maaz, Baumert, Gresch & McElvany, 2010) sind auch Übergänge im späteren Bildungsverlauf wie beispielsweise der Übergang in die Hochschule (z.B. Bornkessel & Asdonk, 2011) oder in die Berufsausbildung und den Arbeitsmarkt (z.B. Pätzold, 2008) von großer Bedeutung, da mit jedem Übergang Entscheidungen getroffen werden, die weitreichende Konsequenzen für individuelle Lebensmöglichkeiten, Bildungsrenditen und gesamtgesellschaftliche Entwicklungen (z.B. steigendes Bildungsniveau der Bevölkerung, Arbeitsmarktintegration und Arbeitslosigkeit) haben. Empirisch wurde wiederholt gezeigt, dass es an solchen Übergängen zur Reproduktion sozialer Ungleichheiten und migrationsbedingten Disparitäten kommt (z.B. Becker, 2011). Auch wenn zur Übergangsthematik bereits eine gut gesicherte empirische Basis vorliegt, besteht die Schwierigkeit einerseits häufig darin, diese Befunde einordnen und vergleichen zu können, da die Datengrundlagen, Anlage der Studien und die Instrumente – und damit die Operationalisierungs- und Analysemöglichkeiten – oftmals sehr unterschiedlich sind. Andererseits eröffnen nur wenige Studien die Möglichkeit längsschnittlicher Analysen aus verschiedenen disziplinären Zugängen unter Berücksichtigung ausleserelevanter Kontrollvariablen.

Das Symposium fokussiert daher verschiedene Übergänge von der Grundschule über die Sekundarstufen 1 und 2, bis in die Hochschule und den Arbeitsmarkt aus soziologischer, erziehungswissenschaftlicher und berufspädagogischer Perspektive. Dabei werden auch verschiedene Subgruppen (z.B. Personen mit Migrationshintergrund) in den Blick genommen. Zentrale gemeinsame Datengrundlage aller Beiträge ist die Studie „Transitionen von der Erstausbildung ins Erwerbsleben“ (TREE). TREE ist in der Schweiz die erste nationale Längsschnittuntersuchung zum Übergang von der Schule ins Erwerbs- und Erwachsenenleben, die auf PISA 2000 aufbaut.

Der erste Beitrag stellt Design, Anlage und Durchführung der TREE-Studie vor und gibt einen Überblick über ausgewählte Befunde zu Übergängen zwischen den Sekundarstufen I und II sowie in den Arbeitsmarkt. Dabei werden vor allem soziokulturelle und geschlechtsspezifische Ungleichheiten thematisiert.

Der zweite Beitrag geht der Frage nach, ob Jugendliche mit und ohne Migrationshintergrund nach dem Ende ihrer Pflichtschulzeit unterschiedliche Bildungsentscheidungen treffen, welche Rolle hierbei die Bildungsaspirationen spielen und ob es Unterschiede zwischen verschiedenen Migrantengruppen bzw. Herkunftsländern gibt.

Der dritte Beitrag analysiert Herkunfts- und Leistungseffekte beim Hochschulzugang und untersucht, inwiefern diese für Absolventinnen und Absolventen mit einer gymnasialen Maturität und für solche mit einer Berufsmaturität unterschiedlich ausfallen.

Der vierte Beitrag untersucht, inwiefern sich durch die veränderte Sozialstruktur und Qualifikation von Personen mit Migrationshintergrund in unterschiedlichen Geburtskohorten ihre Bildungs- und Ausbildungschancen gewandelt haben. Neben den TREE-Daten wird hierbei auch auf Daten der DAB-Panelstudie zu Determinanten der Ausbildungswahl und der Berufsbildungschancen zurückgegriffen.

Abschließend werden die Einzelbeiträge von Prof. Dr. Drs. h.c. Jürgen Baumert diskutiert. Neben Gemeinsamkeiten und Unterschieden werden dabei auch methodische Herausforderungen bei Längsschnittanalysen zu Übergängen und Implikationen der Befunde für die pädagogische Praxis herausgearbeitet.

Zusammenfassend verdeutlichen die Beiträge in diesem Symposium das analytische Potenzial eines multidisziplinären Forschungszugangs zu öffentlich verfügbaren bildungswissenschaftlichen Large-Scale-Datensätzen wie der TREE-Studie (Bergman, Hupka-Brunner, Keller, Meyer & Stalder, 2011; Scharenberg, Hupka-Brunner, Meyer & Bergman, 2016).

#### *Beiträge des Symposiums*

### **Die Längsschnittstudie TREE: Design, Datenbasis und ausgewählte Befunde**

**Thomas Meyer, Sandra Hupka-Brunner**

Universität Bern

Dieser einleitende Beitrag stellt Projektanlage und -kontext von TREE (Transitionen von der Erstausbildung ins Erwerbsleben) vor, gibt einen Überblick über die Struktur und Vielfalt der verfügbaren Daten und leistet eine kurze Synopsis ausgewählter relevanter Befunde.

Projektdesign: TREE ist eine interdisziplinär angelegte schweizerische Multikohortenstudie mit bislang zwei großen, national und (sprach-)regional repräsentativen Stichproben von jungen Menschen, die nach neun Jahren Schulzeit aus der Schulpflicht entlassen werden. Der Stichprobenumfang beim Start des Längsschnitts beträgt 6.343 für die 1. Kohorte (TREE1) und ca. 7.700 für die 2. Kohorte (TREE2). Die Startzeitpunkte der zwei Kohorten liegen 16 Jahre auseinander (Jahre 2000 und 2016). Die Startbefragung (baseline survey) für beide Kohorten bildete ein standardisierter Leistungstest am Ende der obligatorischen Schulzeit. Bei TREE1 war dies der PISA-Test, bei TREE2 ein nationaler standardisierter Leistungstest in Mathematik. Zentraler Untersuchungsgegenstand der Studie sind die nachobligatorischen Ausbildungs- und Erwerbsverläufe. Die Probandinnen und Probanden werden in einem Mixed-Mode-Befragungsdesign (CATI & schriftliche Instrumente) regelmäßig nachbefragt, in den ersten sieben Jahren jährlich, danach in Abständen von drei bis fünf Jahren.

Datenbasis & -struktur: Für die erste Kohorte TREE1 liegen Daten von nunmehr neun Nachbefragungswellen vor, die zwischen 2001 und 2014 durchgeführt wurden. Weitere Nachbefragungen sind geplant. Ausbildungs- und Erwerbsepisoden der Stichprobe werden monatsgenau erfasst, was sehr differenzierte und präzise Verlaufsanalysen ermöglicht. Die Verlaufsdaten sind reich mit Angaben zu Persönlichkeit, sozialem Kontext, Werthaltungen und Präferenzen, Gesundheit, Wohlbefinden, familiärer Situation, allgemeinen Lebensbedingungen etc. kontextualisiert, was vielfältige Modellierungen von Gelingensbedingungen und (Miss-)Erfolgsfaktoren der beobachteten Verläufe erlaubt. Für die 2. Kohorte TREE2 werden ab 2019 erste Daten vorliegen, was erstmals Kohortenvergleiche ermöglichen wird. Die TREE-Daten können als scientific use files

niederschwellig beim schweizerischen Dateninfrastrukturzentrum FORS online bezogen werden. Sie gehören zu den meistgenutzten sozialwissenschaftlichen Daten-Infrastrukturen der Schweiz.

Ausgewählte Befunde: Bislang sind weit über 200 wissenschaftliche Publikationen erschienen, welche auf Analysen der TREE-Daten basieren. Der Beitrag präsentiert Befunde aus drei Analyseschwerpunkten:

Ein erster Schwerpunkt fasst Arbeiten zusammen, die sich mit dem Übergang von der Sekundarstufe I in die Sekundarstufe II beschäftigen. Dieser ist in der Schweiz ähnlich wie in Deutschland stark von der dual-berufsbildenden Struktur des Ausbildungsangebots auf Sekundarstufe II geprägt, aber auch von der im internationalen Vergleich stark segregativ organisierten Sekundarstufe I. Hier geht es unter anderem um die Frage, welche Opportunitätsstrukturen ein so verfasstes Bildungssystem generiert und wie sich diese auf die Bildungsverläufe auswirken.

Ein zweiter Schwerpunkt nimmt die vielfältigen Übergänge in den Arbeitsmarkt in den Blick. Hier weisen die Ergebnisse insgesamt auf einen relativ glatten, reibungslosen Übergang für einen Großteil der TREE-Kohorte (1) hin. Allerdings indizieren vertiefende Analysen u.a. auch, dass der Einstieg in den Arbeitsmarkt über die Zeit eher schwieriger geworden ist, und dass die hohe Segmentierung des Berufsbildungssystems und der mit diesem verzahnten Teil-Arbeitsmärkte auch bedeutsame Risiken bezüglich Nachhaltigkeit und Transferierbarkeit von Qualifikationen bergen.

Ein dritter Forschungsschwerpunkt betrifft die Chancen- und Leistungsgerechtigkeit des schweizerischen Bildungssystems. Eine Vielzahl von TREE-Analyseergebnissen verweist auf einen hohen Einfluss askriptiver Merkmale wie sozialer Herkunft, Geschlecht und Migrationsstatus auf den Bildungserfolg, der in der Summe erhebliche kumulative Effekte generiert und den meritokratischen Anspruch des Bildungssystems stark kompromittiert.

### **Zur Bedeutung von Bildungsaspirationen bei Jugendlichen mit Migrationshintergrund beim Übergang nach dem Ende der Sekundarstufe I**

**Jasper Tjaden<sup>1</sup>, Katja Scharenberg<sup>2</sup>**

<sup>1</sup>Universität Bamberg, <sup>2</sup>Pädagogische Hochschule Freiburg

Theoretischer Hintergrund:

Bildungsvergleichsstudien haben wiederholt Leistungs Nachteile für Schülerinnen und Schüler mit Migrationshintergrund aufgezeigt (Rauch, Mang, Härtig & Haag, 2016; Schwippert, Wendt & Tarelli, 2012; Wendt, Schwippert & Stubbe, 2016). Auch für Übergänge an Schnittstellen im Bildungssystem zeigten sich migrationsbedingte Bildungsbenachteiligungen anhand von ungünstigeren schulischen Platzierungen auf anregungsärmere Schulformen (Alba, Handl & Müller, 1994; Kristen, 2002) oder höheren Risiken, nicht in eine zertifizierende postobligatorische Ausbildung auf der Sekundarstufe II einzumünden (Sacchi, Hupka-Brunner, Stalder & Gangl, 2011). Mit Blick auf schulische Übergänge lassen sich solche Bildungsbenachteiligungen auf Leistungsunterschiede je nach sozialer Herkunft („primäre Herkunftseffekte“) und ein schichtspezifisches, leistungsunabhängiges Entscheidungsverhalten („sekundäre Herkunftseffekte“) zurückführen (Boudon, 1974). Daneben können auch Informationsdefizite bezüglich Anforderungen und institutionellen Hürden im Schulsystem (Relikowski, Schneider & Blossfeld, 2010) oder unterschiedliche Kosten-Nutzen-Abwägungen (Erikson & Jonsson, 1996) für migrationspezifische Unterschiede im Entscheidungsverhalten bei der Wahl von Bildungsgängen bedeutsam sein. Aus bildungssoziologischer Perspektive wird bei der Vorhersage von Bildungsergebnissen und der Erklärung sozialer Determinanten des Bildungserfolgs häufig dem Konzept der Bildungsaspirationen (Stocké, 2009) ein wichtiger Stellenwert zugesprochen. Für Schülerinnen und Schüler mit Migrationshintergrund wurde gezeigt, dass sie durchschnittlich höhere Bildungsaspirationen haben als Einheimische: Sie befinden sich zwar häufig in niedrigen sozialen Positionen in der Aufnahmegesellschaft, ihre Familien erwarten für sie jedoch eine „Aufwärtsmobilität“ und schätzen hohe Bildung als wichtigsten Weg ein, um diese zu verwirklichen (Becker, 2010). Inwiefern Unterschiede in den Bildungsaspirationen zwischen Jugendlichen mit und ohne Migrationshintergrund auch über die Pflichtschulzeit hinaus für Bildungsverläufe und Übergänge über die Schulzeit hinaus bedeutsam sind, ist jedoch bislang weitgehend unerforscht.

Fragestellung:

Unser Beitrag geht folgenden Forschungsfragen nach: 1.) Unterscheiden sich Jugendliche mit und ohne Migrationshintergrund in ihren Bildungsentscheidungen beim Übergang nach dem Ende der Pflichtschulzeit? 2.) Werden ggfs. nachweisbare Unterschiede in den Bildungsentscheidungen durch unterschiedliche Ausprägungen in den Bildungsaspirationen mediiert? 3.) Unterscheidet sich die Bedeutung höherer oder niedrigerer Bildungsaspirationen zwischen verschiedenen Migrantengruppen?

Methode:

Datengrundlage ist die Schweizer Längsschnittstudie „Transitionen von der Erstausbildung ins Erwerbsleben“ (TREE, Kohorte 1). Die Analysetichprobe umfasst 5528 Jugendliche im ersten Jahr nach Schulaustritt.

Zur Vorhersage des Ausbildungsstatus im ersten Jahr nach Ende der Pflichtschulzeit (Berufsbildung vs. Allgemeinbildung vs. Nicht in Ausbildung/Zwischenlösung) wurden individuelle (Alter, Geschlecht, soziale Herkunft), schul- und ausbildungsbezogene (Schulnoten, PISA-Lesekompetenz, besuchter Schultyp in der Sekundarstufe I) sowie soziogeographische Merkmale (Sprachregion) kontrolliert. Effekte der Bildungsaspirationen (Frage: „Was meinst du, welchen Beruf wirst du mit 30 Jahren haben?“; offene Antworten kodiert in ISEI [Ganzeboom, De Graff & Treiman, 1992]) wurden in Mediationsanalysen für nicht-lineare genestete Analysemodelle (Karlson, Holm & Breen, 2012) in STATA 14 geschätzt. Fehlende Werte wurden durch multiple Imputationen (Van Buuren, 2012) ersetzt. Die Daten wurden gewichtet (Sacchi, 2013), um Verzerrungen auszugleichen, die sich aus dem Stichprobenschwund im Längsschnitt ergeben.

Ergebnisse:

In multivariaten Analysen zeigen die durchschnittlichen Marginaleffekte (AME), dass Jugendliche mit Migrationshintergrund eine signifikant niedrigere Wahrscheinlichkeit haben, in eine berufliche Grundbildung (AME=0.420,  $p<.001$ ) einzusteigen als in der Schweiz geborene Personen (AME=0.516). Sie wählen dagegen häufiger einen allgemeinbildenden Ausbildungsgang der Sekundarstufe II (AME=0.328,  $p<.001$ ) als Personen ohne Migrationshintergrund (AME=0.246). Dieser „ethnic choice effect“ wird zu 30 % durch Unterschiede in den Bildungsaspirationen erklärt, wenngleich migrationsbedingte Unterschiede in den Übergangswahrscheinlichkeiten bestehen bleiben (AME=0.056,  $p<.05$ ).

Für beide Ausbildungsoptionen zeigen sich die deutlichsten Unterschiede für Personen aus Spanien/Portugal und dem Balkan/der Türkei im Vergleich zu in der Schweiz geborenen Jugendlichen. Bei Personen, die sich ein Jahr nach Schulaustritt nicht in Ausbildung oder in einer Zwischenlösung befinden, ergeben sich keine migrationspezifischen Benachteiligungen.

Für die Forschung zur Übergangsthematik liefert unsere Studie insofern wichtige Erkenntnisse, als mit dem Übergang nach dem Ende der Sekundarstufe I ein weiterer wichtiger Verzweigungspunkt in der Bildungslaufbahn junger Erwachsener



fokussiert wird. Darüber hinaus werden mit der Analyse von Bildungsaspirationen wichtige Prozessindikatoren bei der Entstehung migrationsbedingter Bildungsungleichheiten beleuchtet.

## **Transitionen von der Primarschule in die Sekundarstufe I und in die Erstausbildung – Längsschnittdaten zu den Bildungs- und Ausbildungschancen von Migranten in unterschiedlichen Geburtskohorten**

**Rolf Becker, David Glauser**  
Universität Bern

Theoretischer Hintergrund:

Die Sozialintegration von Migranten über die Partizipation an Institutionen der Aufnahmegesellschaft wie Bildungssystem und Arbeitsmarkt ist eine der bedeutsamen Herausforderungen in der Schweiz. Insbesondere für junge schulpflichtige Migranten ist die strukturelle Assimilation in das Bildungssystem der Schlüssel zur sozialen Integration (Kalter, 2005). Wenn sie anerkannte Bildungsabschlüsse erwerben und anschließend in den Arbeitsmarkt eintreten können, dann trägt diese strukturelle Assimilation auch zur Systemintegration gesellschaftlicher Ordnungen bei (Esser, 2001). Je besser Zugewanderte für anspruchsvolle berufliche Tätigkeiten mit hohen Qualifikationsanforderungen ausgebildet sind, desto eher können sie ebenso wie die ohnehin privilegierten Einheimischen zur Systemintegration beitragen. Die Fakten sind für die Schweiz andere. Migranten sind – bei großen Unterschieden zwischen einzelnen Nationalitäten, Migrationswellen und Migrationshintergründen – im Bildungs- und Beschäftigungssystem zumeist im Nachteil gegenüber den Einheimischen, obgleich die zugewanderten Familien höhere Bildungsaspirationen aufweisen.

Auch in der Deutschschweiz werden Nachteile von Migranten in der Schule und beim Zugang zur beruflichen Ausbildung in der Regel größtenteils durch die sozioökonomischen Ressourcen des Elternhauses – bemessen an dessen Bildungsniveau und Klassenlage – erklärt (Beck et al., 2010; Becker et al., 2013; Glauser, 2015). Bei Kontrolle ihrer sozialen Herkunft und Leistung werden Vorteile von Migranten („ethnic premia“) beim Übergang auf das Gymnasium auf deren höhere Bildungsaspirationen zurückgeführt, während ihre Schwierigkeiten („ethnic penalties“) beim Übergang in die Berufsausbildung auf unzureichenden Leistungen und Sprachproblemen basieren (Kristen & Dollmann, 2010). Im Aggregat gesehen erklären sich die unterschiedlichen Bildungschancen von Einheimischen und Migranten durch ihre Sozialstrukturen. In der Schweiz haben sich infolge von Migrationswellen die Sozialstrukturen der Migranten im Bildungssystem gewandelt. Während beispielsweise für die um 1985 Geborenen vor allem Kinder von weniger gebildeten Eltern aus den Balkanstaaten das Gros der Migranten im deutschschweizerischen Bildungssystem stellten, waren es etwa für die um 1997 Geborenen neben den Kindern und Jugendlichen aus Deutschland und anderen westlichen EU-Staaten auch Migranten aus Südosteuropa, der Türkei und aus nichteuropäischen Krisengebieten (Glauser, 2015). Da verstärkt besser beruflich und akademisch gebildete Migrantengruppen eingewandert sind, dürfte damit ein Wandel der Bildungschancen für ihre Kinder einhergehen und sich die Einflüsse der ethnischen Herkunft auf deren Bildungserfolge und -wege verschieben (Cattaneo & Wolter, 2012).

Fragestellung:

Haben sich die Bildungs- und Ausbildungschancen von Migranten tatsächlich in der Kohortenabfolge gewandelt? Inwieweit kann dieser Wandel im deutschsprachigen Teil der Schweiz durch die veränderte Sozialstruktur und Qualifikation der Migranten in unterschiedlichen Geburtskohorten erklärt werden?

Methode:

Die empirischen Analysen basieren auf Paneldaten von TREE (2016) und von DAB (Glauser & Becker, 2016). Während TREE die um 1985 Geborenen umfasst, wurden in DAB die um 1997 Geborenen zu ihrem Bildungsweg befragt. Aus Gründen der Vergleichbarkeit beider Datensätze beschränkt sich die Analyse auf den deutschsprachigen Teil der Schweiz. Die Übergänge in die Sekundarstufe I (Vorgymnasium / erweiterte Anforderungen vs. andere Schultypen) und Sekundarstufe II (Berufsausbildung vs. Gymnasium) in Abhängigkeit von sozialer und ethnischer Herkunft werden anhand von Logit-Regressionen modelliert.

Ergebnisse:

Insgesamt ist ein Rückgang migrationsbedingter Nachteile beim Bildungserwerb festzustellen, wobei nach wie vor ausgeprägte Unterschiede zwischen Migrationsgruppen bestehen. „Ethnic premia“ der jüngeren Kohorte (DAB) bestehen sowohl beim besuchten Schultyp auf der Sekundarstufe I wie auch in Bezug auf den Beginn einer Mittelschule für Jugendliche aus nicht südlichen EU-Staaten, welche zudem Vorteile gegenüber Einheimischen aufweisen. Diese Vorteile stehen in Zusammenhang mit der sozial selektiven Einwanderung hochqualifizierter Personen im Zuge der Umsetzung der Personenfreizügigkeit. „Ethnic penalties“ bestehen für beide Kohorten für Kinder der Balkanstaaten, der Türkei und Portugal in Bezug auf den besuchten Schultyp auf der Sekundarstufe I. Besuchen Kinder dieser Herkunft hingegen einen Schultyp mit erweiterten Anforderungen, dann weisen sie keine Nachteile beim Übergang an Mittelschulen auf.

## **Herkunfts- und Leistungseffekte beim Hochschulzugang in der Schweiz**

**Stephan Schumann**  
Universität Konstanz

Theoretischer Hintergrund:

Ein zentrales bildungspolitisches Ziel in demokratischen Gesellschaften ist es, jungen Menschen unabhängig von ihrer Herkunft den leistungsgerechten Zugang zu Bildung und somit zu langfristiger sozialer Teilhabe zu ermöglichen (Hillmert, 2007). Im deutschsprachigen Raum gehören die teils erheblichen sozialen Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs jedoch spätestens seit Veröffentlichung der Ergebnisse aus PISA 2000 praktisch zum Alltagswissen (Baumert & Schümer, 2001; Becker & Lauterbach, 2007). Der vorliegende Beitrag fokussiert vor diesem Hintergrund auf den Hochschulzugang in der Schweiz unter Bezugnahme auf Boudons Konzept primärer und sekundärer Herkunftseffekte (Boudon, 1974). Wie in vielen anderen Ländern auch hat in der Schweiz in den letzten Dekaden eine institutionelle Ausweitung und Öffnung der Hochschulzugangswege und des Hochschulbereichs stattgefunden (z.B. durch die Einrichtung von Berufsmaturitätsschulen und von Fachhochschulen). Ein Ziel dieser Öffnung ist es, den Hochschulzugang für breitere Schichten der Bevölkerung zu gewährleisten. Zugleich weist das Schweizer System einige Besonderheiten auf, die eine ungeprüfte Übertragung z.B. von Befunden aus Deutschland zumindest fraglich erscheinen lassen (z.B. NC-freier Hochschulzugang für alle Schulabgehenden mit einer gymnasialen Maturität).

Fragestellung:

Die Studie geht der Frage nach, ob sich auf dem Weg in die Schweizer Hochschulen Leistungs- und Herkunftseffekte zeigen und ob diese für Absolventinnen und Absolventen mit einer gymnasialen Maturität und für solche mit einer Berufsmaturität unterschiedlich ausfallen.

Methode:

Die Sekundärdatenanalyse erfolgte auf der Basis der nationalen PISA 2000/TREE-Längsschnittdaten (n=2.123 Maturandinnen und Maturanden; davon 47% Frauen sowie 61% mit gymnasialer Maturität vs. 39% mit Berufsmaturität) und wurde dabei mit den im Zeitraum von 2000 (PISA-Studie) bis 2007 (TREE, T7) im Jahresabstand erhobenen Daten durchgeführt. Die Erhebungen erfolgten mittels standardisierter schriftlicher Fragebögen sowie ab 2005 zusätzlich mittels computergestützter Telefoninterviews.

Die soziale Herkunft der Jugendlichen wird mittels des von Ganzeboom et al. (1992) entwickelten International Socio-Economic Index of Occupational Status (ISEI) abgebildet. In den Analysen wird der jeweils höchste Index von Mutter oder Vater verwendet (HISEI). Zur Abschätzung der schulischen Leistung wird auf die in PISA 2000 erfasste Lesekompetenz zurückgegriffen (Baumert et al., 2001).

Ergebnisse:

Mehr als die Hälfte der Absolventinnen und Absolventen mit gymnasialer Maturität beginnt direkt im Anschluss an den Gymnasialbesuch ein Hochschulstudium. Nach einem Jahr tritt mit 28% nochmals ein großer Teil in eine Hochschule ein, so dass zu diesem Zeitpunkt insgesamt 83% studieren. Stark verschieden davon sind die Übergangsprozesse der Berufsmaturandinnen und Berufsmaturanden. Bei rund einem Fünftel kann ein Sofortübertritt beobachtet werden. Im zweiten Jahr nehmen nochmals 15% ein Studium auf.

Die Analysen zeigen einen theoriekonformen Zusammenhang zwischen der sozialen Herkunft und der Lesekompetenz der Jugendlichen (primärer Effekt). Eine höhere Lesekompetenz geht dabei mit deutlich höheren Chancen einher, eine gymnasiale Maturität anstelle einer Berufsmaturität zu erwerben. Unter Kontrolle des Leistungseffekts kann jedoch auch ein starker Einfluss der sozialen Herkunft beobachtet werden. Unter Bezug auf die HISEI-Quartile ist die Chance auf den Erwerb einer gymnasialen Maturität für Jugendliche aus einer hohen Sozialschicht fast dreimal so groß wie für Jugendliche aus einer niedrigen sozialen Schicht (Hinweis auf sekundären Effekt). Dieser Effekt kann als ein Ergebnis selektiver Bildungsentscheidungen während der obligatorischen Schulzeit und am Übergang in die Sekundarstufe II interpretiert werden.

Beim Hochschulzugang kann für die Gruppe der Berufsmaturandinnen und Berufsmaturanden beobachtet werden, dass Jugendliche aus höheren sozialen Schichten eher in die Hochschule eintreten. Da dieser Effekt auch unter Kontrolle der Lesekompetenz nachweisbar ist, kann er als sekundärer Herkunftseffekt interpretiert werden. Ein primärer Effekt lässt sich ebenfalls identifizieren: Berufsmaturandinnen und Berufsmaturanden mit höherer Lesekompetenz beginnen häufiger ein Studium. Für die Absolventinnen und Absolventen mit gymnasialer Maturität sind am Hochschulzugang Effekte der Leseleistung nachweisbar. Darüber hinausgehende Herkunftseffekte zeigen sich dagegen nicht in systematischer Art und Weise.

## F19

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H001

### Prädiktoren des Erfolgs in der Eingangsphase naturwissenschaftlich-technischer Studiengänge

Chair(s): **Elke Sumfleth** (Universität Duisburg-Essen, Deutschland), **Detlev Leutner** (Universität Duisburg-Essen, Deutschland)

DiskutantIn(nen): **Markus Dresel** (Universität Augsburg, Deutschland)

Im Symposium werden ausgewählte Projektergebnisse der Hauptstudie der DFG-Forschergruppe (FOR 2243) „Akademisches Lernen und Studienerfolg in der Eingangsphase von naturwissenschaftlich-technischen Studiengängen“ (ALSTER) vorgestellt, die im Wintersemester 2016/17 und im Sommersemester 2017 durchgeführt wurde. Von den untersuchten fachspezifischen und fächerübergreifenden Bedingungen des Studienerfolgs werden in vier Vorträgen die Parameter Studienzufriedenheit, Ressourcenmanagement, visuelles Modellverständnis und fachliches Vorwissen thematisiert.

Im ersten Beitrag fokussiert Jens Fleischer die zentrale Rolle der Studienzufriedenheit als einer wesentlichen Variable im ALSTER-Rahmenmodell. In diesem Rahmenmodell wird angenommen, dass der Einfluss unterschiedlicher Prädiktoren auf die weiteren Studienerfolgskriterien teilweise durch die Studienzufriedenheit mediiert wird. Außerdem wird der Zusammenhang zwischen Studienzufriedenheit und Studienabbruchneigung untersucht, und Unterschiede zwischen den verschiedenen Studiengängen (Mono-Bachelor: Biologie, Chemie, Physik, Bauingenieurwesen) in den Effekten werden überprüft.

Im zweiten Beitrag geht es um Ressourcenmanagement. Mit Hilfe eines neu entwickelten Situational-Judgement-Instruments zur Erfassung von Kompetenzen und Strategiedefiziten im Bereich des Ressourcenmanagements untersucht Julia Waldeyer die Bedeutung dieser Variable für den Studienerfolg, auch in Abhängigkeit von unterschiedlichen Leistungsanforderungen (Vergleich der Studiengänge Erziehungswissenschaft und Bauingenieurwesen), und sie geht der Frage nach, ob Ressourcenmanagementkompetenz als Mediator zwischen Gewissenhaftigkeit und Studienerfolg fungiert.

Im dritten Beitrag steht die Bedeutung des visuellen Modellverständnisses für den Studienerfolg in den Studiengängen Chemie und Bauingenieurwesen im Vordergrund. Das Verstehen von fachspezifischen Modellen ist eine charakteristische Anforderung in naturwissenschaftlich-technischen Studienfächern. Thomas Dickmann hat einen Test zur Erfassung des visuellen Modellverständnisses entwickelt und zeigt, dass dieses ein bedeutsamer Prädiktor für den Klausurerfolg und das erworbene Fachwissen ist. Es wirkt nicht ausschließlich direkt auf das Fachwissen, sondern auch als Mediator auf den Zusammenhang zwischen fachlichem Vorwissen und dem erworbenen Fachwissen.

Im vierten Beitrag gibt Daniel Averbek einen Überblick über den Einfluss von kognitiven und affektiven Variablen auf den Studienerfolg in Chemie. Die Ergebnisse zeigen die hohe Bedeutung der kognitiven Faktoren Vorwissen, mathematische Fähigkeiten und Abiturnote. Detailanalysen machen deutlich, dass der Studienerfolg im Modul Allgemeine Chemie, das an fast allen Universitäten den Kern des Studiums im ersten Semester ausmacht, großen Einfluss auf den Studienerfolg in den später folgenden Teildisziplinen der Chemie hat. Dadurch kommt diesem Modul im Studienverlauf eine richtungsweisende Bedeutung zu.

Die vier Beiträge des Symposiums werden abschließend von Markus Dresel zusammenfassend diskutiert.

#### Beiträge des Symposiums

### Vorhersage des Studienerfolgs in MINT-Fächern – Studienzufriedenheit und Studienabbruchneigung

**Jens Fleischer<sup>1</sup>, Daniel Averbek<sup>1</sup>, Elke Sumfleth<sup>1</sup>, Detlev Leutner<sup>2</sup>, Matthias Brand<sup>1</sup>**  
<sup>1</sup>Universität Duisburg-Essen, <sup>2</sup>Universität Duisburg-Essen

Vor dem Hintergrund der vergleichsweise geringen Studierendenzahlen in den MINT-Studiengängen stellen die vergleichsweise hohen Studienabbruchquoten in diesen Fächern ein Problem dar (Heublein et al., 2012; OECD, 2011). Die Identifizierung von Prädiktoren des Studienerfolgs hat somit gerade in den MINT-Studiengängen eine besondere Relevanz. Die zahlreichen Einzelbefunde zu Prädiktoren des Studienerfolgs wurden in verschiedenen Modellen zusammengefasst (z. B. Heublein & Wolter, 2011). Bisher fehlen jedoch systematische fachspezifische Untersuchungen, die neben einer querschnittlichen Betrachtung auch eine längsschnittliche Modellierung des Studienerfolgs in den MINT-Fächern ermöglichen.

Ziel des hier beschriebenen Teilprojektes der Forschergruppe „ALSTER“ ist es daher, übergreifend eine systematisch angelegte fachspezifische Modellierung des Studienerfolgs zu entwickeln. Hierzu wurde ein Rahmenmodell erstellt, welches auf spezifische Anforderungen naturwissenschaftlich-technischer Studiengänge fokussiert. Die Studienzufriedenheit spielt hierbei eine zentrale Rolle: Zum einen stellt sie ein Studienerfolgskriterium dar (Machado et al., 2011; Trapmann et al., 2007), zum anderen steht sie sowohl mit Studienleistungen als auch dem Verbleib im Studium in Zusammenhang (Bean & Bradley, 1986; Wiers-Jenssen et al., 2002). Im Rahmenmodell wird daher angenommen, dass der Einfluss unterschiedlicher Prädiktoren auf die weiteren Studienerfolgskriterien teilweise durch die Studienzufriedenheit mediiert wird.

Gegenstand des Beitrags ist die Untersuchung der prädiktiven Validität sowohl kognitiver und als auch nicht kognitiver Variablen für die Vorhersage unterschiedlicher Aspekte der Studienzufriedenheit in der Studieneingangsphase naturwissenschaftlich-technischer Studiengänge. Ferner wird geprüft, inwiefern die Studienzufriedenheit bei der Vorhersage der Studienabbruchneigung eine Mediatorvariable darstellt, und inwiefern sich die Effekte zwischen den untersuchten Studiengängen unterscheiden.

Für die Hauptstudie, die im Wintersemester 2016/2017 an zwei deutschen Universitäten durchgeführt wurde, liegen Daten vom Anfang und Ende des ersten Semesters von Bachelor-Studierenden der Fächer Chemie (n = 275), Biologie (n = 162), Physik (n = 106) und Bauingenieurwesen (n = 208) vor. Die Studienzufriedenheit wurde zusätzlich Mitte des Semesters erhoben.

Zur Vorhersage der Studienzufriedenheit zum Ende des Semesters wurde ein Mehrgruppenpfadmodell spezifiziert, in dem das Studieninteresse (Schiefele et al., 1993;  $\beta = .34$ ;  $p < .001$ ), das Engagement im Studium (Schaufeli et al., 2006;  $\beta = .27$ ;  $p < .001$ ) und die Mathematikkompetenz ( $\beta = .15$ ;  $p < .001$ ) in allen untersuchten Studiengängen signifikante Effekte auf die Zufriedenheit mit den Studieninhalten (Westermann et al., 1996) aufwiesen. Die Erwartungskomponente der Motivation zeigte hingegen lediglich in Chemie und Biologie signifikante Effekte (Kosovich et al., 2015;  $\beta = .24$ ;  $p < .001$ ). Bei der Vorhersage der Zufriedenheit mit der Bewältigung der Studienanforderungen (Westermann et al., 1996) erwiesen sich die Kostenkomponente der Motivation (Kosovich et al., 2015;  $\beta = -.29$ ;  $p < .001$ ) und die emotionale Erschöpfung (Maslach et al., 1996;  $\beta = -.29$ ;  $p < .001$ ) als signifikante Prädiktoren in allen untersuchten Studiengängen. Die Mathematikkompetenz zeigte hingegen lediglich in

Physik einen signifikanten Effekt ( $\beta = .22$ ;  $p < .001$ ) während die Subskala Neurotizismus des NeoFFI in allen untersuchten Studiengängen einen signifikanten Effekt zeigte (Borkenau & Ostendorf, 1993;  $\beta = -.17$ ;  $p < .001$ ), mit Ausnahme der Physik.

Bei der Vorhersage der Studienabbruchneigung (Fellenberg & Hannover, 2006) zum Ende des Semesters zeigten sich neben direkten auch signifikante indirekte Effekte mediiert über die Zufriedenheit mit den Studieninhalten zur Mitte des Semesters für die Mathematikkompetenz ( $\beta_{\text{Bio}} = -.04$ ;  $\beta_{\text{CH/PH/Bau}} = -.01$ ), das Studieninteresse ( $\beta_{\text{Bio}} = -.19$ ;  $\beta_{\text{CH/PH/Bau}} = -.06$ ), die Erwartungskomponente der Motivation ( $\beta_{\text{Bio}} = -.12$ ;  $\beta_{\text{CH/PH/Bau}} = -.04$ ) sowie das Engagement im Studium ( $\beta_{\text{Bio}} = -.21$ ;  $\beta_{\text{CH/PH/Bau}} = -.07$ ), die für Biologie größer ausfielen als für die anderen untersuchten Studiengänge (für alle Koeffizienten gilt  $p < .01$ ).

In weiteren Analysen sollen die zusätzlichen Kriterien für Studienerfolg (Fachwissen und Klausurnoten) sowie ein zusätzlicher Messzeitpunkt (Ende des zweiten Semesters) in den Blick genommen werden.

## **Die Bedeutung von Ressourcenmanagement für den Studienerfolg**

**Julia Waldeyer<sup>1</sup>, Jens Fleischer<sup>1</sup>, Detlev Leutner<sup>1</sup>, Joachim Wirth<sup>2</sup>**

<sup>1</sup>Universität Duisburg-Essen, <sup>2</sup>Ruhr-Universität Bochum

Ein Studium stellt Lernende vor umfangreiche Herausforderungen im Bereich des selbstregulierten Lernens, insbesondere im Bereich des Ressourcenmanagements. Studierende müssen weitgehend eigenständig planen wann, wo, wie und mit welchen Strategien und Hilfs-mitteln sie lernen. Sie müssen ihren Lernprozess und ihren Lernfortschritt selbst überwachen und gegebenenfalls nachjustieren – und dies alles unter, im Vergleich zum Lernen in der Schule, erschwerten Bedingungen. Betrachtet man hierzu die einschlägige empirische Lernstrategieliteratur, werden zwei Punkte deutlich. Zum einen nehmen Ressourcenmanagementstrategien eine zentrale Rolle für die Bewältigung von akademischen Lernanforderungen – und damit letztlich für die Vorhersage von Lernerfolg – ein (z. B. Schiefele et al., 2003; Richardson, Abraham & Bond, 2012). Zum anderen mangelt es an Instrumenten, die Ressourcenmanagementkompetenzen valide diagnostizieren (vgl. Spörer & Brunstein, 2006). Um künftig untersuchen zu können, welchen Beitrag Ressourcenmanagement tatsächlich zur Vorhersage von Studienerfolg leistet, wurde ein neues Instrument entwickelt und validiert, das auf dem Situational-Judgement-Ansatz basiert und Ressourcenmanagementkompetenz in Form von konditionalem Strategiewissen als auch in Form von Defiziten bei der Strategieanwendung erfasst. Das validierte Instrument wurde im Rahmen einer Längsschnittstudie der Forschergruppe ALSTER über die beiden ersten Studiensemester zu insgesamt vier Messzeitpunkten administriert. Dabei wurden unter anderem folgenden Forschungsfragen adressiert:

FF1: Welche Bedeutung hat Ressourcenmanagementkompetenz für die Vorhersage von Studienerfolg im ersten Studiensemester?

FF2: Gibt es zwischen Studienfächern mit unterschiedlichen Leistungsanforderungen (Bauingenieurwesen/Erziehungswissenschaft) Unterschiede in der Bedeutsamkeit von Ressourcenmanagementkompetenz für die Vorhersage von Studienerfolg?

FF3: Fungiert Ressourcenmanagementkompetenz als Mediator zwischen Gewissenhaftigkeit und Studienerfolg?

Das Instrument zeigt, basierend auf einer Stichprobe von  $N = 380$  Erstsemesterstudierenden der Erziehungswissenschaft (88 % weiblich) und des Bauingenieurwesens (35 % weiblich), hohe psychometrische Qualität. So konnte die theoretisch angenommene fünf-faktorielle Struktur der Daten (Zeitmanagement, Anstrengungsregulierung, Motivation, Hilfesuchen, Lernumgebung) mittels konfirmatorischer Faktorenanalysen zu Beginn des ersten Semesters empirisch bestätigt werden ( $\lambda = .42-.72$ ;  $p < .001$ ; CFI: .977; TLI: .975; RMSEA: .027;  $\chi^2 = 734.996$ ,  $p < .001$ ), wobei sich das Modell für die weiteren Messzeitpunkte und für beide Fächer als invariant erwies. Zudem wiesen die einzelnen Faktoren, mit Ausnahme des Faktors Lernumgebung ( $\alpha = .57$ ), zufriedenstellende Reliabilitäten auf ( $\alpha = .72-.76$ ), die über die weiteren Messzeitpunkte hinweg tendenziell zunahmen. Und schließlich zeigte sich für die Bauingenieure, dass die Faktoren des Ressourcenmanagements auch über kognitive Grundfähigkeiten und die Abiturdurchschnittsnote hinaus, Studienerfolg in Form von Fachwissen und Klausurnoten im ersten Semester signifikant vorhersagt (exemplarisch für Anstrengungsregulation und Fachwissen:  $R^2 = .41$ ,  $p < .001$ , inkrementelles  $R^2(\text{Anstrengungsregulierung}) = .01$ ,  $p < .05$ ; exemplarisch für Anstrengungsregulierung und Klausurnoten:  $R^2 = .33$ ,  $p < .001$ , inkrementelles  $R^2(\text{Anstrengungsregulierung}) = .10$ ,  $p < .05$ ). Erwähnenswert ist hierbei, dass vier der fünf Faktoren ebenfalls gegenüber der LIST-Skala Ressourcenmanagement eine signifikante Zunahme der aufgeklärten Varianz bei der Vorhersagen der Noten und des Fachwissens im ersten Studiensemester der Bauingenieure verzeichneten. Darüber hinaus zeigten erste Analysen, dass der Effekt von Gewissenhaftigkeit auf den Studienerfolg der Bauingenieure (Klausurnoten) – wie erwartet – zumindest teilweise über die Faktoren des Ressourcenmanagements mediiert wird. Die indirekten Effekte sind klein (exemplarisch für Anstrengungsregulierung:  $\beta = -.09$ ), werden aber statistisch signifikant ( $p < .05$ ). Im Rahmen des Symposiums werden diese und weitere Ergebnisse vorgestellt.

## **Studienerfolg und visuelles Modellverständnis – Eine Längsschnittstudie**

**Thomas Dickmann<sup>1</sup>, Maria Opfermann<sup>2</sup>, Stefan. Rumann<sup>1</sup>, Elmar Dammann<sup>3</sup>, Martin Lang<sup>1</sup>, Carsten Schmuck<sup>1</sup>**

<sup>1</sup>Universität Duisburg-Essen, <sup>2</sup>Ruhr-Universität Bochum, <sup>3</sup>Hochschule Furtwangen

In nationalen und internationalen Studien zum Studienabbruch (Chen, 2013; Heublein, 2017) wird die hohe Drop-Out-Quote, insbesondere in MINT-Fächern, bemängelt. Um diesem „Schwund“ entgegenzuwirken, sollte eine Aufgabe naturwissenschaftsdidaktischer Forschung sein, Prädiktoren für den (mangelnden) Studienerfolg herauszufinden und deren spezifische Rollen sowie Zusammenhänge zu untersuchen.

In der Lehr-Lern-Forschung, speziell im Bereich des multimedialen Lernens, werden in diesem Zusammenhang Visualisierungen als besonders lernförderlich angesehen und sind aus naturwissenschaftlichen Lehrmaterialien nicht mehr wegzudenken. So schreiben eine Vielzahl von Autoren (Coll & Lajum, 2011; Harrios & Treagust, 2000; Ramadas, 2009; Wu & Shah, 2004) Visualisierungen eine besonders hohe Bedeutung zu, da diese substantiell für nachhaltiges Lernen in der Chemie und den Ingenieurwissenschaften seien. Gleichzeitig sollten Visualisierungen allerdings nach aktuellen Erkenntnissen zur Informationsverarbeitung (z.B. Cognitive Load Theory; Sweller, Ayres & Kalyuga, 2011) sowie gängigen Design-Prinzipien (z.B. Cognitive Theory of Multimedia Learning; Mayer, 2009) gestaltet und mit dem zugehörigen Text kombiniert werden (z.B. Integratives Modell des Text- und Bildverständnisses; Schnotz, 2005).

Die bislang noch ungeklärte Frage ist aber weiterhin, dass trotz der in vielerlei Hinsicht nachgewiesenen Lernförderlichkeit von Visualisierungen noch nicht geklärt werden konnte, welche Arten von Visualisierungen, unter welchen Umständen und für welchen Lernertyp geeignet sind (Höffler, Schmeck & Opfermann, 2013). Diese Forschungslücke möchte die vorliegende Studie bearbeiten. Sie befasst sich mit der Fähigkeit von Lernenden, unter der Berücksichtigung von domänenspezifischen

Eigenheiten, relevante Informationen aus unterschiedlichen Visualisierungen herauszufiltern, zu übersetzen und aufeinander zu beziehen. Diese Fähigkeit bezeichnen wir als visuelles Modellverständnis.

In der hier vorgestellten Studie stehen folgende Forschungsfragen im Blickpunkt:

- Welche Eigenschaften der Lernenden sind Prädiktoren für visuelles Modellverständnis?
- Kann visuelles Modellverständnis Studienerfolg (zumindest partiell) vorhersagen?
- Gibt es studienfachbezogene Unterschiede bezüglich des Einflusses von visuellem Modellverständnis?

Zur Beantwortung dieser Forschungsfragen wurde zunächst ein Instrument zur Erfassung des visuellen Modellverständnisses entwickelt. Die Items für diesen Test wurden auf Basis einer umfassenden Lehrbuchanalyse entworfen (Dickmann et al., 2015) und sind in drei gleichgroße Teilbereiche (chemiespezifisches, allgemeines und ingenieurwissenschaftliches visuelles Modellverständnis) untergliedert. Das finale Instrument beinhaltet 33 Items (11 je Bereich) und besitzt ein gute interne Konsistenz (Cronbach's  $\alpha = .853$ ).

Insgesamt nahmen 248 Studierenden der Chemie (weiblich: 40,3%; Alter: 20,9 Jahre) und 188 Studierende der Ingenieurwissenschaften (weiblich: 35,1%; Alter: 20,8 Jahre) an der Studie teil. Die drei Messzeitpunkte befanden sich am Anfang und am Ende des ersten Semesters sowie am Ende des zweiten Semesters. Neben dem visuellen Modellverständnis wurden fachspezifische Leistungstests durchgeführt, und weitere personenbezogene Variablen, wie zum Beispiel kognitive Fähigkeiten (verbal und figural), räumliches Vorstellungsvermögen und Motivation mithilfe umfassender Testbatterien erhoben.

Die Ergebnisse der gerade abgeschlossenen Hauptstudienhebung belegen zum einen die Reliabilität des visuellen Modellverständnistests zu allen Messzeitpunkten. Ferner sind konvergente und diskriminante Validitätskriterien erfüllt. Gleichzeitig zeigen Regressions- und Mediationsanalysen, dass visuelles Modellverständnis einen Prädiktor für Studienerfolg (erhoben über Klausurnoten und Fachleistungstests) darstellt. So ist visuelles Modellverständnis in multiplen Regressionsanalysen zur Varianzaufklärung der Leistung in den Fachtests hinter dem Vorwissen der zweitstärkste Prädiktor (Chemie:  $\beta = .270$ ; Ingenieurwissenschaften:  $\beta = .328$ ). Darüber hinaus ist die Gesamtskala des visuellen Modellverständnistests in der Lage, die Klausurnote in der Chemie ( $\beta = .456$ ,  $p < .001$ ) und in den Ingenieurwissenschaften ( $\beta = .305$ ,  $p < .001$ ) vorherzusagen. Die Ergebnisse nachgeschalteter Mediationsanalysen zeigen, dass das Vorwissen nicht ausschließlich direkt auf Fachwissen wirkt, sondern sowohl in Chemie, als auch in den Ingenieurwissenschaften partiell über visuelles Modellverständnis mediiert wird.

Weitere Ergebnisse, insbesondere die längsschnittlichen Datenanalyse, sollen auf der GEBF präsentiert werden, um die Rolle des visuellen Modellverständnisses auf Studienerfolg tiefergehend zu analysieren.

## **Der Einfluss von kognitiven und affektiv-motivationalen Faktoren auf den Studienerfolg von Chemiestudierenden**

**Daniel Averbeck, Jens Fleischer, Elke Sunfleth, Detlev Leutner, Matthias Brand**  
Universität Duisburg-Essen

Die Studieneingangsphase chemiebezogener Studiengänge wird bestimmt von geringen Studierendenzahlen bei gleichzeitig hohen Quoten von Studienabbrüchen. So liegt der prozentuale Anteil aller Hochschulzugangsberechtigten dieser Fächergruppe bei 1,4% (GDCh, 2015). Darüber hinaus wechseln 42% dieser Studierenden bereits nach zwei Semestern den Studiengang oder verlassen die Hochschule ohne einen Abschluss (Heublein et al., 2017). Entsprechend kommt der Studieneingangsdiagnostik in Hinblick auf die Voraussetzungen der Studierenden eine besondere Bedeutung zu, um frühzeitige und kriteriumsgeleitete Unterstützungsmaßnahmen entwickeln und implementieren zu können.

Diesbezüglich sollen in dem hier dargestellten Teilprojekt der DFG-geförderten Forschergruppe „ALSTER“ Prädiktoren für den Studienerfolg von Chemiestudierenden in der Studieneingangsphase untersucht werden.

Theoretischer Hintergrund

Untersuchungen zu fachspezifischen Studienerfolgsprediktoren in der Chemie sind bisher nur vereinzelt durchgeführt worden. Zu den wichtigsten Prädiktoren gelten weithin die Abiturnote (Tai, Ward & Sadler, 2006) sowie das chemische Vorwissen (Freyer, 2013; Hailikari & Nevgi, 2010). Darüber hinaus zeigen sich weitere Zusammenhänge zu verschiedenen kognitiven Faktoren wie beispielsweise der Intelligenz (Tai, Sadler & Loehr, 2005), den mathematischen Fähigkeiten (Busker et al., 2011) oder dem räumlichen Vorstellungsvermögen (Wu & Shah, 2004). Die Einflüsse von affektiv-motivationalen Variablen auf den Studienerfolg bleiben dagegen meist unberücksichtigt. Da den einzelnen Studien zudem häufig unterschiedliche Rahmenbedingungen zugrunde liegen (z. B. Betrachtung verschiedener Teilgebiete der Chemie oder unterschiedlicher Kohorten) oder unterschiedliche Operationalisierungen des Studienerfolgs genutzt werden, führen die resultierenden Ergebnisse zu unterschiedlichen Einflussfaktoren. Ein Transfer der Ergebnisse auf andere chemische Teilgebiete beziehungsweise deren Verallgemeinerung lässt sich aus diesen mit unterschiedlichen Zielen verfolgten Studien kaum ableiten.

Demzufolge werden in diesem Forschungsprojekt die Einflüsse kognitiver sowie affektiv-motivationaler Faktoren auf den Studienerfolg in den chemischen Teilbereichen der Allgemeinen, Physikalischen, Analytischen sowie der Anorganischen und Organischen Chemie innerhalb der Studieneingangsphase von Chemiestudierenden untersucht. Darüber hinaus werden die Zusammenhänge zwischen den einzelnen Teilgebieten in Hinblick auf die Performanz sowie den Wissenszuwachs aufgeklärt.

Methode

Die Datenerhebung erfolgte in einem Längsschnittstudien-Design mit drei Messzeitpunkten innerhalb der ersten beiden Semester. 275 Erstsemesterstudierende an zwei unterschiedlichen deutschen Universitäten bearbeiteten adaptierte allgemeine Fragebögen und neu entwickelte, teilgebietsspezifische Tests. Am Beginn des ersten Semesters (T1) wurden kognitive Fähigkeiten, affektiv-motivationale Faktoren, Personenvariablen sowie das Vorwissen (Testleistung zu Beginn des Semesters) in den Teilgebieten der Allgemeinen, Analytischen und Physikalischen Chemie erhoben. Am Ende des ersten Semesters (T2) wurden das in diesen Teilgebieten erworbene Fachwissen (Testleistung zum Ende des Semesters) sowie das Vorwissen der Studierenden in den weiterführenden Teilgebieten der Anorganischen und Organischen Chemie erfasst. Am Ende des zweiten Semesters (T3) wurde abschließend das korrespondierende Fachwissen erhoben.

Ergebnisse

Erste Ergebnisse der durchgeführten multivariaten-mehrgruppen Regressionsanalysen zeigen, dass zum einen das Vorwissen im jeweiligen Teilgebiet, aber auch das Vorwissen in der Allgemeinen Chemie die zentralen Prädiktoren für das erworbene Wissen in der Physikalischen bzw. Analytischen Chemie am Ende des ersten Semesters darstellen. Das Vorwissen in der

Analytischen Chemie ( $\beta = 0.395$  ;  $p < .001$ ) und das Vorwissen in der Allgemeinen Chemie ( $\beta = 0.321$  ;  $p < .001$ ) zeigen dabei einen ähnlich starken Einfluss auf das erworbene Wissen in der Analytischen Chemie. In der Physikalischen Chemie ist der Einfluss des teilgebietsspezifischen Vorwissens ( $\beta = 0.351$  ;  $p < .001$ ) größer als der des Vorwissens in der Allgemeinen Chemie ( $\beta = 0.222$  ;  $p < .001$ ). Darüber hinaus zeigt sich, dass die mathematischen Fähigkeiten der Studierenden, die Abiturnote und das Geschlecht ebenfalls einen moderaten Einfluss auf das erworbene Fachwissen in allen drei Teilbereichen am Ende des ersten Semesters ausüben. Alle weiteren Eingangsvoraussetzungen zeigen entweder keinen Einfluss auf den Wissenserwerb oder werden vollständig über das Vorwissen mediiert.

Mit Blick auf die weiterführenden chemischen Teilgebieten der Organischen und Anorganischen Chemie innerhalb des zweiten Semesters zeigen die Analysen, dass das erworbene Fachwissen in der Allgemeinen Chemie eine zentrale Einflussgröße auf das jeweilige Vorwissen darstellt.

Die Ergebnisse zum zweiten Semester werden im Vortrag vorgestellt.

## F20

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Kollegienhaus H120

### Entwicklung und Determinanten geschlechterspezifischer Motivation im MINT-Bereich

Chair(s): **Elisa Oppermann** (Freie Universität Berlin, Deutschland)

DiskutantIn(nen): **Rebecca Lazarides** (Universität Potsdam)

Im Zuge der technologischen und gesellschaftlichen Veränderungen kommt den Bereichen Mathematik, Informatik, Naturwissenschaften und Technik (MINT) eine zunehmende Bedeutung zu. Der Bedarf an MINT-Fachkräften steigt stetig und auch im persönlichen Leben wird eine grundlegende MINT-Kompetenz immer wichtiger. Gleichzeitig zeichnen sich bereits in der Grundschule, insbesondere unter Mädchen, ein geringes Interesse an MINT-Fächern sowie geringe Fähigkeitsüberzeugungen ab (Eccles, 1994; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; OECD, 2016; Wigfield et al., 1997). So berichten Mädchen typischerweise ein niedrigeres Selbstkonzept als Jungen sowie häufig auch ein geringeres Interesse in Mathematik (Eccles, 1994; Marsh et al., 2005; Wigfield et al., 1997) und den Naturwissenschaften (Eccles, 1994; Frome, Alfeld, Eccles, & Barber, 2006). Die Entwicklung dieser Geschlechterunterschiede in der MINT-Motivation, sowie Determinanten dieser, sind bisher jedoch kaum erforscht. Vor diesem Hintergrund zielt das Symposium darauf ab, zu einem besseren Verständnis der Entwicklung und Determinanten der geschlechterspezifischen Motivation im MINT-Bereich MINT beizutragen. In vier Einzelbeiträgen wird die Lernmotivation von Mädchen und Jungen in unterschiedlichen Nationen sowie in unterschiedlichen Abschnitten des Bildungsverlaufs untersucht und deren Entwicklung sowie mögliche Einflussfaktoren in einer gemeinsamen Diskussion betrachtet.

Der erste Beitrag von Gaspard et al. untersucht Verläufe in den geschlechterspezifischen Motivationsprofilen anhand eines US-amerikanischen Datensatzes von Schülerinnen und Schülern von der ersten bis zur zwölften Klassenstufe u.a. im Fach Mathematik.

Oppermann et al. untersuchen im zweiten Beitrag Geschlechterunterschiede in der naturwissenschaftlichen Lernmotivation von Kindern im Vorschulalter sowie Zusammenhänge mit den naturwissenschaftlichen Aktivitäten und den motivationalen Überzeugungen der betreuenden Fachkräfte.

Der dritte Beitrag von Ehrtmann et al. betrachtet den Einfluss der individuellen Geschlechtsrollenorientierung und der Klassenzusammensetzung auf die geschlechterspezifische Motivation von Schülerinnen und Schülern in der 6. - 8. Klasse in den Fächern Mathematik und Deutsch.

Der Beitrag von Keller et al. analysiert Geschlechtsunterschiede in der domänenspezifischen Motivation von 15-Jährigen Schülerinnen und Schülern in der Gruppe der Spitzenleister in Mathematik weltweit.

Die Beiträge werden abschließend vor dem Hintergrund ihrer Bedeutsamkeit für die pädagogische sowie motivationspsychologische Forschung von Prof. Lazarides diskutiert.

#### Beiträge des Symposiums

### Profile in den Verläufen der Erwartungen und Wertüberzeugungen von Schülerinnen und Schülern für Mathematik und die Muttersprache: Geschlechtsunterschiede und Zusammenhänge mit beruflichen Aspirationen

**Hanna Gaspard<sup>1</sup>, Fani Lauer<sup>2</sup>, Norman Rose<sup>1</sup>, Allan Wigfield<sup>3</sup>, Jacquelynne Eccles<sup>4</sup>**

<sup>1</sup>Universität Tübingen, <sup>2</sup>Universität Bonn, <sup>3</sup>University of Maryland, <sup>4</sup>University of California, Irvine

#### Theoretischer Hintergrund & Fragestellung

Die Erwartungen und Wertüberzeugungen von Schülerinnen und Schülern in verschiedenen Schulfächern nehmen über die Schulzeit hinweg ab (z.B. Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). Dieser Abfall in der Motivation wird durch eine Wechselwirkung von entwicklungsbedingten und schulischen Faktoren erklärt (Wigfield et al., 2015). Beispielsweise postuliert die Interessentheorie, dass sich die Interessen der Schülerinnen und Schüler zunehmend ausdifferenzieren (Schiefele, 2009). Intraindividuelle Unterschiede in den Interessen der Schülerinnen und Schüler sollten somit mit dem Alter zunehmen, wobei Interessen in manchen Bereichen stark ausgeprägt bleiben sollten, während andere abnehmen. Die Interessentheorie legt außerdem nahe, dass die Abnahme in verschiedenen Schulfächern, die sich im Mittel finden lässt, die Verläufe der einzelnen Schülerinnen und Schüler nicht ausreichend repräsentiert. Stattdessen sollten sich Subgruppen von Schülerinnen und Schülern mit unterschiedlichen Verläufen in verschiedenen Schulfächern identifizieren lassen. Solche Profile in motivationalen Verläufen könnten zudem mit Geschlechtsunterschieden zusammenhängen und Unterschiede in Studien- und Berufswahlen erklären.

In der vorliegenden Studie wurde daher untersucht, ob sich in den Verläufen der Erwartungen und Wertüberzeugungen in Mathematik und der Muttersprache über die Schulzeit hinweg differenzierte Profile identifizieren lassen. Zudem wurden Zusammenhänge dieser Profile mit dem Geschlecht und weiteren Eingangsmerkmalen der Schülerinnen und Schüler sowie ihren beruflichen Aspirationen am Ende der Schulzeit untersucht.

#### Methode

Zur Untersuchung dieser Fragestellung wurden Daten von 1069 Schülerinnen und Schüler der US-amerikanischen „Childhood and Beyond“-Studie reanalysiert. Diese Studie liefert mithilfe eines sequentiellen Kohortendesigns mit drei Kohorten Daten von der ersten bis zur zwölften Klassenstufe. Erwartungen und Wertüberzeugungen in Mathematik und Englisch wurden anhand von fünf Items zum Selbstkonzept und zwei Items zum intrinsischen Wert operationalisiert, wobei der Wortlaut in den beiden Domänen und über die Schuljahre gleich war. Profile in den Verläufen von Erwartungen und Wertüberzeugungen in den beiden Fächern wurden anhand von Growth-Mixture Modellen untersucht. Anschließend wurden Unterschiede der gefundenen Profile in verschiedenen Schülercharakteristika (Geschlecht, elterliche Bildung, durch die Lehrkraft eingeschätzte Mathematik- und Lesefähigkeiten in Klasse 1-4) und den beruflichen Aspirationen in Klasse 12 untersucht.

#### Ergebnisse & Diskussion

Für die Verläufe des Selbstkonzepts wurden zwei Profile gefunden, die als „Moderate Abnahme in Mathematik/Hohes Selbstkonzept in Englisch“ (63,4 %) und „Moderate Abnahme in Mathematik/Starke Abnahme in Englisch“ (36,6 %) beschrieben werden können. Im Vergleich zu dem Profil „Moderate Abnahme in Mathematik/Hohes Selbstkonzept in Englisch“ waren Jungen im Profil „Moderate Abnahme in Mathematik/Starke Abnahme in Englisch“ stärker vertreten. Zudem waren die elterliche Bildung und die durch die Lehrkraft eingeschätzten Lesefähigkeiten in diesem Profil vergleichsweise niedriger. Dieses

Profil ging zudem mit höheren beruflichen Aspirationen im Bereich Mathematik/Naturwissenschaften und niedrigeren Aspirationen im sozialen Bereich in Klasse 12 einher.

Für die Verläufe des intrinsischen Werts wurden drei Profile gefunden, die als „Starke Abnahme in Mathematik/Anfängliche Abnahme in Englisch“ (37,0 %), „Moderate Abnahme in Mathematik/Starke Abnahme in Englisch“ (29,7 %) und „Stabiler Verlauf in Mathematik und Englisch“ (33,4 %) beschrieben werden können. Jungen waren im Profil „Stabiler Verlauf in Mathematik und Englisch“ stärker vertreten als in den anderen beiden Profilen. Zudem zeigten Schülerinnen und Schüler des Profils „Starke Abnahme in Mathematik/Anfängliche Abnahme in Englisch“ niedrigere Mathematikfähigkeiten als Schülerinnen und Schüler im Profil „Moderate Abnahme in Mathematik/Starke Abnahme in Englisch“. Schülerinnen und Schüler des Profils „Starke Abnahme in Mathematik/Anfängliche Abnahme in Englisch“ berichteten im Vergleich zu den anderen beiden Profilen niedrigere berufliche Aspirationen im Bereich Mathematik/Naturwissenschaften und im Vergleich zum Profil „Stabiler Verlauf in Mathematik und Englisch“ höhere Aspirationen im sozialen Bereich.

Die Ergebnisse zeigen, dass Profile in den Verläufen von Erwartungen und Wertüberzeugungen in Mathematik und der Muttersprache über die Schuljahre zunehmend differenziert werden. Es fanden sich bedeutsame Zusammenhänge dieser Profile mit den Charakteristika der Schülerinnen und Schüler sowie mit ihren beruflichen Aspirationen, was die Validität dieser Profile unterstützt.

### **Welche Rolle spielen Fähigkeitsüberzeugungen pädagogischer Fachkräfte für die Lernmotivation von Mädchen und Jungen in den Naturwissenschaften?**

**Elisa Oppermann<sup>1</sup>, Martin Brunner<sup>2</sup>, Yvonne Anders<sup>1</sup>**

<sup>1</sup>Freie Universität Berlin, <sup>2</sup>Universität Potsdam

Theoretischer Hintergrund und Fragestellung

Lernen in den Naturwissenschaften beginnt lange vor der Einschulung. Kinder erkunden naturwissenschaftliche Phänomene in ihrer Umwelt und entwickeln dadurch erste motivationale Überzeugungen bezüglich der Naturwissenschaften als Inhaltsbereich sowie sich selbst als Lernende (Mantzicopoulos, Patrick, & Samarapungavan, 2008). Die Bedeutung früher Erfahrungen für die Entwicklung bereichsspezifischer Lernmotivation ist gut belegt (Wigfield et al., 1997). Bisher ist jedoch wenig darüber bekannt, welche Rolle pädagogische Fachkräfte hierbei spielen. Forschungsbefunde zeigen, dass Fachkräfte sich unsicher in den Naturwissenschaften und der Vermittlung dieser fühlen (Greenfield et al., 2009) und folglich die Naturwissenschaften meiden (Appleton & Kindt, 1999; Harlen & Holroyd, 1997). Die Fähigkeitsüberzeugungen der Fachkräfte äußern aber sich möglicherweise nicht nur in der Häufigkeit naturwissenschaftlicher Lernangebote, sondern könnten durch Prozesse wie das Modelllernen auch an die Kinder weitergeben werden (vgl. Bandura, 1977). Insbesondere für junge Kinder spielt Modelllernen eine wichtige Rolle für den Erwerb von Fähigkeiten und Überzeugungen (Schunk, 1987; Zimmerman & Ringle, 1981). Nach der sozial-kognitive Theorie sollten bei über 90% weiblicher Fachkräften vor allem Mädchen hierfür empfänglich sein, da Kinder gleichgeschlechtliche Modelle eher imitieren (Bandura, 1977). Die Zusammenhänge zwischen den naturwissenschaftlichen Fähigkeitsüberzeugungen der Fachkräfte, der Häufigkeit der angebotenen Lerngelegenheiten in den Naturwissenschaften und der naturwissenschaftlichen Motivation der Kinder sind bisher nicht erforscht. Die vorliegende Untersuchung soll diese Zusammenhänge sowie geschlechterspezifische Muster näher betrachten.

Methode

Die Studie basiert auf Daten von 277 Kindern (Alter 5-6 Jahre) und 348 Fachkräften aus 5 Bundesländern in Norddeutschland. Die naturwissenschaftliche Motivation der Kinder wurde mithilfe der Y-CSM Skala (Oppermann, Brunner, Eccles, & Anders, 2017) in 20-minütigen Einzelinterviews erhoben. Die Skala besteht aus den Subskalen naturwissenschaftliche Selbstwirksamkeitserwartung (15 Items,  $\alpha=0.87$ ) und Lernfreude (13 Items;  $\alpha=0.86$ ). Angaben zur Selbstwirksamkeitserwartung der Fachkräfte in der Vermittlung von Naturwissenschaften ( $\alpha=0.83$ ) sowie zur Häufigkeit naturwissenschaftlicher Aktivitäten (Einzelitem) wurden durch einen Fragebogen erhoben.

Die Zuordnung von Fachkraft- und Kind-Daten erfolgte anhand der Angaben der Fachkräfte zur Betreuungsintensität. Da pro Kind nur die Fachkraft ausgewählt wurde, die das Kind am häufigsten betreut, sowie einige Kinder von keiner der Fachkräfte in unserer Stichprobe betreut wurden, verringert sich der Datensatz auf  $N=234$  Kinder und  $N=88$  Fachkräfte. Die Berechnung der Zusammenhänge zwischen Fachkraft- und Kind-Variablen erfolgte mithilfe von Pfadmodellen in Mplus. Für die Analyse geschlechterspezifischer Zusammenhangsmuster wurden Mixture Modelle herangezogen. Die Mehrebenenstruktur wurde in allen Analysen berücksichtigt und fehlende Werte wurden mit „full information maximum likelihood“ geschätzt.

Ergebnisse und Diskussion

Die Ergebnisse des Pfadmodells zeigen, dass Fachkräfte mit höherer Selbstwirksamkeitserwartung in den Naturwissenschaften häufiger naturwissenschaftliche Aktivitäten anbieten ( $\beta=.35, p<.01$ ). Mit Blick auf die Motivation der Kinder zeigt sich ein signifikanter Effekt der Selbstwirksamkeitserwartung der Fachkräfte ( $\beta=.22, p<.01$ ), nicht aber der Aktivitäten. Ergebnisse der Mixture Modelle offenbaren allerdings geschlechterspezifische Zusammenhangsmuster: Der Zusammenhang zwischen der Selbstwirksamkeitserwartung der Fachkräfte und der kindlichen Motivation ist bei Mädchen stärker ausgeprägt (Selbstwirksamkeitserwartung:  $\Delta\beta=-.29, p=.03$ ; Lernfreude:  $\Delta\beta=-.24, p=.05$ ), wohingegen der Zusammenhang zwischen den Aktivitäten und der kindlichen Motivation bei Jungen stärker ausfällt (Selbstwirksamkeitserwartung:  $\Delta\beta=.31, p=.02$ ; Lernfreude:  $\Delta\beta=.31, p=.07$ ).

Die Ergebnisse auf Fachkräftebene bestätigen, dass die naturwissenschaftlichen Fähigkeitsüberzeugungen ausschlaggebend für die Aktivitäten der Fachkräfte in den Naturwissenschaften sind. Das Ergebnismuster der Zusammenhänge zwischen Fachkraft- und Kindvariablen legt zudem nahe, dass die Fähigkeitsüberzeugungen sowie die Aktivitäten der Fachkräfte in den Naturwissenschaften nicht für Jungen und Mädchen gleichermaßen relevant sind. Stattdessen zeigen sich geschlechterspezifische Effekte. Frühere Studien konnten bereits für den Primarbereich sowie die frühe sprachliche Bildung zeigen, dass der Einfluss der Fachkräfte (bzw. Lehrkräfte) auf Kind-Outcomes vom Geschlecht des Kindes abhängig ist (Beilock, Gunderson, Ramirez, & Levine, 2010; Wolter, Braun, & Hannover, 2015). Dies ist jedoch die erste Studie, die dieses Ergebnismuster für die frühe naturwissenschaftliche Bildung aufzeigt.

### **Der Einfluss individueller Geschlechtsrollenorientierung und der Klassenzusammensetzung nach Geschlecht auf die Motivation in Mathematik und Deutsch**

**Lisa Ehartmann<sup>1</sup>, Ilka Wolter<sup>1</sup>, Bettina Hannover<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für Bildungsverläufe (LIfBi), <sup>2</sup>Freie Universität Berlin



Die Motivation, die Schülerinnen und Schüler in einer fachlichen Domäne zeigen, trägt zur Erklärung interindividueller Unterschiede in Kompetenzen, zur Vorhersage von Leistungskurs- und berufsbezogenen Wahlen und der Kompetenzentwicklung bei (z.B. Eccles & Wigfield, 2002; Lazarides, Rohowski, Ohlemann & Ittel, 2015; Lazarides & Watt, 2014; Wentzel & Wigfield, 1998). Entsprechend hat sich die domänenspezifische Motivation auch für die Erklärung von Geschlechtsunterschieden in Kompetenzen und Selbstkonzepten als sehr bedeutsam erwiesen (Eccles, 2011). In dieser Studie wurden Variablen untersucht, von denen wir angenommen haben, dass sie ihrerseits die Motivation von Mädchen und Jungen vorhersagen: a) die Geschlechtskonnotation der Fachdomäne, b) die Klassenzusammensetzung nach Geschlecht, genauer die Abweichung von einer Gleichverteilung, mit der Geschlecht als soziale Kategorie für die jeweilige Minoritätsgruppe im Klassenzimmer salienter wird (McGuire & Padawer-Singer, 1976), und c) die individuelle Geschlechtsrollenorientierung, d.h. das Ausmaß, in dem die Person gesellschaftliche Normen über die traditionelle Arbeitsteilung zwischen Männern und Frauen und differentielle an Männer und Frauen gerichtete Erwartungen bejaht. Für die maskulin konnotierte Domäne Mathematik haben wir eine stärkere Motivation bei Jungen als bei Mädchen und für die weiblich konnotierte Domäne Deutsch eine stärkere Motivation bei Mädchen als bei Jungen erwartet. In beiden Fachdomänen sollten Mädchen und Jungen in Abhängigkeit ihrer Geschlechtsrollenorientierung differentiell auf die Klassenkomposition nach Geschlecht reagieren: Jugendliche mit einer traditionellen Orientierung sollten in der Domäne, die nach Geschlechtsstereotypen nicht zu ihrem Geschlecht gehört, wenig motiviert sein, und zwar unabhängig davon, wie die Klasse nach Geschlecht zusammengesetzt ist. Jugendliche mit einer egalitären Orientierung hingegen sollten in ihrer Motivation gegenüber der geschlechtsinkongruenten Domäne nur in dem Maße beeinträchtigt sein, wie Geschlecht aufgrund einer Unterrepräsentanz ihrer eigenen Geschlechtsgruppe im Klassenzimmer salient ist. Diese Interaktionshypothese haben wir vor dem Hintergrund der Befunde von Zander und Wolter (2016) formuliert, die fanden, dass der Einfluss der geschlechtsbezogenen Klassenkomposition auf die Motivation davon moderiert wurde, für wie bedeutsam die Jugendlichen ihre Geschlechtszugehörigkeit für ihre Identität erachteten.

Diese Annahmen haben wir in einer Stichprobe von N = 8317 (47,6 % weiblich) Schüler\_innen der sechsten Klassenstufe aus dem Nationalen Bildungspanel (Blossfeld, Roßbach & von Maurice, 2011; SC3) getestet. Durch vier Items wurde erfasst, inwiefern die Jugendlichen eher traditionell oder egalitär in Bezug auf Geschlechterrollen eingestellt sind. In der achten Klasse wurden sie zu ihrer Motivation in den Domänen Mathematik und Deutsch mit jeweils vier Items befragt. Außerdem wurde die geschlechtsbezogene Zusammensetzung der Klasse in der achten Klasse erfasst.

Es wurden Mehrebenenanalysen unter Spezifizierung der cross-level Interaktion aus der geschlechtsbezogenen Klassenkomposition und der Geschlechtsrollenorientierung und des Geschlechts der Schüler\_innen mit Mplus berechnet. Die Ergebnisse zeigten weder für Mathematik noch für Deutsch signifikante Effekte der Klassenzusammensetzung auf die Motivation der Mädchen oder Jungen. Auch die erwartete dreifach-Interaktion aus geschlechtsbezogener Klassenzusammensetzung, Geschlechtsrollenorientierung sowie Geschlecht der Schüler\_innen wurde weder in Mathematik noch in Deutsch signifikant.

Die Ergebnisse unserer Studie sind mit den Befunden von Dresel, Stöger und Ziegler (2006) vereinbar, die keinen direkten Effekt der geschlechtsbezogenen Klassenkomposition auf motivationale Variablen zeigen konnten. Vor dem Hintergrund der Befunde von Zander und Wolter (2016) hatten wir jedoch erwartet, dass Einflüsse der Klassenzusammensetzung nicht direkt, sondern vermittelt über individuelle Einstellungsmerkmale nachweisbar sein sollten. Anders als bei Zander und Wolter (2016) wurde in dieser Studie nicht die Zentralität gemessen, die Schülerinnen und Schüler ihrem Geschlecht beimaßen, sondern die Einstellungen der Jugendlichen gegenüber Geschlechtsrollen erfasst. Es scheint, dass Effekte der Klassenzusammensetzung unabhängig davon sind, ob Jungen und Mädchen egalitäre oder traditionelle Vorstellungen über Geschlecht haben, sondern nur davon bestimmt sind, für wie bedeutsam sie ihre Geschlechtszugehörigkeit für ihre Identität erachten. Die Diskussion gibt einen Ausblick auf zukünftige Studien, in denen der potentielle Einfluss geschlechtsbezogener Einstellungen und Identitätsaspekte auf die Motivation von Mädchen und Jungen gleichzeitig berücksichtigt werden sollte.

### **Wer gehört zu den Spitzenleistern in Mathematik? Eine Metaanalyse der Geschlechtsunterschiede in Leistung und Leistungsmotivation von Schülerinnen und Schülern weltweit**

**Lena Keller<sup>1</sup>, Francis Preckel<sup>2</sup>, Jacquelynne Eccles<sup>3</sup>, Martin Brunner<sup>1</sup>**  
<sup>1</sup>Universität Potsdam, <sup>2</sup>Universität Trier, <sup>3</sup>University of California, Irvine

Theoretischer Hintergrund: Spitzenleister in Mathematik sind aufgrund ihrer mathematischen Fähigkeiten dazu prädestiniert, erfolgreich Berufe im Bereich der Mathematik, Informatik, Naturwissenschaften und Technik (MINT) auszuüben (Park, Lubinski, & Benbow, 2007). Fachkräfte im MINT-Bereich tragen zur wissenschaftlichen und technologischen Entwicklung eines Landes bei und sind somit wichtig für dessen langfristiges wirtschaftliches Wohl und Wettbewerbsfähigkeit (Halpern et al., 2007). Obwohl die Nachfrage nach Fachkräften im MINT-Bereich seit Jahren hoch ist, sind Frauen in diesen Berufen immer noch unterrepräsentiert (National Science Board, 2016). Nach dem Erwartung-x-Wert-Modell für leistungsbezogene Entscheidungen (Eccles et al., 1983), werden berufliche Aspirationen durch individuelle Erfolgserwartungen und subjektive Werte geformt. Folglich sollten sich Individuen am wahrscheinlichsten für einen Beruf im MINT-Bereich entscheiden, wenn sie die höchsten Erfolgserwartungen für Berufe im MINT-Bereich haben und sie diesen Berufen die höchsten subjektiven Werte zuschreiben. Geschlechtsunterschiede in der Leistung und Leistungsmotivation in der Gruppe der Spitzenleister in Mathematik sollten demnach auch zu Geschlechtsunterschieden in deren späteren Berufsentscheidungen führen. Bisherige Metaanalysen zu Geschlechtsunterschieden in Leistung und Leistungsmotivation fokussierten entweder nur auf eine Domäne, z.B. auf Mathematik, und analysierten keine Profile über Domänen hinweg oder bezogen sich lediglich auf ein Land und nicht auf mehrere Länder oder nur auf die Gesamtpopulation und nicht auf Spitzenleister in Mathematik.

Fragestellung: In der vorliegenden Studie werden Geschlechtsunterschiede in der Gruppe der Spitzenleister in Mathematik in mehreren Leistungs- und Leistungsmotivationsdomänen weltweit untersucht. Dabei wird den Fragen nachgegangen, welche Geschlechterverteilung in der Gruppe der Spitzenleister in Mathematik vorliegt, ob bzw. in welchem Ausmaß es Geschlechtsunterschiede in Leistung und Leistungsmotivation in dieser Gruppe gibt sowie ob Geschlechtsunterschiede in Leistungsprofilen vorliegen.

Methode: Um diese Forschungsfragen zu beantworten, wurden repräsentative Daten von fünf Zyklen des Programme for International Student Assessment (PISA) verwendet (80 Länder, N = 175.744, 41% weiblich, ca. 15 Jahre alt) und länderspezifische, gewichtete Effektgrößen berechnet (Cohens d). Für die Analysen wurden Leistungstests in Mathematik, Lesen und Naturwissenschaften sowie 25 Selbstberichtskaalen zu den Einstellungen der Schüler/Innen gegenüber Mathe, Lesen und Naturwissenschaften verwendet. Leistungsprofile wurden durch die Subtraktion des individuellen Leistungswerts in einem Bereich von einem Leistungswert in einem anderen Bereich gebildet. Jugendliche in den Top 5% ihres jeweiligen Landes in Mathematik wurden als Spitzenleister in Mathematik definiert.

Ergebnisse: Lediglich 2 von 5 Jugendliche in der Gruppe der Spitzenleister in Mathematik waren Mädchen. Jungen zeigten etwas bessere Leistungen in Mathematik als Mädchen ( $d_{\text{Median}} = 0,15$ ), während Mädchen besser im Lesen abschnitten ( $d_{\text{Median}} = 0,61$ ). Geschlechtsunterschiede in der naturwissenschaftlichen Leistung waren vernachlässigbar klein ( $d_{\text{Median}} = 0,01$ ). Jungen zeigten ein Mathematik-orientiertes Leistungsprofil, während sich die Leistungsprofile der Mädchen ausgeglichener über die Domänen verteilten. Mädchen hatten außerdem eine positivere Einstellung gegenüber dem Lesen. Im Interesse an naturwissenschaftlichen Themen zeigte sich, dass über die Länder hinweg Jungen stärker an Physik interessiert waren, während Mädchen ein stärkeres Interesse an Botanik und Humanbiologie zeigten.

Diskussion: Mädchen in der Gruppe der Mathematik-Spitzenleister besitzen die kognitiven und motivationalen Voraussetzungen für Berufe sowohl im MINT-Bereich als auch außerhalb des MINT-Bereichs. Insgesamt ist ihr Leistungs- und Leistungsmotivationsprofil stärker verbal ausgeprägt. Diese Ausrichtung könnte dazu führen, dass sich Mädchen eher für Berufe außerhalb des MINT-Bereichs entscheiden. Folglich könnten die motivationalen und kognitiven Eigenschaften mathematisch leistungsstarker Mädchen, neben anderen Faktoren, zur Unterrepräsentation von Frauen im MINT-Bereich beitragen.

## F21

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Bernoullianum 30, H148

### Aktuelle Verfahren zur Analyse von Logdaten

Chair(s): **Steffen Brandt** (opencampus.sh und Kiron Open Higher Education, Deutschland), **Andreas Rausch** (Universität Mannheim)

DiskutantIn(nen): **Andreas Frey** (Friedrich-Schiller-Universität Jena)

Durch den vermehrten Einsatz computerbasierter Testverfahren stehen der Bildungsforschung heute in zunehmenden Maße sogenannte Logdaten zur Verfügung, welche die Interaktion des Probanden mit dem computerbasierten Testsystem aufzeichnen und es so ermöglichen, nicht nur den Lösungsvorschlag einer Probandin im Sinne eines Produkts zu beurteilen, sondern in einem gewissen Maße auch die Prozesse zu untersuchen, die zu einer bestimmten Lösung geführt haben. Aufgrund der enormen Menge verfügbarer Daten ergeben sich neue Verfahren zur Datenauswertung, aber auch Notwendigkeiten, bestehende Verfahren anzupassen, um weiterhin sinnvoll interpretierbare Auswertungsergebnisse zu erhalten. Das Symposium stellt verschiedene derzeit angewendete Verfahren zur Auswertung von Logdaten vor und zeigt damit die Möglichkeiten dieser Verfahren sowie die Voraussetzungen, die für die Anwendungen dieser Verfahren gegeben sein müssen. So sind je nach Verfahren gegebenenfalls unterschiedliche Logdaten-Formate erforderlich. Insbesondere für die Implementierung des Testsystems ist es daher wichtig, diese zu kennen, um einerseits relevante Interaktionen überhaupt beobachten zu können und sie andererseits auch in adäquater Form aufzuzeichnen.

Der erste Beitrag des Symposiums von Kroehne, Artelt, Goldhammer, Hahnel, Mahlow und Schoor befasst sich zunächst grundlegend mit der Systematik von Logdaten und stellt ein generelles Framework zur Aufbereitung von Logdaten bzw. zur formalen Definition von Indikatoren aus Logdaten (oder allgemeiner von Prozessdaten) vor. Das Framework wird anhand eines praktischen Beispiels, einem computerbasierten Test zur Erfassung von Multiple Document Literacy, illustriert und die Umsetzung eines auf dem Framework basierenden R-Paketes diskutiert.

Der zweite Beitrag von Domnick, Stadler und Greiff nutzt verschiedene aus den Logdaten extrahierte Indikatoren, um im Rahmen einer Regressionsanalyse zu untersuchen, ob diese eine verbesserte Differenzierung von Leistungsdaten ermöglichen. Angewendet wird das Verfahren auf einen komplexen Problemlösetest, der aufgrund der Kürze des Tests mit lediglich 5 Items, viele Probanden mit gleicher Punktzahl enthält.

Im dritten Beitrag von Bergrab, Rausch und Kögler werden Logdaten-Indikatoren, die den Aktivitätsverlauf der getesteten Personen beschreiben, analysiert. Verwendet wird dazu ein auf sogenannten n-grams basierendes Verfahren, auf dessen Basis Übergangswahrscheinlichkeiten für die Aktivitätsverläufe berechnet werden, um die Testteilnehmenden verschiedenen Gruppen mit einem ähnlichen Aktivitätsverlauf zuzuordnen. Angewendet wird das Verfahren auf Daten aus einem Test zum Problemlöseverhalten von Industriekaufleuten, insbesondere mit der Fragestellung inwieweit identifizierte Aktivitätsverläufe Einfluss auf den Problemlöseerfolg haben und wie sie sich mit den aus der Theorie bekannten Problemlösestrategien in Übereinstimmung bringen lassen.

Der vierte Beitrag von Abele beschäftigt sich ebenfalls mit der Untersuchung von Aktivitätsverläufen bzw. Verhaltensmustern. Zur Gruppierung der Testteilnehmenden hinsichtlich ihres Verhaltensmusters wird hier jedoch eine latente Klassenanalyse durchgeführt und es wird untersucht, inwieweit die über die latente Klassenanalyse identifizierten Verhaltensmuster der Gruppen mit theoretisch erwarteten Problemlösestrategien im Einklang stehen. Zusätzlich wird betrachtet, inwieweit bestimmte Verhaltensmuster (bzw. Problemlösestrategien) in Abhängigkeit von der Schwierigkeit der Aufgaben die Lösungswahrscheinlichkeit erhöhen bzw. verringern. Angewendet wird das Verfahren auf Daten aus einer Computersimulation zur Erfassung des diagnostischen Problemlösens von Kfz-Mechatronikern.

Eine abschließende Diskussion der Beiträge erfolgt durch Andreas Frey.

*Beiträge des Symposiums*

### Formale Definition von Indikatoren aus Log-Daten am Beispiel eines computerbasierten Tests zur Erfassung von Multiple Document Literacy

**Ulf Kröhne<sup>1</sup>, Cordula Artelt<sup>2</sup>, Frank Goldhammer<sup>3</sup>, Carolin Hahnel<sup>3</sup>, Nina Mahlow<sup>4</sup>, Cornelia Schoor<sup>5</sup>**

<sup>1</sup>DIPF – Deutsches Institut für Internationale Pädagogische Forschung, <sup>2</sup>LifBi – Leibniz-Institut für Bildungsverläufe e.V. und Universität Bamberg, <sup>3</sup>DIPF – Deutsches Institut für Internationale Pädagogische Forschung und ZIB – Zentrum für internationale Vergleichsstudien, <sup>4</sup>LifBi – Leibniz-Institut für Bildungsverläufe e.V., <sup>5</sup>Universität Bamberg

Theoretischer Hintergrund

Zu den Vorteilen computerbasierter Assessments wird die Verfügbarkeit zusätzlicher Informationen in Form von Log-Daten gezählt. Im Gegensatz zu Ergebnisdaten enthalten Log-Daten Informationen über den Ablauf einer Testsitzung – im besten Fall erlauben sie Einblicke in den Aufgabenbearbeitungsprozess einer Person und können somit diagnostisch wertvolle Informationen liefern. Aktuelle Arbeiten zu Log-Daten basieren vielfach auf heuristischen Vereinfachungen im Umgang mit diesen neuartigen Datenformen. So werden Indikatoren, die aus Log-Daten abgeleitet werden, häufig ad hoc als Teil eines eher explorativen Vorgehens extrahiert. Die theoriegeleitete Definition von Indikatoren und darauf aufbauend, die Integration der unterschiedlichen Informationsquellen in gemeinsame Modelle stellt hingegen eine aktuelle Herausforderung in der Forschung mit Log-Daten dar. Hierfür fehlt bisher ein theoretisches Framework als zentrale Gelingensbedingung, welches a) die vielfältigen, atomaren Informationen in ihrer assessmentplattformübergreifenden Heterogenität fasst und b) einen Bezug zu (Verhaltens- und kognitiven) Prozessen der Aufgaben- und Testbearbeitung herstellen kann. Von einem solchen Framework ist zu erwarten, dass die zur Analyse von Log-Daten notwendigen Schritte der Datenspeicherung, -aufbereitung und -analyse studienübergreifend reproduzierbar und kommunizierbar werden, so dass sich über Projekt-, Domänen- und Disziplinergrenzen hinweg ein aufeinander bezogener Forschungsstand aufbauen kann.

Fragestellung

Im Vordergrund dieses Beitrags steht die Vorstellung eines solchen theoretischen Frameworks zur formalen Definition von Indikatoren aus Log-Daten. Ausgehend von unterschiedlich strengen Definitionen „vollständiger“ Log-Daten wird eine Modellierung eingeführt, welches die atomaren Log-Ereignisse kontextualisiert und zur Unterscheidung benannter Abschnitte des Aufgabenbearbeitungsprozesses (sog. Zustände) integriert. Aufbauend auf diesen Zuständen, welche eine abstrakte und plattformübergreifende Beschreibung von Ausschnitten der Interaktion zwischen Assessmentplattform und Zielperson darstellen, können Indikatoren definiert werden.

Nach der Herleitung wird dieses Verfahren anhand von Aufgaben eines computerbasierten Tests zur Erfassung von Multiple Document Literacy (MDL) illustriert, die in dem Verbundprojekt MultiTex entwickelt wurden. Besonderes Augenmerk liegt dabei auf der Gegenüberstellung der verbalen Beschreibung von Indikatoren und deren Übertragung in eine formale Repräsentation.

#### Methode

Die schrittweise Rekonstruktion des Aufgabenbearbeitungsprozesses in Form von Wechseln, bzw. Sequenzen, von Zuständen stellt eine Methode dar, mit der empirische Log-Daten ausgewertet und die Werte der Indikatoren für Testpersonen ermittelt werden können. Dies kann bspw. konkret durch die Anreicherung von erhobenen Log-Daten um die rekonstruierten Zustände erfolgen, welche zum Zeitpunkt des jeweiligen Log-Ereignisses über den Interaktionsprozess zwischen Assessmentplattform und Zielperson angenommen wurde. Diese Rekonstruktion kann mit Hilfe sog. finiter Zustandsautomaten operationalisiert werden, deren Zustandsänderungen jeweils mit konkreten atomaren Log-Ereignissen identifiziert sind. Über das Framework hinaus wird im Beitrag zudem der Entwicklungsstand eines R-Pakets berichtet, mit welchem vorhandene Log-Daten fallweise um die rekonstruierte Sequenz von Zuständen angereichert werden können.

#### Ergebnisse

Der Beitrag wird exemplarisch über die Anwendbarkeit des theoretischen Frameworks für die Analyse der Aufgaben zur Erfassung von MDL berichten. Dazu werden Indikatoren, die im Prozess der Aufgabenkonstruktion von Inhaltsexperten theoriegeleitet beschrieben wurden, so umformuliert, dass sie aus Zusammenfassungsfunktionen (bspw. die kumulierte Gesamtzeit auf einem Zustand, die Häufigkeit des Eintretens bestimmter Zustände oder die Abfolge von Zuständen in Form von N-Grammen) der mit finiten Zustandsautomaten rekonstruierten Abfolge von Zuständen der Aufgabenbearbeitung ermittelt werden können. Als Illustration der formalen Definition dienen dabei graphische Repräsentationen der Zustandsautomaten. Es werden Erweiterungen der finiten Zustandsautomaten, insbesondere Wächterausdrücke, Variablen und orthogonale (unabhängige) Bereiche vorgestellt, welche für die Übertragung vorliegender Indikatoren im konkreten Test notwendig sind. Zentrales Ergebnis des Vortrags ist schließlich die Auflistung von Zusammenfassungsfunktionen, welche für die Ableitung von beobachteten Werten der Indikatoren aus den angereicherten Log-Daten notwendig sind, um das theoretische Framework konkret anwenden zu können.

Abschließend werden Möglichkeiten und Grenzen des Frameworks sowie dessen Umsetzung im R-Paket diskutiert. Es wird eine Systematik dargestellt, wie die Werte der Indikatoren aus Log-Daten in gemeinsame Analysen mit den Ergebnisdaten integriert und im Hinblick auf ihre diagnostische Eignung untersucht werden können.

### **Die besten der besten: Der Nutzen von Logdaten zur Differenzierung bei Deckeneffekten**

**Florian Krieger, Matthias Stadler, Samuel Greiff**  
University of Luxembourg

#### Theoretischer Hintergrund

Besonders im Rahmen bildungswissenschaftlicher Large-Scale Erhebungen werden zunehmend kürzere Messinstrumente eingesetzt, um mehr Konstrukte in der zur Verfügung stehenden Zeit messen zu können (z.B. Brähler, Zenger, & Kemper, 2015). Obwohl die Validität dieser Maße ausreichend gut gewährleistet werden kann, bleibt das Problem mangelnder Differenzierung der Leistungen von Schüler\*innen durch die reduzierte Anzahl an Aufgaben (z.B. Axelrod, 2002). Besonders Schüler\*innen mit maximaler Leistung stellen ein Problem dar, da hier nicht weiter differenziert werden kann und die geringe Anzahl an Aufgaben in kurzen Messinstrumenten keinen Platz für ausreichend schwere Aufgaben lässt.

#### Fragestellung

Ziel dieser Studie war es daher zu untersuchen ob Verhaltensunterschiede in den Aufgaben eine zusätzliche Möglichkeit bieten zwischen Schüler\*innen mit maximaler Leistung zu differenzieren. Dabei werden die Logdaten genutzt die bei computerbasierter Testung aufgezeichnet werden, um zusätzliche Informationen über die Probanden zu erhalten. Besonders komplexe Problemlöseaufgaben sind hier von Nutzen, da sie eine große Bandbreite an unterschiedlichen Verhaltensweisen ermöglichen.

#### Methode

Um die Fragestellung zu untersuchen bearbeitete eine bevölkerungsrepräsentative Stichprobe von N = 1476 finnischen Schüler\*innen eine Reihe von kontextfreien komplexen Problemen (nach dem MicroDYN Ansatz; Greiff, Fischer, Stadler, & Wüstenberg, 2015) die absichtlich kürzer gehalten war als normal üblich (5 statt 9 Aufgaben). Dadurch erreichte ein relativ großer Anteil der Schüler\*innen (n = 329; 22%) die maximale Punktzahl. Zusätzlich wurden die Häufigkeit der Interaktionen mit den Problemen, die Verwendete Zeit, und ob die Vary-One-Thing-At-A-Time (VOTAT; Tschirgi, 1980) Strategie angewandt wurde. Als Kriterium für die Nützlichkeit der Prozessdaten wurden die Durchschnittsnoten der Schüler\*innen erfragt. Diese wurde mittels latenter Regressionen durch die drei Verhaltensindikatoren vorhergesagt.

#### Ergebnisse

Trotz einer null-Varianz in der Gesamtleistung, variierten die Schüler\*innen mit maximaler Punktzahl signifikant bezüglich der Anzahl der Interaktionen, der verwendeten Zeit und der verwendeten Strategie. Erwartungsgemäß zeigten die Anzahl der Interaktionen negative ( $\beta = -.13^*$ ), die verwendete Zeit ( $\beta = .15^*$ ) und die Strategienutzung positive ( $\beta = .23^{**}$ ) Zusammenhänge mit der Durchschnittsnote der Schüler\*innen ( $R^2=.08$ ).

Die Ergebnisse deuten darauf hin, dass Verhaltensindikatoren aus Logdaten genutzt werden können um auch zwischen Schüler\*innen mit maximaler Punktzahl zu differenzieren. Dies bietet neben diagnostischen Möglichkeiten auch Ansatzpunkte für Interventionen. Weitere Studien sollten diese Ergebnisse auf Schüler\*innen mit minimalen Punktzahlen erweitern.

### **Analyse von n-grams aus Logdaten computerbasierten Problemlösens zur Identifikation erfolgskritischer Aktivitätsverläufe**

**Michael Berggrab<sup>1</sup>, Andreas Rausch<sup>2</sup>, Kristina Kögler<sup>3</sup>**

<sup>1</sup>Otto-Friedrich-Universität Bamberg, <sup>2</sup>Universität Mannheim, <sup>3</sup>Goethe-Universität Frankfurt

#### Theoretischer Hintergrund

Logdaten aus computerbasierten Assessments bieten die Möglichkeit, Aktivitätsverläufe der Testteilnehmenden zu rekonstruieren und diese theoriegestützt oder datengetrieben nach Mustern zu untersuchen, die in einem Zusammenhang mit der Lösungsqualität der Tests stehen bzw. die Lösungsqualität ggf. erklären können (He and von Davier, 2015; Rausch et al., in print). Datengetriebene Ansätze werden häufig den Bereichen „Educational data mining“ oder „Learning analytics“

zugeordnet (Biswas et al., 2010; Lajoie et al., 2015; Martin and Sherin, 2013). Dabei kommen in der Regel rechenintensive Verfahren wie Mustersequenzanalysen, künstliche neuronale Netzwerke oder Markov-Ketten zum Einsatz (He and von Davier, 2016; Kinnebrew et al., 2014; Lajoie et al., 2015). Im nächsten Schritt sind erfolgreiche Aktivitätsmuster inhaltlich zu interpretieren, um Strategien erfolgreicher Problemlöser zu identifizieren und diese in pädagogischen Interventionen auch Leistungsschwächeren näher zu bringen.

#### Fragestellung

Ziele der Studie sind (1.) die datengetriebene Identifikation erfolgskritischer Aktivitätsverläufe in einem Problemlösetest der kaufmännischen Domäne auf Basis von Logdaten. Es wird (2.) untersucht, inwiefern sich auf Basis unterschiedlicher Aktivitätsverläufe gebildete Gruppen von Testteilnehmenden in der auf Basis der Lösungsprodukte ermittelten Problemlösekompetenz unterscheiden. Schließlich wird (3.) untersucht, inwiefern sich die identifizierten Aktivitätsverläufe klassischen Problemlösestrategien zuordnen lassen.

#### Methode

Im Projekt DomPL-IK wurde die domänenspezifische Problemlösekompetenz in dem kaufmännischen Kernbereich des Controllings anhand dreier je 30-minütiger Problemszenarien gemessen, die in einer computerbasierten Büroumgebung zu bearbeiten waren. An der Studie haben sich 780 kaufmännische Auszubildende (50.1 % weiblich; Alter: M = 21.3; SD = 2.69) freiwillig beteiligt und erklärten schriftlich ihr Einverständnis. Die Szenarien erforderten jeweils das Auffinden und Beurteilen domänenspezifischer Informationen, die Anwendung domänenspezifischer Algorithmen (z. B. Errechnen von Plankosten), das Abwägen von Alternativen sowie das Treffen und fachlich adäquate Kommunizieren einer Entscheidung. Dazu stand den Probanden eine Vielzahl relevanter und irrelevanter Dokumente zur Verfügung. Zur Bearbeitung bot die computerbasierte Bürosimulation gängige Werkzeuge wie eine Tabellenkalkulation, Taschenrechner, Notizblock etc. Jeder Mausclick und jeder Tastaturanschlag wurden mit einem Zeitstempel aufgezeichnet (Rausch et al., 2016).

Zur Bearbeitung der Forschungsfragen werden die resultierenden Logdaten mittels text-mining-Methoden aufbereitet. In einem ersten Schritt erfolgt mittels einer modellfreien Sequenzanalyse eine erste Klassifizierung in stärkere und schwächere Testteilnehmende: Die Logdaten je Testteilnehmenden werden als einzelne Sequenz gewertet und anschließend in n-grams zerteilt. Hierbei werden die Sequenzen entweder als uni-gram in einzelne Aktivitäten (z. B. Email schreiben oder Dokument lesen) oder mehrere längere zusammenhängende Aktivitätsverläufe (erst Dokument lesen, dann Email schreiben, dann Email abschicken) zerlegt. Mittels der n-grams kann eine erste Gruppierung der Testteilnehmenden vorgenommen werden. Hierfür werden in erster Linie die Häufigkeiten der Aktionsmuster herangezogen (sog. feature selection methods; Oakes et al., 2001). Dies erlaubt eine erste deskriptive Analyse. In einem weiteren Schritt werden die Logdaten mittels sog. Hidden Markov Modelle (HMM) analysiert, um auch die tieferliegende latente Struktur der Daten zu untersuchen, Übergangswahrscheinlichkeiten für die Aktionsmuster zu ermitteln und die Testteilnehmenden auch unter der Berücksichtigung von Kovariaten zu gruppieren (Helske and Helske, 2017). HMMs haben demnach den Vorteil, dass in einem probabilistischen Rahmen die inneren Zustände bzw. die hinter den Aktivitätsverläufen liegenden Problemlösestrategien analysiert werden können (Rabiner, 1989).

#### Ergebnisse

Erste Analysen legen nahe, dass sich erfolgreiche und weniger erfolgreiche Problemlöser hinsichtlich der Problemlösestrategien unterscheiden (Rausch et al., in Druck). Die in n-grams aufgeteilten Sequenzen zeigen darüber hinaus, welche Aktivitätsfolgen typisch für verschiedene Gruppen sind. So zeigen beispielsweise erste Ergebnisse der Sequenzanalyse auf Basis von Chi-Quadrat-Tests zur Klassifizierung (Oakes et al., 2001), dass diejenigen, die sich während des Assessments durch die in den Szenarien eingebetteten Distraktoren ablenken ließen, auch weitere irrelevante Dokumente benutzten. Weitere Analysen werden im Rahmen des Beitrags präsentiert und diskutiert. Werden computerbasierte Testumgebungen auch zu Lernzwecken genutzt, können diese Analysen online durchgeführt und für automatisierte Interventionen genutzt werden (z. B. durch intelligente tutorielle Systeme; Weber, 2012).

### **Diagnosestrategien in beruflichen Kontexten: Erfassung und Analyse der Anwendungs- und Erfolgswahrscheinlichkeit mithilfe computergenerierter Logdaten und latenter Klassenanalysen**

**Stephan Abele<sup>1</sup>, Matthias von Davier<sup>2</sup>**

<sup>1</sup>Technische Universität Dresden, <sup>2</sup>National Board of Medical Examiners, Philadelphia, USA

#### Theoretischer Hintergrund

Im Vortrag wird eine Studie zu Diagnosestrategien, d.h. zum professionellen Handeln bei Diagnoseproblemen vorgestellt. Diagnoseprobleme repräsentieren Situationen, in denen die Ursache(n) für einen unerwünschten Zustand zu finden sind. Diagnoseprobleme sind in unterschiedlichen Berufen relevant: Lehrkräfte müssen die Gründe von Lernschwierigkeiten bestimmen, Mediziner die Ursachen von Krankheiten, Kfz-Mechatroniker die Ursachen von Kfz-Defekten etc. Um Diagnoseprobleme zu lösen, werden Diagnosestrategien benötigt. Der Vortrag fokussiert Diagnosestrategien von Kfz-Auszubildenden, integriert allerdings Erkenntnisse unterschiedlicher technischer und medizinischer Berufe (z.B. Konradt 1995; Coderre et al. 2003). Diagnosestrategien sind Ansätze, den diagnostischen Problemlöseprozess zu strukturieren. Diagnosestrategien dienen dazu, problemrelevante Informationen zu repräsentieren, daran anknüpfend Ursachenhypothesen zu formulieren, diese Hypothesen zu prüfen und deren Gültigkeit zu evaluieren (Abele 2017). Um diese Prozessanforderungen zu bewältigen, sind aufeinander abgestimmte mentale Problemlöseaktivitäten und beobachtbares Problemlöseverhalten (z.B. elektrotechnische Messung) erforderlich. Bei der computerbasierten Strategie werden Informationen aus einem computerbasierten Expertensystem (externe Informationsquelle) abgerufen und umgesetzt, um Diagnoseprobleme zu lösen. Diese Strategie ist nur anwendbar, wenn das Expertensystem relevante Informationen enthält. Bei der fallbasierten Strategie werden zur Problemlösung frühere Diagnosefälle auf ein aktuelles Diagnoseproblem übertragen. Die fallbasierte Strategie liegt v.a. bei Problemen vor, die in der Diagnosepraxis häufig vorkommen. Bei der modellbasierten Strategie werden mentale Modelle gebildet, die Ausschnitte relevanter Kfz-Systeme repräsentieren und eine systematische Diagnose erlauben. Diese Strategie ist anspruchsvoll, aber prinzipiell bei jedem Diagnoseproblem anwendbar.

#### Fragestellung

In der Studie wurde untersucht, von was die Anwendungs- und die Erfolgswahrscheinlichkeit der Strategien abhängen. Es wurde vermutet, dass die Diagnosestrategien bei der Bearbeitung von leichten und schwierigen Diagnoseproblemen vorkommen (H1). Zudem wurde unterstellt, dass die modellbasierte Strategie generell seltener vorkommt als die anderen Strategien (H2). Der Grund hierfür wird im mentalen Aufwand der Strategien gesehen, der bei der modellbasierten Strategie höher ist als bei den anderen Strategien. Vermutet wurde allerdings, dass die modellbasierte Strategie aufgrund ihrer vielseitigen Anwendbarkeit bei schwierigen Diagnoseproblemen trotz dieses Mehraufwandes häufiger auftritt als bei leichten

(H3). Schließlich wurde angenommen, dass die Erfolgswahrscheinlichkeit der Strategien bei leichten Diagnoseproblemen ähnlich hoch ist (H4a), bei schwierigen Problemen sollte die modellbasierte Strategie hingegen erfolgreicher sein als die anderen Strategien (H4b).

#### Methode

Zur Hypothesenprüfung wurden Kfz-Auszubildende des dritten Ausbildungsjahrs einbezogen (N=370). Den Auszubildenden wurden unterschiedlich schwierige Diagnoseprobleme vorgegeben. Basierend auf theoretischen Analysen wurde unterstellt, dass sich die Strategien in idiosynkratischen beobachtbaren Problemlöseverhaltensmustern manifestieren. Diese Verhaltensmuster wurden a priori theoretisch bestimmt. Das empirische Problemlöseverhalten der Probanden wurde mit einer authentischen Kfz-Computersimulation erfasst (z.B. Abele 2017). Anschließend wurde anhand computergenerierter Logdaten für jeden Probanden bestimmt, ob die theoretisch bestimmten Verhaltensweisen bei einem bestimmten Diagnoseproblem vorlagen. Die Identifikation der empirischen Verhaltensmuster erfolgte mit latenten Klassenanalysen. Um die Erfolgswahrscheinlichkeit der Strategien zu ermitteln, wurde anhand schriftlicher Probandenprotokolle zudem der individuelle Problemlöseerfolg ermittelt (Interrater-Reliabilität:  $\kappa > 0,92$ ).

#### Ergebnisse

Die latenten Klassenanalysen erbrachten Verhaltensmuster, die (weitgehend) im Einklang mit den theoretisch erwarteten Verhaltensmustern standen und sich im Sinne der drei theoretisch identifizierten Strategien interpretieren ließen. Im Einklang mit H1 traten bei der Lösung eines Diagnoseproblems unterschiedliche Strategien auf. Beobachtet wurde auch, dass die modellbasierte Strategie beim leichten Diagnoseproblem seltener vorgekommen war als die anderen Strategien. Beim schwierigen Diagnoseproblem trat die fallbasierte Strategie allerdings überraschend seltener auf als die modellbasierte. H2 konnte also nicht bestätigt werden. Dagegen korrespondierten die Ergebnisse mit H3: Die Anwendungswahrscheinlichkeit der modellbasierten Strategie war beim schwierigen Diagnoseproblem höher als beim leichten. Die Erfolgswahrscheinlichkeit der Strategien war beim leichten Diagnoseproblem hoch, entgegen H4a war die computerbasierte Strategie aber weniger erfolgreich als die anderen Strategien. Beim schwierigen Diagnoseproblem war die modellbasierte Strategie erwartungsgemäß erfolgreich, die beiden anderen Strategien dagegen nicht/kaum. Die Ergebnisse legen u.a. nahe, in Bildungskontexten unterschiedliche Problemlösestrategien und auch die Bedingungen zu lehren, unter denen die Strategien anwendbar und erfolgreich sind.

## F22

Zeit: Samstag, 17.02.2018: 10:30 - 12:15 · Ort: Musikwissenschaftliches Seminar, Hörsaal 006

### Personalisierte Lernkonzepte implementieren und weiterentwickeln: Tiefenstrukturelle Dimensionen in der Schul- und Unterrichtsentwicklung

Marco Galle<sup>1</sup>, Kurt Reusser<sup>1</sup>, Rita Stebler<sup>1</sup>, Christine Pauli<sup>2</sup>

<sup>1</sup>Universität Zürich, Schweiz; <sup>2</sup>Universität Freiburg, Schweiz; [marco.galle@ife.uzh.ch](mailto:marco.galle@ife.uzh.ch)

Im längsschnittlichen Forschungsprojekt perLen (personalisierte Lernkonzepte in heterogenen Lerngruppen) werden Schulen untersucht, die personalisierte Lernkonzepte implementieren und diese weiterentwickeln (perLen-Schulen). Personalisierte Lernkonzepte haben zum Ziel, jedem Lernenden einen persönlichen und passgenauen Zugang zum Lernen zu ermöglichen (Bray & McClaskey, 2015). Im Beitrag werden Ergebnisse einer strukturierenden Inhaltsanalyse (Kuckartz, 2016) von längsschnittlichen Leitfadenterviews unter folgender Forschungsfrage präsentiert: Wie häufig beschreiben Schulleitende und Lehrpersonen tiefenstrukturelle (der verbesserten Lehr-/Lernqualität betreffende) Veränderungen vergleichend zu Veränderungen der (organisationalen und lernmethodischen) Oberflächenstrukturen der Schul- und Unterrichtsqualität? Um welche Veränderungen handelt es sich und wie lassen sich diese charakterisieren?

#### Theorie

Um personalisierte Lernkonzepte nachhaltig in den Unterrichtsalltag zu implementieren und weiterzuentwickeln, bedarf es einer komplexen und vielschichtigen Schulentwicklung. Im Kern umfasst sie gegenseitig beeinflussende Entwicklungsdimensionen:

- Unterrichtsentwicklung
- Entwicklung der Zusammenarbeit von Lehrpersonen
- Entwicklung organisationaler Strukturen
- Entwicklung des Schullebens
- Personalentwicklung (Prain et al., 2015).

Forschungsfokus ist eine gelingende und lernwirksame Unterrichtsentwicklung. Hier sind – aus einer didaktisch-psychologischen Perspektive – nicht nur Dimensionen der Oberflächenstrukturen wichtig, sondern insbesondere tiefenstrukturelle Dimensionen (Reusser, 2008). Während unter Oberflächenstrukturen die Bereitstellung der Lerninfrastruktur und das schulorganisatorische sowie unterrichtsmethodische Handeln der Lehrpersonen zu verstehen ist, beziehen sich die Tiefendimensionen auf „psychologische Prozesse und Merkmale des Lehrens und Lernens, welche dem Unterricht als psychologisch-didaktische Qualitäts[merkmale] zugrunde liegen“ (Reusser & Pauli, 2010, S. 19). In personalisierten Lernkonzepten lassen sich Oberflächenstrukturen z.B. an veränderten Schul- und Unterrichtsstrukturen erkennen: alters- und/oder leistungsdurchmischte Lerngruppen, kleinere Input-Räume und Grossraumbüros für über 30 Schüler/innen (Lernlandschaften) sowie Wahlmöglichkeiten von Lernaufgaben und Lernmethoden. Merkmale der Tiefenstrukturen im personalisierten Unterricht sind z.B.:

- Individuell-adaptive, pädagogisch und fachdidaktisch gehaltvolle Unterrichts- und Lernunterstützungsangebote für Lernende und Lerngruppen
- Ganzheitliches Fördern und Bilden der Persönlichkeit unter mehrdimensionalen (fachlichen und überfachlichen) Kompetenzkriterien
- Ermöglichen von selbstgesteuerten und eigenverantwortlichen Lernens
- Bildend wirkende Lehrperson und Lerngemeinschaft (Stebler, Pauli & Reusser, 2017).

#### Methode

Von den perLen-Schulen, die am Forschungsprojekt teilnahmen, wurden elf perLen-Fallschulen ausgewählt, bei welchen längsschnittliche und leitfadengestützte Einzel- und Gruppeninterviews durchgeführt wurden (Schuljahr 2012/13 (t1), 2013/14 (t2), 2014/15 (t3)). Hierbei führte das Forschungsteam insgesamt 41 Interviews mit Schulleitenden (t1) und Lehrpersonen (t1, t2, t3).

Die Datenanalyse erfolgte mit einer computergestützten, strukturierenden Inhaltsanalyse (Kuckartz, 2016): In einem ersten Schritt wurde das Datenmaterial einer Basiscodierung unterzogen. Das hierfür verwendete Kategoriensystem wurde deduktiv anhand eines theoretischen Arbeitsmodells unterrichtszentrierter Schulentwicklung (u.a. Reusser, 2011; Rolff, 2013) konstruiert, welches die folgenden fünf Entwicklungsdimensionen umfasst: Unterrichtsentwicklung, Entwicklung der Zusammenarbeit der Lehrpersonen und der organisationalen Strukturen sowie Entwicklung des Schullebens und Personalentwicklung. Während der Datenanalyse wurden die Kategorien induktiv erweitert. Eine Qualitätsüberprüfung mittels Inter-Coder-Übereinstimmung ergab eine zufriedenstellende 80%ige Übereinstimmung.

Im zweiten Schritt wurde damit begonnen, die codierten Textstellen (Fundstellen) der Unterrichtsentwicklung anhand der zwei Kategorien Oberflächenstruktur und Tiefenstruktur feinzucodieren. Auch hier werden die Kategorien deduktiv erstellt und während des Codierens induktiv ergänzt. Bis zur Präsentation werden die Ergebnisse aller perLen-Fallschulen mit Inter-Coder-Übereinstimmung vorliegen.

#### Ergebnisse

Erste Auswertungen zweier perLen-Fallschulen zeigen, dass tiefenstrukturelle Dimensionen im Zusammenhang mit einer Schulentwicklung hin zu personalisierten Lernkonzepten genannt werden. Die Interviewten berichten viel weniger über tiefenstrukturelle Veränderungen als über organisationale und lernmethodische Massnahmen.

Eine besonders häufig erwähnte, tiefenstrukturelle Dimension ist das Fördern und Begleiten des selbstständigen und eigenverantwortlichen Lernens der Schüler/innen. Hier beschreiben die Lehrpersonen ein verändertes Lehrverhalten:

„Eben deshalb heisst es Lernbegleiter, obwohl wir Lehrpersonen sind. [...] Es ist eben nicht ein ‚wir zeigen vor und sie machen nach‘, sondern es ist ein nebenher gehen oder ein mitgehen. [...] Man lässt den Schüler/innen die Möglichkeit, eigenständige Lernwege zu entdecken [und] die Eigenverantwortung für ihr Lernen zu übernehmen. Das ist teilweise schwierig anzuschauen. Doch [führen] wir immer wieder Gespräche [mit den Schüler/innen] und bleiben [an ihren Lernwegen] dran“ (Fallschule\_K).

### **Matthias Forell, Gabriele Bellenberg, Grit im Brahm**

Ruhr-Universität Bochum, Deutschland; [matthias.forell@rub.de](mailto:matthias.forell@rub.de), [grit.imbrahm@rub.de](mailto:grit.imbrahm@rub.de)

Die Dauer der Schulbesuchszeit bis zum Abitur spielt in den gegenwärtigen Bildungsdebatten zum Gymnasium eine prominente Rolle. Die entsprechende KMK-Regelung sieht eine Mindestlernzeit von 265 Wochenstunden Unterricht ab Klasse 5 vor. Dadurch kann die Hochschulzugangsberechtigung über das Abitur entweder nach acht oder nach neun Jahren erworben werden. Nachdem zwischen dem Schuljahr 2001/02 und 2008/09 fast alle westdeutschen Bundesländer – mit Ausnahme von Rheinland-Pfalz – die Schulzeit am Gymnasium generell auf acht Jahre (G8) verkürzt haben, findet seit dem Schuljahr 2011/12 wieder ein Rollback zu neun Jahren bis zum Abitur (G9) statt, der schließlich dazu geführt hat, dass mittlerweile in nahezu allen dieser Bundesländern – das Saarland ausgenommen – G9 entweder flächendeckend wiedereingeführt wurde und G8 ersetzt hat, als Wahloption in unterschiedlichen Varianten oder im Rahmen eines Modellversuchs ermöglicht wurde.

Nordrhein-Westfalen legte sich bei der Verkürzung der gymnasialen Schulzeit zum Schuljahr 2005/06 darauf fest, die Oberstufe mit drei Schuljahren unangetastet zu lassen und die Sekundarstufe I von sechs auf fünf Schuljahre zu reduzieren. Diese Maßnahme stellte die Lehrkräfte vor allem in deren curricularer Unterrichtsplanung und -gestaltung vor große Herausforderungen. Die Rückkehr zum neunjährigen gymnasialen Bildungsgang im Rahmen eines Schulversuchs ab 2011/12 (G9neu) und dem damit verbundenen zusätzlichen Schuljahr in der Sekundarstufe I ermöglichte den Lehrkräften daraufhin neue curriculare Gestaltungsmöglichkeiten nicht nur für das hinzugewonnene Schuljahr, sondern auch für den gesamten Verlauf der Mittelstufe von Jahrgang 7 bis 10.

Der Beitrag geht der Frage nach, wie Lehrerinnen und Lehrer der Mittelstufe in den Kernfächern Deutsch, Mathematik und Englisch dieses Mehr an Lernzeit in ihrem Unterricht nutzen. Dabei werden zum einen fachbezogene unterrichtliche Prozesse und Muster transparent, zum anderen werden Unterschiede im professionellen Handeln der Lehrkräfte zwischen den Bildungsgängen herausgearbeitet und in Zusammenhang mit der jeweiligen Schülerklientel gesetzt. Zur Einordnung der Bedeutung von unterrichtlicher Lernzeit im Wirkungsgefüge von Lernvoraussetzungen, Lernprozess und Lernerfolg wird auf Carrolls Modell schulischen Lernens zurückgegriffen, in dem die zur Verfügung stehende Lernzeit und die Gestaltung dieser durch die Lehrkraft in Relation zu der Zeit gesetzt wird, die von den Schülerinnen und Schülern zum Lernen benötigt wird.

Als Sample dienen hierzu Unterrichtsreihenbögen von Gymnasiallehrkräften (n = >300 mit >1.000) dokumentierten Unterrichtsreihen), die im Rahmen der wissenschaftlichen Begleitung des o.g. Schulversuchs mittels eines querversetzten Kontrollgruppendesigns im Längsschnitt von Jahrgangsstufe 8 bis 10 in beiden Bildungsgängen (G8 und G9neu) erhoben wurden. Diese geben sowohl Auskunft über die gelehrt Inhalte und Kompetenzen der Lehrkräfte als auch über die unterrichtete Stundenanzahl pro Reihe, was die Möglichkeit eröffnet, sie auch als Rückkopplungsinstrument zur Entwicklung des professionellen Handelns der einzelnen Lehrkraft zu nutzen.

Die Dokumentenanalysen der Unterrichtsreihen zeigen deutliche Unterschiede in der Verteilung und im Umfang der Unterrichtszeit sowie in der inhaltlichen Gestaltung des Unterrichts zwischen den Kernfächern innerhalb eines Bildungsgangs wie auch zwischen den beiden Bildungsgängen. So ergibt der Abgleich mit den Stundentafeln der jeweiligen Gymnasien und deren Kerncurricula mit den eingereichten Unterrichtsreihenbögen, dass in beiden Bildungsgängen die reale von der nominalen Unterrichtszeit fachabhängig unterschiedlich stark differiert und zudem Fachinhalte zwischen den Jahrgangsstufen verschoben werden. Darüber hinaus offenbart der Vergleich der fachbezogenen Lernzeitnutzung zwischen den beiden Bildungsgängen eine zeitliche Anpassung der Lerninhalte an die Schülerschaft in G9neu im Verlauf der Sekundarstufe I. Demnach werden in G9neu Unterrichtsreihen mit gleichem Inhalt zu einem späteren Zeitpunkt und auch länger unterrichtet als dies im G8-Bildungsgang der Fall ist. Außerdem profitieren die drei Kernfächer unterschiedlich stark von dem zeitlichen Zugewinn.

### **Leistungskonsolidierung, Leistungssteigerung – oder etwas ganz anderes? Die Effekte einer Klassenwiederholung auf die Leistungsentwicklung von Schülerinnen und Schülern.**

**Paul Fabian<sup>1</sup>, Katja Scharenberg<sup>2</sup>, Wilfried Bos<sup>1</sup>**

<sup>1</sup>TU Dortmund, IFS, Deutschland; <sup>2</sup>Pädagogische Hochschule Freiburg, Institut für Soziologie, Deutschland; [paul.fabian@tu-dortmund.de](mailto:paul.fabian@tu-dortmund.de)

Politische und gesellschaftliche Debatten um Klassenwiederholungen erscheinen stark ideologisch aufgeladen. Politik und Interessenverbände (z.B. Deutscher Philologenverband, Verband Bildung und Erziehung e.V.) vertreten hierzu sehr gegensätzliche Positionen (Fabian 2012). Klassenwiederholungen sind eine national wie international höchst umstrittene Homogenisierungsmaßnahme (Tillmann 2008). In Deutschland haben einige Bundesländer (z.B. Hamburg, Berlin) das Wiederholen abgeschafft, in anderen (z.B. Bayern, Bremen) liegt die Wiederholerquote über dem Bundesdurchschnitt. Im internationalen Vergleich ist die Quote der SchülerInnen, die im Verlauf ihrer Bildungskarriere mindestens eine Klassenstufe wiederholt haben, in Deutschland mit 18,1% überdurchschnittlich hoch (OECD-Durchschnitt: 6,2%; Reiss et al. 2016). Die Forschungslage zur Effektivität von Klassenwiederholungen hingegen ist diffus. Im englischsprachigen Raum wurden Effekte von Klassenwiederholungen auf die Leistungsentwicklung umfangreich erforscht (z.B. Xia & Kirby 2009). In Deutschland finden sich nur wenige belastbare Studien, die konkrete Aussagen zur Effektivität von Klassenwiederholungen zulassen (z.B. Ehmke et al. 2008, 2010, 2017): Im Vergleich zu den gezeigten Leistungen im Vorjahr tritt im wiederholten Jahr zwar eine Leistungssteigerung (Bless et al. 2004) ein, die jedoch zum Ende des wiederholten Jahres abflacht und sich ins Gegenteil verkehrt (verfehlte Leistungskonsolidierung, Ehmke et al. 2008, 2010, 2017).

Als Grundlage zur Analyseplanung zur Wirkung von Klassenwiederholungen haben Palowski et al. (2016) das heuristische Wirkungsmodell der Klassenwiederholung vorgeschlagen. Es unterscheidet quantitative und qualitative Zugänge zur Bestimmung der Wirkung von Klassenwiederholungen. Je nach gewähltem empirischen Zugang werden unterschiedliche Adressatengruppen (z.B. Individuum, Eltern, Mitarbeitende aus bspw. der Schulsozialarbeit und -psychologie, Schuladministration) angesprochen, zugleich lässt nur der qualitative Ansatz die Erfassung individueller Bewältigung und Folgen einer Klassenwiederholung zu. So wichtig die dabei gewonnenen Erkenntnisse für die Adressatengruppen auch sein mögen - zur Steuerung des Bildungswesens können sie nicht oder nur selten herangezogen werden. Hier ermöglicht ein quantitativer Ansatz, sowohl die Häufigkeit von Klassenwiederholungen zu bestimmen als auch zu erfassen, ob und inwiefern dadurch den Funktionen der Schule (Qualifikations- und Allokationsfunktion; Fend 2006) Rechnung getragen wird.

Unser Beitrag widmet sich der Frage, wie sich KlassenwiederholerInnen ohne eine Klassenwiederholung entwickelt hätten. Auf Basis der hierzu bislang vorliegenden Forschungsbefunde wird erwartet, dass die Klassenwiederholung keinen Effekt auf die Leistungsentwicklung ausübt.

Datengrundlage bildet die National Educational Panel Study (NEPS, Startkohorte 3; Blossfeld & von Maurice 2011). Da die meisten Klassenwiederholungen in den Jahrgangsstufen 7 bis 9 stattfinden (vgl. Reiss et al. 2016), werden SchülerInnen



betrachtet, die die 7. Jahrgangsstufe wiederholten (n=128). Unter Nutzung des Rubin Causal Models (Rubin 1974) sowie des Propensity Score Matchings (Rosenbaum & Rubin 1985; Stuart 2010) und auf Basis von Informationen von SchülerInnen, Eltern sowie Klassenlehrkräften können KlassenwiederholerInnen mit jenen SchülerInnen verglichen werden, die ihnen in zentralen Hintergrundmerkmalen ähneln, jedoch regulär versetzt wurden. Mit der resultierenden Analysestichprobe (n=336) wurden lineare Regressionsanalysen nach dem Makro-Modell der Bedingungsfaktoren schulischer Leistungen (Helmke & Schrader 2010) durchgeführt, um möglichen Fehlspezifikationen in der Berechnung des Propensity Scores oder Balance-Problemen nach dem Matching Rechnung zu tragen (sog. „doppelt-robuste Methode“; Stuart 2010). Dazu wird die Kriteriumsvariable (Leseleistung in Klassenstufe 9) innerhalb des gematchten Samples in drei schrittweise spezifizierten Modellen auf die Klassenwiederholung (M1), dann auf relevante Kovariaten (M2) sowie schließlich gemeinsam auf die Klassenwiederholung und die relevanten Kovariaten (M3) regressiert.

Die Ergebnisse zeigen in keinem der Modelle signifikante Effekte (M1:  $\beta=0.03$ ,  $p=0.78$ ; M3:  $\beta=0.09$ ,  $p=0.38$ ) der Klassenwiederholung auf die Leseleistung in Jahrgangsstufe 9. Basierend auf unseren Analysen kann unsere Hypothese, dass eine Klassenwiederholung nicht zur Verbesserung der Leseleistung beiträgt, erhärtet werden. KlassenwiederholerInnen würden in der 9. Jahrgangsstufe die gleiche Leseleistung erreichen, hätten sie die Klasse nicht wiederholt. Die Ergebnisse bestätigen die Befunde der Studie von Klemm (2009), der Klassenwiederholungen als teure und unwirksame Maßnahme bezeichnete. Für nachfolgende Analysen legen unsere Befunde nahe, die Effektivität von Klassenwiederholungen auch für die Mathematikleistung zu prüfen.

## **Gestaltungsmöglichkeiten von Lernzeiten – Eine Studie zur Erfassung von Lernzeitkonzepten an Ganztagsgymnasien**

**Friederike Gilsbach<sup>1</sup>, Inga Wehe<sup>2</sup>, Ferdinand Stebner<sup>1</sup>, Silvia-Iris Beutel<sup>2</sup>, Joachim Wirth<sup>1</sup>**

<sup>1</sup>Ruhr-Universität Bochum, Deutschland; <sup>2</sup>Technische Universität Dortmund, Deutschland; [friederike.gilsbach@rub.de](mailto:friederike.gilsbach@rub.de)

Die Entwicklung von Halbtags- zu Ganztagsgymnasien in NRW ist eine Reaktion auf den festgestellten Zusammenhang zwischen der soziokulturellen Herkunft der Schülerinnen und Schüler (SuS) und dem Bildungserfolg: demnach absolvieren SuS bildungsferner Eltern seltener das Gymnasium (Beutel et al., 2015). Um diesem Trend entgegenzuwirken, soll die Heterogenität der Schülerschaft an Ganztagsgymnasien verstärkt berücksichtigt und deren Bildungschancen durch individuelle Fördermaßnahmen verbessert werden (Holtappels, 2006; Züchner & Fischer, 2014). Instrumente zur Erreichung dieser Ziele an Ganztagsgymnasien sind das Mehr an Zeit und die vollständige Verlagerung der Hausaufgaben in z.B. Lernzeiten (Börner et al., 2011; MSW, 2015).

Obwohl Lernzeiten ein zentrales Gestaltungselement des Ganztagsgymnasiums darstellen und an fast allen (sehr individuell) angeboten werden (Altermann et al., 2016), gibt es derzeit weder in der Literatur noch in der Praxis ein einheitliches Verständnis darüber, was eine (gute) Lernzeitkonzeption auszeichnet. Daraus resultieren Widersprüchlichkeiten und Unsicherheiten hinsichtlich der Ausgestaltung dieser Lerngelegenheit auf Seiten der Lehrkräfte (Stötzel & Tabel, 2012).

Aufgrund der diffusen und empirisch defizitären Befundlage ist es das Ziel der vorliegenden Hospitationsstudie, den Bestand aktueller Lernzeitkonzepte nordrhein-westfälischer Ganztagsgymnasien zu erfassen.

Um dieses Ziel zu erreichen, hospitierten zwei Beobachter mit Hilfe eines neu entwickelten Hospitationsbogens an elf ausgewählten Ganztagsgymnasien in den Jahrgangsstufen 5 (2 Gymnasien), 6 (8 Gymnasien) und in einer jahrgangsgemischten Lerngruppe (Schüler der Jahrgangsstufen 6-9).

Der eingesetzte Hospitationsbogen umfasst 28 Beobachtungskriterien, die aus der Literatur zur Thematik abgeleitet und u.a. an die Kriterien der Checkliste QUIGS SEK 1 (= Qualitätsentwicklung in Ganztagschulen der Sekundarstufe I; Institut für soziale Arbeit e.V., 2010) angelehnt wurden. Die Kriterien sind drei Kategorien zuordenbar:

1. Vorbereitete Lernumgebung, Lehrerrolle, Regeln. Items zu den äußeren Rahmenbedingungen: z.B. Ort, Zeit, Materialien, Personen, Regularien
2. Individuelle Förderung: Differenzierung, Forder- und Förderplanung. Items zu den Lernvorgängen: z.B. Aufgabenformate, Sozialformen, individuelle Unterstützung, Verknüpfung der unterschiedlichen Lerngelegenheiten (Fachunterricht – Lernzeit)
3. Begleitung durch Dokumente. Items zu den Instrumenten der Lernplanung und –reflexion sowie Dokumente der Lernbegleitung: z.B. Wochenplan, Portfolio

Die Ergebnisse zeigen drei Lernzeit-Typen:

In „fachgebundenen Lernzeiten“ (4 Gymnasien) bearbeiten SuS im Klassenverband in Begleitung der Fachlehrkraft Aufgaben zu einem schriftlichen Unterrichtsfach. An drei Gymnasien erfolgt eine Aufgabendifferenzierung hinsichtlich Schwierigkeit und Umfang, an jeweils einem wird hinsichtlich der Verbindlichkeit (Pflicht- und Wahlaufgaben) bzw. gar nicht differenziert.

In „freien Lernzeiten Typ I“ (2) werden SuS bei der Bearbeitung von Aufgaben zu allen schriftlichen Unterrichtsfächern im Klassenverband von der Klassenlehrerin/ dem Klassenlehrer (1) oder der Fachlehrkraft eines schriftlichen Faches (1) unterstützt. Es wurde keine Form der Differenzierung beobachtet.

In „freien Lernzeiten Typ II“ (5) bearbeiten SuS im Klassenverband (3) oder in klassen- (1) bzw. jahrgangsübergreifenden (1) Lerngruppen Aufgaben zu allen Unterrichtsfächern. Dabei werden sie von der Klassenlehrerin/ dem Klassenlehrer (2), sie unterrichtenden (1) oder weiteren Fachlehrkräften (1) begleitet. Zwei Gymnasien differenzieren hinsichtlich Verbindlichkeit, die übrigen entweder hinsichtlich der Aufgabenschwierigkeit und des Umfangs (1) oder es findet keine Differenzierung (1) statt.

Die Ergebnisse bestätigen trotz vieler konzeptioneller Gemeinsamkeiten auf der deskriptiven Ebene das in der Theorie vorgefundene Bild der Vielfältigkeit zur Strukturierung und Ausgestaltung der Lernzeiten an Ganztagsgymnasien.

Aufgrund der geringen beobachtbaren Differenzierungs-, Forder- und Fördermaßnahmen sowie der starken, verbindlichen Strukturierung dieser Lerngelegenheit (vgl. Stillarbeitsphasen im Fachunterricht) werden SuS wenig individuell gefördert und selten zum selbstregulierten Lernen angeregt. Außerdem werden ihnen nur geringe Partizipationsmöglichkeiten eingeräumt. Es wird erwartet, dass konzeptuelle Veränderungen, z.B. Aufgabengestaltung, Austausch zwischen Fachlehrern sowie die Verknüpfung der Lernzeiten mit einem Training zum selbstregulierten Lernen das förderliche Potenzial dieser zusätzlichen Zeit zum Lernen mehr ausschöpfen würden. In einer zukünftigen Studie soll ein Lernzeitkonzept entwickelt werden, das ein Training zum selbstregulierten Lernen mit dem Fachunterricht und den Lernzeiten inhaltlich verbindet.

# G01

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus S103

## **Ethnizität und "learning gender": Warum streben Jugendliche mit Migrationshintergrund geschlechtsuntypische Berufe an?**

**Alexandra Wicht, Matthias Siembab**

Universität Siegen, Deutschland; [wicht@soziologie.uni-siegen.de](mailto:wicht@soziologie.uni-siegen.de), [siembab@soziologie.uni-siegen.de](mailto:siembab@soziologie.uni-siegen.de)

### 1 Einführung und Forschungsstand

Die geschlechtsspezifischen Berufsorientierungen Jugendlicher sind ein wesentlicher Faktor für die horizontale Segregation westlicher Arbeitsmärkte in typische »Frauenberufe« und »Männerberufe« und damit verbundene Ungleichheiten, wie beispielsweise die unterschiedliche Entlohnung von Frauen und Männern (Charles & Grusky, 2004; Busch, 2013). Die Geschlechterkonnotation von Berufen ist jedoch nicht universell gültig, sondern abhängig von den soziokulturellen sowie strukturellen Gegebenheiten einer Gesellschaft. Nationalspezifische Besonderheiten von Arbeitsmärkten bilden somit einen entscheidenden Kontext, der hinter dem Rücken der Akteure Einfluss auf deren geschlechtsspezifische berufliche Orientierungen nimmt. In unserem Beitrag untersuchen wir Unterschiede in den geschlechtsspezifischen Berufsorientierungen von Jugendlichen mit und ohne Migrationshintergrund, um die Bedeutung verschiedener nationalgesellschaftlicher Kontexte herausstellen zu können. Bislang ist noch nichts darüber bekannt, inwiefern das Zusammenspiel von Geschlecht und Migration einen Beitrag zur Erklärung von Selbstselektionsprozessen Jugendlicher in geschlechtsspezifische Bewerberpools leisten kann.

### 2 Theorie und Hypothesen

Wir stützen uns auf eine allgemeine Theorie des sozialen Lernens (Akers, Krohn, Lanza-Kaduce, & Radosevich, 1979; Lent, Brown, & Hackett, 1994) und verbinden diese mit der Theorie der soziokulturellen Einbettung Jugendlicher in spezifische Kontexte (Granovetter, 1985; Portes & Rumbaut, 2001; Schulz, 2013). Aus der Perspektive der Theorie des sozialen Lernens sind Jugendliche mit Migrationshintergrund für die Erklärung geschlechtstypischer Berufsorientierungen von besonderem Interesse, da ihre Vorstellungen („frame of reference“) in Bezug auf die beruflichen Karrieren von Männern und Frauen vom jeweiligen soziostrukturellen Kontext des Heimatlands sowie des Ziellands beeinflusst sein können (Wicht, 2016). Entsprechend entwickeln wir Hypothesen mit Bezug auf das Ziel- und Herkunftsland.

Die ersten beiden Hypothesen nehmen das Zielland (Deutschland) in den Blick: Unterschiede zwischen Jugendlichen mit Migrationshintergrund und Deutschen sollten mit zunehmender Migrationsgeneration abnehmen und zwar (1) aufgrund zunehmender Informationen über die strukturellen Bedingungen des deutschen Arbeitsmarktes und/oder (2) aufgrund von Sozialisationsprozessen in der aufnehmenden Gesellschaft. Die dritte Hypothese berücksichtigt den strukturellen Kontext im jeweiligen Herkunftsland: (3) Eine niedrige Erwerbsbeteiligung von Frauen im Herkunftsland sollte die Herausbildung eher geschlechtsuntypischer Berufsorientierungen begünstigen. Denn zum einen verfügen Jungen dadurch über ein weiteres Spektrum potenzieller Berufswahloptionen, zum anderen orientieren sich Mädchen – mangels weiblicher Vorbilder auf dem Arbeitsmarkt – möglicherweise eher an den Karriereplänen von Jungen.

### 3 Daten und Methode

Wir verwenden die in Klasse 9 erhobenen Querschnittsdaten der Startkohorte 4 »Schule und Ausbildung« des Nationalen Bildungspanels und untersuchen getrennt nach Geschlecht, ob sich die geschlechtstypischen Berufsorientierungen von Jugendlichen mit und ohne Migrationshintergrund unterscheiden. Die Geschlechtstypik des angestrebten Berufs bilden wir über den Frauenanteil in dem Beruf ab, von dem Jugendliche glauben, dass sie ihn erlernen werden (realistische berufliche Aspirationen). Um die verschiedenen Hypothesen untersuchen zu können, greifen wir zusätzlich auf Informationen über berufliche Interessen und Berufswünsche (idealistische Aspirationen) zurück sowie Daten über die Erwerbsbeteiligung von Frauen in den jeweiligen Herkunftsländern.

### 4 Ergebnisse

Unsere vorläufigen Ergebnisse zeigen, dass insbesondere Jungen mit Migrationshintergrund häufiger geschlechtsuntypische Berufe anstreben als Deutsche – also Berufe mit einem vergleichsweise höheren Frauenanteil. Für Mädchen lassen sich hingegen nahezu keine statistisch signifikanten Effekte nachweisen. Die Hypothese bezüglich des Ziellands lässt sich zunächst stützen: Je länger die Jugendlichen bzw. ihre Familien in Deutschland leben, desto eher gleichen sie ihre Berufsorientierungen denen der Deutschen an, d.h. sie streben bezogen auf den deutschen Kontext geschlechtstypischere Berufe an. Dies scheint dabei am ehesten auf Sozialisationsprozesse im Zielland zurückzuführen zu sein. Auch die Annahmen zu den strukturellen Bedingungen im Herkunftsland bestätigen sich in unseren ersten Analysen – ebenfalls nur für männliche Migranten: Jungen aus Ländern mit niedriger Frauenerwerbsquote streben eher Berufe mit einem höheren Frauenanteil an. In weiteren Analysen soll unter anderem der Dissimilaritätsindex als Maß für die Geschlechtersegregation des Arbeitsmarktes im Heimatland als zusätzliches Kontextmerkmal aufgenommen werden.

## **Der Einfluss der Familiensprache Deutsch auf die Deutschleistung von SchülerInnen mit Migrationshintergrund**

**Benedict C. O. F. Fehringer, Stefanie Dotzel, Karina Karst, Stefan Münzer**

Universität Mannheim, Deutschland; [b.fehringer@uni-mannheim.de](mailto:b.fehringer@uni-mannheim.de)

### Theoretischer Hintergrund

Mittlerweile haben rund 20% der Einwohner in Deutschland einen Migrationshintergrund. Bei Kindern und Jugendlichen ist dieser Anteil mit einem Drittel noch höher (Statistisches Bundesamt, 2016). Die Abschwächung sozialer Ungleichheiten und die Integration in das deutsche Bildungssystem von Kindern mit Migrationshintergrund gelingen nur bedingt. So zeigen Studien, dass ebendiese Kinder vor allem in der Sprachkompetenz deutlich zurückliegen und weniger qualifizierte Schulabschlüsse erreichen (Hartig, Jude & Wagner, 2008; Autorengruppe Bildungsberichtserstattung, 2008). Im Zusammenhang mit dem Bildungserfolg wird der Beherrschung der Unterrichtssprache eine besondere Bedeutung zugesprochen (Rauch, Schastak & Richter, 2016). Studien unterstreichen die Relevanz des Deutschen auch als Familiensprache in Hinblick auf die Schulleistung der SchülerInnen. Bei ansonsten vergleichbarem familiärem Hintergrund und ähnlich ausgeprägten kognitiven Grundfähigkeiten haben z.B. HauptschülerInnen, die zu Hause nicht Deutsch sprechen, knapp 31 Punkte weniger im PISA-

Test erzielt als HauptschülerInnen, deren Familiensprache Deutsch ist (Stanat, 2006). Aussagen über GymnasialschülerInnen wurden nicht getroffen.

#### Fragestellung

In der vorliegenden Studie wurden vier Gruppen von SchülerInnen bezüglich ihrer Leistung in einem Deutschtest verglichen. Dabei wurden SchülerInnen mit Migrationshintergrund und Familiensprache Deutsch von Kindern mit Migrationshintergrund ohne deutsche Familiensprache und von deutschen Kindern unterscheiden. Es wurde auch überprüft, ob es einen Unterschied macht, wenn ausschließlich Deutsch oder Deutsch und andere Sprachen zu Hause gesprochen werden.

#### Methode

In dieser Studie wurden N=696 GymnasialschülerInnen (393 m, 286 w, 17 kA) der fünften Klassenstufe zweier Jahrgänge (2015, 2016) aus vier Gymnasien in 29 Klassen in Mannheim untersucht. Die Erhebung fand im Rahmen eines vom Land Baden-Württemberg geförderten Projekts zum Umgang mit Heterogenität statt. Die SchülerInnen bearbeiteten den Duisburger Sprachstandstest (Theunissen, Schoppengerd, Witzke, Endell, & Wenning, 2005). Er umfasst die Bereiche Hörverstehen, Leseverstehen, Grammatik, Satzbau, Wortschatz und Rechtschreibung (Rechtschreibung geht nicht in das Gesamtergebnis ein). Darüber hinaus wurden der Migrationshintergrund und die Familiensprachen abgefragt. Ausgehend von diesen Angaben wurden die SchülerInnen in vier Gruppen eingeteilt: Deutsche (N=250), Migranten, deren Familiensprache ausschließlich Deutsch ist (N=146), Migranten, die neben Deutsch noch mindestens eine andere Familiensprache sprechen (N=180) und Migranten, die zu Hause kein Deutsch sprechen (N=120). Andere Sprachen waren z.B. Türkisch, Arabisch, Russisch oder Polnisch. Mehrebenenanalysen untersuchten die Unterschiede dieser Gruppen bezüglich des Gesamtergebnisses und den einzelnen Testbereiche.

#### Ergebnisse

Die Mehrebenenanalysen mit den SchülerInnen auf Ebene 1 und den Klassen auf Ebene 2 wurden mit einem Random-Intercept-Only-Model berechnet, wobei das Geschlecht kontrolliert wurde. Vergleichsmodelle mit einem Random Slope waren im Hinblick auf die Informationskriterien AIC und BIC nicht überlegen. Die Gruppenvariable wurde als Treatmentvariable kodiert mit der Gruppe der Migranten, deren Familiensprache ausschließlich Deutsch ist als Referenzgruppe. Sowohl bei dem Gesamtergebnis als auch bei den einzelnen Testbereichen gab es keinen signifikanten Unterschied zwischen den deutschen SchülerInnen und den Migranten, deren Familiensprache ausschließlich Deutsch ist. Dagegen gab es signifikante Unterschiede ( $p < .05$ , nach Bonferroni-Korrektur) zwischen dieser Migrantengruppe und den beiden anderen, die zu Hause entweder neben Deutsch noch weitere Sprachen oder gar kein Deutsch sprechen. Ausnahmen waren die Kompetenz in Rechtschreibung und Satzbau, die sich nicht signifikant zwischen den Gruppen unterschieden. Im Bereich des Hörverstehens gab es nur einen signifikanten Unterschied ( $p < .05$ ) zwischen den Migranten, die zu Hause ausschließlich Deutsch sprechen und denen, die gar kein Deutsch zu Hause sprechen.

#### Diskussion

Die Ergebnisse zeigen, dass die deutsche Sprachkompetenz von Migranten signifikant besser ist, wenn die SchülerInnen auch zu Hause Deutsch sprechen. Darüber hinaus konnte gezeigt werden, dass es dabei einen Unterschied macht, ob Deutsch die einzige Familiensprache ist oder nicht. Im ersten Fall konnten keine Unterschiede in der gemessenen Deutschkompetenz zwischen Deutschen und Migranten gefunden werden. Migranten, die nur teilweise zu Hause Deutsch sprechen, zeigen hingegen signifikant schlechtere Leistungen. Inwieweit dieser Effekt womöglich durch weitere Variablen, wie z.B. das Herkunftsland bestimmt wird, müssen Folgestudien zeigen.

## G02

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus S104

### Wissenschaftsverständnis im Grundschulalter – Messbar und veränderbar?

**Julia Schiefer, Jessika Golle, Maike Tibus, Ulrich Trautwein, Kerstin Oschatz**

Universität Tübingen, Deutschland; [julia.schiefer@uni-tuebingen.de](mailto:julia.schiefer@uni-tuebingen.de)

#### Theoretischer Hintergrund

Naturwissenschaften und naturwissenschaftliche Erkenntnisse spielen eine entscheidende Rolle in unserem Alltag (OECD, 2016). Um sich an Diskussionen zu gesellschaftswissenschaftlichen Fragen beteiligen zu können, sind nicht nur Wissen und Kenntnisse in den MINT-Fächern, sondern ein angemessenes Wissenschaftsverständnis (WSV) von großer Bedeutung (OECD, 2016). Entscheidend für die Entwicklung des WSV sind sowohl epistemische Überzeugungen, als auch ein grundlegendes Verständnis für die naturwissenschaftlich-forschenden Methoden (Elby et al., 2016). Die Förderung des WSV ist ein normatives Bildungsziel, bereits bei Kindern in der Grundschule (Mullis & Martin, 2013).

Interventionen bieten neben dem Schulunterricht eine wichtige Möglichkeit, um das WSV von Schüler/innen zu untersuchen und zu fördern. Dabei stellt sich die Frage wie grundlegende Aspekte wie ein epistemisches Verständnis bereits bei Grundschulkindern gezielt gefördert werden können. Zudem existieren kaum Messinstrumente zur Erfassung des WSV (z.B. Koerber et al., 2015). Bestehende Instrumente erfassen meist nur begrenzte Aspekte oder weisen eine eingeschränkte Zuverlässigkeit und Gültigkeit auf.

In dem Beitrag wird die Entwicklung und Validierung eines neuen Messinstruments zur Erfassung des WSV sowie die Entwicklung, Evaluation und Implementation einer 10-wöchigen Intervention zur Förderung des WSV vorgestellt (Oschatz & Schiefer, 2017). Die Intervention ist Teil eines extracurricularen Enrichment-Programms und fokussiert sowohl auf die explizite Anregung wissenschaftstheoretischer Reflexion sowie die Vermittlung von naturwissenschaftlichen Methoden und Arbeitsweisen. Es wurde untersucht, ob durch die Intervention epistemischen Überzeugungen, naturwissenschaftlichen Methodenkompetenzen sowie motivationalen Dispositionen von Grundschulkindern gefördert werden können.

#### Methode

##### Stichproben und Design

Das Instrument (scientific inquiry cycle test, SIC-Test) wurde an einer Stichprobe von N = 878 Dritt- und Viertklässler/innen skaliert und validiert (Schiefer et al., 2017). Die Effektivität der Intervention wurde mittels kontrolliert randomisierter Studien sowohl unter kontrollierten Bedingungen durch die Programmentwicklerinnen (N=65 Dritt- und Viertklässler/innen), als auch unter Praxisbedingungen (N=117 Kinder, N=10 Kursleiter/innen des Enrichment-Programms) untersucht.

##### Instrumente

Der SIC-Test fokussiert auf die Erfassung des Verständnisses für den zyklischen Prozess naturwissenschaftlicher Erkenntnisgewinnung (EAP-Reliabilität = .64). Epistemische Überzeugungen wurden mit dem Fragebogen von Conley et al. (2004) erhoben. Zudem kamen Verfahren zur Erfassung der kognitiven Grundfähigkeiten (CFT20-R, Weiß, 2006 und BEFKI, Schroeders et al., 2015) und des Leseverständnisses (ELFE, Lenhard & Schneider, 2006) zum Einsatz. Weitere Instrumente zielten auf die Erfassung motivationaler Dispositionen (need for cognition, Preckel & Strobel, 2017; epistemic curiosity, Litman, 2008) sowie Experimentierstrategien (z.B. Mayer et al., 2014).

##### Statistische Analyse

Der SIC-Test wurden mit einem 2PL-Modell für dichotome Items in Mplus skaliert (Muthén & Muthén, 1998-2012). Die angenommene Eindimensionalität wurde mittels konfirmatorischer Item-Faktor-Analysen geprüft. Die Konstruktvalidität wurde anhand von Korrelationsanalysen untersucht. Die Interventionseffekte wurden mittels linearer multipler Regressionen (AVs: standardisierte Posttestwerte, UVs: Gruppenzuweisung [Intervention=1, Kontrollgruppe=0], standardisierte Prätestwerte) analysiert und können als Effektstärken (ES) interpretiert werden.

##### Ergebnisse und Diskussion

Die Ergebnisse der Item-Faktor-Analysen zeigen einen akzeptablen Modellfit (RMSEA = .035;  $\chi^2/df = 2.10$ ,  $\chi^2(54) = 113.662$ ,  $p < .000$ , CFI = 0.89, TLI = 0.86). Positive Zusammenhänge zeigen sich zwischen dem SIC-Test und Intelligenz, Textverständnis, Experimentierstrategien sowie epistemischen Überzeugungen. Die Resultate weisen darauf hin, dass der SIC-Test das Verständnis für den Wissenschaftszyklus zufriedenstellend erfasst und im Rahmen der Interventionsstudien eingesetzt werden kann.

Hierbei zeigten sich im ersten Schritt signifikant positive Effekte auf die epistemischen Überzeugungen ( $0.43 < ES < 0.61$ ) sowie die Wissbegierde der Kinder ( $ES = 0.34$ ). Die Intervention scheint somit das Verständnis der Kinder für epistemische Prozesse zu fördern. Unter Praxisbedingungen fanden sich Effekte auf das WSV (SIC-Test:  $ES = 0.31$ , Experimentierstrategien:  $ES = 0.33$ ) sowie Freude am Denken ( $ES = 0.25$ ). Die Effekte auf die epistemischen Überzeugungen und die Wissbegierde konnten nicht repliziert werden. Die Ergebnisse liefern Hinweise auf die Effektivität der Intervention und eine mögliche Implementation in die Praxis. Der Fördererfolg scheint jedoch stark von den Lehrenden abzuhängen. Die Ergebnisse werden hinsichtlich ihrer Bedeutung für die naturwissenschaftliche Förderung von Grundschulkindern diskutiert.

### Erfassung und spielerische Förderung des naturwissenschaftlichen Verständnisses im Bereich Statik bei Vorschulkindern

**Anke Maria Weber, Miriam Leuchter**

Universität Koblenz-Landau, Deutschland; [weber-a@uni-landau.de](mailto:weber-a@uni-landau.de)

#### Theorie

Das Spiel mit Bauklötzen ist besonders für junge Kinder sehr wichtig. Es verbessert ihr räumliches Vorstellungsvermögen (Casey et al., 2008) und macht die Kinder mit unterschiedlichen physikalischen Gesetzen vertraut (Adams & Nesmith, 1996). Kinder verfügen bereits früh über ein intuitives Wissen über physikalische Prinzipien (Baillargeon, 1995; Wilkening, Huber, & Cacchione, 2006). Karmiloff-Smith (1992) fand heraus, dass Vorschulkinder oft ein intuitives Wissen darüber besitzen, wie symmetrische und asymmetrische Objekte auf einer Balkenwaage balanciert werden können, ohne dass sie in der Lage wären, ihren Gedankengang zu verbalisieren. Mit zunehmender Erfahrung entwickeln sie eine verbalisierbare Mitte-basierte-Theorie und versuchen, alle Objekte mit ihrer geometrischen Mitte auf der Balkenwaage zu platzieren. Dabei scheitern sie an

asymmetrischen Objekten. Diese Theorie wird später modifiziert, indem die Kinder die Masse eines asymmetrischen Objekts einbeziehen, um es zu balancieren. Ab dann können die Kinder sowohl symmetrische als auch asymmetrische Objekte balancieren. Dieser Prozess muss nicht im Kindergarten abgeschlossen sein, sondern kann sich bis in die Grundschulzeit und darüber hinaus fortsetzen (Bonawitz, van Schijndel, Friel, & Schulz, 2012). Dennoch hat die Forschung gezeigt, dass spezifische Interventionen, die es Kindern erlauben, Erfahrungen mit dem Balancieren von Objekten zu sammeln, diese Entwicklung unterstützen können (Bonawitz et al., 2012; Pine, Messer, & Godfrey, 1999). Außerdem kann auch Scaffolding den Lernprozess von Kindern anregen (van de Pol, Volman, & Beishuizen, 2010). Die Untersuchung des kindlichen Verständnisses von Gleichgewicht kann zum Wissensgewinn über die Wandlung vom intuitiven Wissen hin zu verbalisierbaren Theorien beitragen.

#### Forschungsfragen

Es wird untersucht, (a) ob Kinder eine Massemittelpunktstrategie nutzen, um Objekte zu balancieren, (b) ob strukturiertes Material ihr Stabilitätsverständnis hinsichtlich einer solchen Strategie verbessern kann und (c) ob sie zusätzlich von verbalen Erklärungen profitieren.

#### Methode

167 Kinder zwischen fünf und sechs Jahren ( $M = 5.50$ ,  $SD = 0.49$ ) wurden in einem Messwiederholungsdesign mit Prä- und Posttest untersucht. Zu beiden Messzeitpunkten bearbeiteten die Kinder einen von Plöger & Leuchter (in Vorb.) entwickeltes Verfahren zur Messung der angewandten Strategie. Die Kinder in den beiden Experimentalgruppen erhielten strukturierte Materialien in Form von Bildkarten, anhand derer sie symmetrische und asymmetrische Zusammenstellungen von Bauklötzchen nachbauen sollten. Experimentalgruppe 1 (EG1,  $n = 55$ ) erhielt nur die strukturierten Materialien, während die Kinder aus der zweiten Experimentalgruppe (EG2,  $n = 55$ ) zusätzlich verbale Erklärungen über die Stabilität bzw. Instabilität der dargebotenen Bauklötzstrukturen erhielten. Die Kontrollgruppe ( $n = 57$ ) spielte frei mit Bauklötzchen.

#### Ergebnisse

Im Prätest wendeten nur wenige Kinder eine Massemittelpunktstrategie an, um die Stabilität von Bauklötzkonstruktionen zu beurteilen,  $M = 8.52$ ,  $SD = 4.18$  von 23 Items. Der Gebrauch der Strategie stieg im Posttest auf  $M = 10.39$ ,  $SD = 4.99$ . Eine moderierte Regression zeigte einen Anstieg des Statikverständnisses hinsichtlich der Massemittelpunktstrategie vom Prä- zum Posttest für die EG2: Material und verbale Unterstützung,  $\beta = .76$ ,  $p < .001$ . Der Anstieg war zudem höher verglichen mit der EG1: Material,  $\beta = -.51$ ,  $p < .05$  und der KG: freies Spiel,  $\beta = -.57$ ,  $p < .05$ .

#### Diskussion

Nur wenige Kinder nutzten zunächst eine Massemittelpunktstrategie, was konsistent mit bisheriger Forschung ist (Bonawitz et al., 2012). Die Intervention verbesserte das Statikverständnis der Kinder nur in der Gruppe mit zusätzlicher verbaler Unterstützung. Ein direktes Erleben des Materials reichte folglich nicht aus, sondern es bedarf zusätzlicher verbaler Scaffolds.

## Der Einfluss des Mathematikschulbuchs auf die Leistungsentwicklung von Grundschulkindern

**Ann-Katrin van den Ham, Inga Niedermeyer, Aiso Heinze**

IPN Kiel, Deutschland; [vandenham@ipn.uni-kiel.de](mailto:vandenham@ipn.uni-kiel.de)

#### Theoretischer Hintergrund und Fragestellung

Mathematikschulbücher spielen eine wichtige Rolle in der täglichen Schulpraxis. So wird dem Schulbuch eine vermittelnde Rolle zwischen dem intendierten Curriculum und dem durch die Lehrkraft implementierten Curriculum zugeschrieben (Schmidt et al., 1997, Valverde et al., 2002). Untersuchungen konnten zeigen, dass Schulbücher in der Regel die Themen und Intentionen des jeweiligen Mathematikcurriculums repräsentieren, sich jedoch in ihren konkreten Inhalten, deren Präsentation und den pädagogischen Ansätzen unterscheiden (Schmidt et al., 1997, Fan et al., 2013). Auch konnte gezeigt werden, dass Schulbücher einen Einfluss auf den Mathematikunterricht haben. So gaben 86% der Grundschullehrkräfte in Deutschland an, das Mathematikschulbuch als Basis für ihre Unterrichtsvorbereitung zu nutzen (Mullis et al., 2012). Schmidt et al. (2001) berichten einen Zusammenhang zwischen dem Themenumfang in Schulbüchern und der durch Lehrkräfte aufgewendeten Instruktionszeit zu dem Thema. Obwohl angenommen wird, dass das Schulbuch einen Einfluss auf die Schülerleistungen hat (Valverde et al., 2002), gibt es diesbezüglich nur wenige Studien und uneinheitliche Ergebnisse (Fan et al., 2013). So konnten Van Steenbrugge, Valcke und Desoete (2013) keinen Unterschied in den Schülerleistungen zwischen Klassen mit verschiedenen Mathematikschulbüchern feststellen, während Törnroos (2005) eine positive Korrelation zwischen den Lerngelegenheiten im Schulbuch und den dazugehörigen Schülerleistungen fand. Die Aussagekraft vieler Studien ist aufgrund der querschnittlichen Designs und der Tatsache, dass die einbezogenen Schulbücher unterschiedliche Curricula abbilden, begrenzt. In dieser Studie soll daher untersucht werden, welchen Einfluss Mathematikschulbücher, die auf das gleiche Curriculum ausgerichtet sind, auf die Entwicklung der Schülerleistungen in den ersten drei Schuljahren haben.

#### Methode

Es wurden Daten aus einer Längsschnittstudie in Schleswig Holstein reanalysiert. Von den insgesamt 127 Schulklassen wurden jene 93 Klassen (1664 Lernenden) ausgewählt, die eines der vier meist vorkommenden Schulbuchreihen verwendeten. Schulbücher A (Welt der Zahl) und B (Denken und Rechnen) bestehen aus einem einzigen Buch pro Klassenstufe und sind flexibel einsetzbar. Die Schulbücher C (Flex und Flo) und D (Einstern) sind pro Klassenstufe in mehrere Einzelhefte gegliedert. Bei Schulbuch C unterscheiden sich die Hefte thematisch und sind ebenfalls flexibel einsetzbar. Bei Schulbuch D ist eine lineare Bearbeitung der Hefte vorgesehen. Zu Beginn der ersten Klassenstufe wurden die mathematischen, sprachlichen und kognitiven Vorläuferfähigkeiten erhoben (Cronbach's alpha: .72-.95). Außerdem wurden Lehrer- und Elternfragebögen eingesetzt. Am Ende der ersten und zweiten Klassenstufen wurde jeweils ein Arithmetiktest eingesetzt. Diese Tests wurden mithilfe von IRT-Modellen unter Berücksichtigung eines Hintergrundmodells skaliert (EAP-PV Reliabilität: .93 und .94). Zusätzlich konnten die Summenwerte aus den Vergleichsarbeiten (VerA 3) am Ende der dritten Klassenstufe genutzt werden.

Für die Analyse der Schulbucheffekte wurden Multilevel Random Intercept-Modelle berechnet. Auf Individualebene wurden die Vorläuferfähigkeiten und auf Klassenebene die aggregierten Vorläuferfähigkeiten sowie die Lehrerqualifikation (Mathematik studiert: ja/nein) berücksichtigt. Die Schulbücher wurden als Dummy-Variablen in das Modell aufgenommen (D ist Referenzkategorie).

#### Ergebnisse

Insgesamt ergaben sich deutliche Effekte des Schulbuches auf die Schülerleistung (zusätzlich erklärte Varianz auf Klassenebene: 12% in Klasse 1, 20% in Klasse 2 und 10% in Klasse 3). Außerdem unterscheiden sich die Schulbücher in ihren

Effekten. Der Effekt von Schulbuch D ist in Klasse 1 signifikant negativ im Vergleich zu Schulbuch A und in Klasse 2 signifikant negativ im Vergleich zu Schulbuch A, B und C. Auch in den Vergleichsarbeiten am Ende der Klasse 3 ist der Effekt von Buch D signifikant negativ im Vergleich zu Schulbuch A und C. Die Ergebnisse deuten darauf hin, dass das Mathematikschulbuch als substantieller Faktor für die Entwicklung der Schülerleistung anzusehen ist. Ausgehend vom Stand der Forschung (s.o.) kann vermutet werden, dass dieser Effekt über die Qualität der unterrichtlichen Lerngelegenheiten vermittelt wird. Für die weitere Forschung stellt sich u.a. die Frage, welche Lehrkraftmerkmale die Schulbuchnutzung beeinflussen. Als Implikation für die Schulpraxis ist zu diskutieren, wie die Schulbuchqualität gesichert werden kann.

### **Die vorschulische Entwicklung von Schriftproduktion und Schriftsprachverständnis. Eine empirische Untersuchung kindlicher Konzepte von Schreiben und Schrift im Kindergarten**

**Rita Balakrishnan**

Leibniz Universität Hannover, Deutschland; [balakrishnan@psychologie.uni-hannover.de](mailto:balakrishnan@psychologie.uni-hannover.de)

Der Schriftspracherwerb beginnt weit vor dem Eintritt in die Schule, das belegen internationale Studien sehr eindrücklich. Bereits ab zwei oder drei Jahren, noch bevor sie selbst Buchstaben produzieren können, unterscheiden Kinder das Schreiben und das Malen sowohl im eigenen Produkt (Levin & Bus, 2003; Treiman & Yin, 2011) als auch im Produktionsprozess voneinander (Brenneman, Massey, Machado & Gelman, 1996;). Sie differenzieren schon im Vorschulalter Buchstaben und Ziffern (Bialystok, 1992; Dockrell & Teubal, 2007) und verwenden Charakteristika ihres Zielschriftsystems wie z.B. die Linearität oder Schreibrichtung (Chan, Juan & Foon, 2008; Ferreiro & Teberosky, 1982; Rowe & Wilson, 2015). Dabei werden unterschiedliche Phasen der vorschulischen Schreibentwicklung beschrieben: Mit dem Fokus auf kindliche Hypothesen z.B. die Phasen 1) undifferenzierte Schrift, 2) formal eingeschränkte Schrift und 3) Phonetisierung (Tolchinsky, 2004). Mit dem Fokus auf einem kindlichen statistischen Wissen über graphotaktische Merkmale wird zwischen 1) vor-phonologischer Schreibweise und 2) ausgedachter Schreibweise unterschieden (Treiman, 2017). Untersuchungen mit deutschsprachigen Kindern sind allerdings bislang stark unterrepräsentiert, obwohl die wenig schulische Ausrichtung des deutschen Kindergartens zur Folge hat, dass der Erwerbsprozess bis zum Eintritt in die Schule vergleichsweise lange weitgehend ungesteuert verläuft.

In der vorliegenden Studie wird die vorschulische Schreibentwicklung von Kindern untersucht, die in deutschen Bildungsinstitutionen mit der deutschen Schriftsprache aufwachsen. Dabei wird neben der Frage nach der Übertragbarkeit der internationalen Ergebnisse vor allem der Frage nachgegangen, wie sich die Entwicklung des Schriftbildes von Kindern beschreiben lässt, die noch nicht systematisch mit der Schrift in Kontakt gekommen sind, welche Entwicklungslogik sich daraus ableiten lässt und welche Auskunft individuelle Entwicklungsverläufe über die kindliche Herangehensweise an und das Verständnis von Schreiben und Schrift liefern können. In einer qualitativ eingebetteten Mixed-Methods-Längsschnittstudie wurden 41 Kinder im Alter von drei bis sechs Jahren über ein Jahr bei der Umsetzung von Schreibaufgaben videographiert. An drei Messzeitpunkten wurden sie aufgefordert, ihren Namen zu schreiben sowie zwei Tiere jeweils zu malen und zu schreiben. Für Detailanalysen wurden 16 der Kinder als Fallbeispiele darüber hinaus interviewt. Die explorativen Analysen der Schreibprozesse und Kommentare dieser Fallbeispiele (mit dem offenen Kodieren) bilden den Schwerpunkt der Studie. Des Weiteren finden im Sinne der deduktiven Kategorienbildung die internationalen Befunde Eingang in die quantitativen Analysen der Gesamtstichprobe.

Die Ergebnisse zeigen, dass Kinder zwischen drei und sechs Jahren Malen und Schreiben anhand ihrer Produkte quasi durchgängig voneinander unterscheiden, hier ist bereits keine Entwicklung mehr zu beobachten. Mit zunehmendem Alter steigt die Nutzung konventioneller Merkmale der Schrift wie Linearität, Schreibrichtung oder Anordnung in einer Zeile. Die Entwicklung des Schriftbildes zeigt, dass Vorschulkinder Buchstaben als wesentliches Merkmal von Schrift entdecken und bezogen auf den eigenen Namen lernen, das korrekte Schriftbild zu reproduzieren, während sie sich bei unbekanntem Schriftzügen teilweise sogar an den Wortlaut herantasten. Hinsichtlich zweier Dimensionen, die in der Entwicklung von Schriftproduktion und Schriftsprachverständnis eine wesentliche Rolle spielen, unterscheiden sich die Kinder untereinander: 1) Einsicht in die äußeren Merkmale von Schrift und 2) Einsicht in die Verbindung von Phonem und Graphem. Es scheint unterschiedliche Herangehensweisen an das sich zu eigen Machen der Schriftsprache zu geben. Einige Kinder beschäftigen sich intensiver/ länger mit den äußeren Merkmalen der Schrift. Andere Kinder entdecken und erkunden schneller den Zusammenhang vom Geschriebenen zum Gesprochenen. Man kann diese beiden Dimensionen im Erwerbsprozess voneinander unterscheiden. Die Daten zeigen, dass die ersten Schritte beim Erfassen des Zusammenhangs von Laut und Zeichen gemacht werden können, noch bevor die verwendeten ‚Schriftzeichen‘ überhaupt die Form von Buchstaben annehmen. Andersherum kann gibt es Kinder, die ihre Kenntnis um die äußeren Merkmale der Schrift (wie Linearität, Schreibrichtung oder die konventionellen Buchstabenformen) bereits weit entwickelt haben, während ihnen die Einsicht in die Phonem-Graphem-Korrespondenz noch weitgehend fehlt.

## G03

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus S105

### Zusammenhänge zwischen der professionellen Kompetenz von Mathematiklehrkräften und ihrer Fortbildungsaktivität

**Eric Richter, Dirk Richter, Alexandra Marx**

Universität Potsdam, Deutschland; [eric.richter@uni-potsdam.de](mailto:eric.richter@uni-potsdam.de)

#### Theoretischer Hintergrund und Fragestellung

Lehrerfortbildungen sind formale, zumeist auf Freiwilligkeit basierende Lerngelegenheiten. Es obliegt daher jeder Lehrkraft selbst, diese Angebote zu ergreifen, um ihre professionelle Kompetenz zu entwickeln (Kwakman, 2003). Überraschend ist, dass sich nur wenige Studien (implizit oder explizit) der Frage widmen, ob Lehrerfortbildungen überhaupt die Lehrkräfte erreichen, die sie am dringendsten benötigen (Brunner et al., 2006; Desimone, Smith & Ueno, 2006; Richter et al., 2010). Die bestehenden Befunde deuten dabei darauf hin, dass Fortbildungen vorrangig durch Lehrkräfte genutzt werden, die beispielsweise bereits über ein höheres Fachwissen verfügen (Desimone et al., 2006) oder ein höheres berufliches Engagement aufweisen (Richter et al., 2010). Demnach machen eher Lehrkräfte von Fortbildungsangeboten Gebrauch, die bereits über eine hohe professionelle Kompetenz verfügen. Vor diesem Hintergrund untersucht die vorliegende Studie, ob es sich bei den Teilnehmenden von fachbezogenen Lehrerfortbildungen hinsichtlich der professionellen Kompetenz um eine selektive Gruppe handelt und, ob die Zeit, die die Lehrkräfte in fachbezogene Fortbildungen investieren, mit Merkmalen ihrer professionellen Kompetenz zusammenhängt.

#### Methode

Die Daten der vorliegenden Untersuchung stammen aus der COACTIV-Erhebung, welche als Ergänzungsstudie im Rahmen von PISA 2003-2004 durchgeführt wurde. Die Stichprobe umfasst 229 Mathematiklehrkräfte der Sekundarstufe I allgemeinbildender Schulen. Die Merkmale der professionellen Kompetenz (Wissen, Überzeugung, Motivation, Selbstregulation) wurden mithilfe von validen und reliablen Wissenstests sowie standardisierter Erhebungsinstrumente erfasst (Kunter et al., 2013). Von den befragten Lehrkräften besuchten im Untersuchungszeitraum 154 (67.2%) eine formale Lehrerfortbildung, wobei 76 (33.1%) mindestens eine mathematik-spezifische Lehrerfortbildung und 78 (34.1%) eine nicht-fachbezogene Lehrerfortbildung besuchten. 75 Lehrkräfte (32.8%) besuchten hingegen keine Fortbildung.

#### Ergebnisse

Vorläufige Gruppenvergleichsanalysen zeigten, dass es sich bei den Teilnehmerinnen und Teilnehmern an Mathematik-Fortbildungen um eine selektive Stichprobe handelt, welche sich in Teilen signifikant von den anderen beiden Gruppen unterscheidet. So erreichten die Teilnehmenden an Mathematik-Fortbildungen eine bessere Abiturnote ( $F = 3.39$ ,  $df = 2$ ,  $p < .05$ ) und zeigten signifikant höhere Werte im Bereich Enthusiasmus für Mathematik ( $F = 4.12$ ,  $df = 2$ ,  $p < .05$ ). Vorläufige Korrelationsanalysen zeigten zudem, dass die Zeit, die Mathematiklehrkräfte bei mathematik-spezifischen Fortbildungen verbrachten, signifikant mit allen Facetten der professionellen Kompetenz zusammenhängt: Fachwissen ( $r = .26$ ,  $p < .01$ ), fachdidaktisches Wissen ( $r = .18$ ,  $p < .05$ ), konstruktivistische ( $r = .22$ ,  $p < .01$ ) und transmissive Überzeugungen über das Lernen ( $r = -.26$ ,  $p < .01$ ), Enthusiasmus für Mathematik ( $r = .26$ ,  $p < .01$ ) sowie berufliches Engagement ( $r = .16$ ,  $p < .05$ ). Keine Zusammenhänge zeigten sich hingegen zwischen der professionellen Kompetenz und der Zeit, die Lehrkräfte bei nicht-fachbezogenen Lehrerfortbildungen verbrachten.

#### Diskussion

Die Befunde zeigen, dass die Nutzung fachbezogener Lehrerfortbildungen mit Merkmalen der professionellen Kompetenz von Lehrkräften zusammenhängt. Aufgrund des querschnittlichen Forschungsdesign lässt sich jedoch nicht zweifelsfrei ergründen, ob die höhere Ausprägung der professionellen Kompetenz Ursache oder Folge der differentiellen Nutzung von fachbezogenen Lehrerfortbildungen ist. Hinweise darauf, dass es sich eher um die Nutzungsursache handeln könnte, bietet jedoch zum einen die Analyse der individuellen Kontrollvariablen sowie zum anderen die Ergebnisse von Desimone et al. (2006) und Richter et al. (2010). Eine mögliche Erklärung für diesen Befund könnte sein, dass sich Lehrkräfte mit einer höheren professionellen Kompetenz selbstbewusster fühlen und motivierter sind, fachbezogene Fortbildungen aufzusuchen, während Lehrkräfte mit einer geringeren professionellen Kompetenz eher weniger Interesse an fachbezogenen Fortbildungen besitzen könnten oder Angst davor haben könnten, den fachlichen Anforderungen hierfür nicht zu entsprechen. Implikationen für Forschung und Praxis werden diskutiert.

### Bildungswissenschaftliche Kontroversen und epistemologische Sensitivierung - Effekte auf die Veränderung epistemologischer Überzeugungen bei Lehramtsstudierenden

**Eric Klopp, Robin Stark**

Universität des Saarlandes, Deutschland; [e.klopp@mx.uni-saarland.de](mailto:e.klopp@mx.uni-saarland.de)

Eine in der Lehrerbildung geforderte Kompetenz ist die an bildungswissenschaftlichen Erkenntnissen orientierte Reflexion von Lehr- und Lernprozessen (KMK, 2004). Eine zentrale Voraussetzung hierfür sind angemessene epistemologische Überzeugungen (Lunn-Brownlee et al., 2017), die subjektive Annahmen über die Natur des Wissens beschreiben. Im entwicklungspsychologischen Ansatz werden drei Stufen epistemologischer Überzeugungen unterschieden: Absolutismus, Multiplizismus und Evaluatismus (Hofer & Pintrich, 1997). Absolutismus bezeichnet die Überzeugung, dass es nur eine objektive Wahrheit geben kann, während Multiplizismus von beliebig vielen, subjektiven Wahrheiten ausgeht. Evaluatismus beschreibt die Überzeugung, dass eine Wahrheit von Evidenz gestützt und kontextabhängig gewichtet werden muss. Somit scheint eine gezielte Förderung von Evaluatismus bei Lehramtsstudierenden angezeigt.

Ausgehend von dem Modell epistemologischen Wandels (Bendixen & Rule, 2004) haben Rosman et al. (2016) eine Intervention zur Förderung evaluatistischer Überzeugungen vorgestellt, die auf dem Prinzip auflösbarer, wissenschaftlicher Kontroversen beruht. Da Kontroversen unvereinbar mit Absolutismus und Multiplizismus sind, soll zusammen mit der anschließenden Präsentation einer Lösungsstrategie epistemologischer Wandel hin zu Evaluatismus evoziert werden. Nach Klopp und Stark (2017) lässt sich die Wirkung der Intervention durch eine epistemologische Sensitivierung (Porsch & Bromme, 2011), bei der die möglichen Gründe für bildungswissenschaftliche Kontroversen vor Darbietung derselben diskutiert werden, steigern.

Die vorliegende Studie untersucht die Frage, ob sich die Art der Bearbeitung der Kontroversen auf den epistemologischen Wandel auswirkt, wobei auf Methoden zur Förderung kritischen Denkens zurückgegriffen wird, die sich zur Veränderung epistemologischer Überzeugungen bewährt haben (Valanides & Angeli, 2005). In einer freien Bearbeitung werden die Gründe für die jeweilige Kontroverse durch eine offene Reflexionsfrage erarbeitet, während diese in einer schematischen Bearbeitung durch eine Reihe gezielter Reflexionsfragen erarbeitet werden, wobei immer Feedback über die Auflösung der Kontroverse gegeben wird.

An einem randomisierten Experiment nahmen insgesamt  $n=106$  (23 weiblich,  $M_{\text{Alter}}= 22.48$  [ $sd=4.29$ ]) Lehramtsstudierende teil, die fünf auflösbare, bildungswissenschaftliche Kontroversen erhielten. In zwei Zwischengruppenfaktoren wurde die Bearbeitung der Kontroversen (frei/schematisch) sowie die Vorgabe der Sensitivierung (mit/ohne) variiert. Absolutismus, Multiplizismus und Evaluatismus wurden in einem Pre-Posttest-Design mittels des Fragebogens von Barzilai und Weinstock (2015) erhoben. Zudem wurde eine Argumentationsaufgabe dargeboten, um das Niveau epistemologischer Überzeugungen handlungsnah zu erfassen (Mason & Sirica, 2006).

Für Absolutismus zeigte sich lediglich ein signifikanter Interaktionseffekt ( $F[1,99]=8.19$ ,  $p=.005$ ,  $\eta^2=.08$ ). Weiterhin ergab sich ein einfacher Haupteffekt der Bearbeitung in der Bedingung mit epistemologischer Sensitivierung ( $F[1,55]=4.45$ ,  $p=.040$ ,  $\eta^2=.07$ ). Auch zeigte sich ein einfacher Haupteffekt der Sensitivierung in der Bedingung freie Bearbeitung ( $F[1,49]=9.36$ ,  $p=.004$ ,  $\eta^2=.16$ ). Ein Rückgang absolutistischer Überzeugungen war nur bei Vorliegen einer Sensitivierung in Kombination mit einer freien Bearbeitung zu beobachten.

Hinsichtlich Multiplizismus ergab sich ein Effekt der Sensitivierung ( $F[1,99]=5.66$ ,  $p=.019$ ,  $\eta^2=.05$ ), ein Rückgang erfolgte nur ohne Sensitivierung. Weiterhin zeigte sich ein einfacher Haupteffekt der Sensitivierung bei freier Bearbeitung ( $F[1,49]=6.08$ ,  $p=.017$ ,  $\eta^2=.11$ ). Hier erfolgte ein Rückgang multiplizistischer Überzeugungen nur in der neutralen Bedingung bei freier Bearbeitung; in der Sensitivierungsbedingung nahm Multiplizismus hingegen noch zu.

Beim Evaluatismus ergab sich weder ein signifikanter Effekt der Bedingungen noch der Interaktion. Das gleiche Befundmuster ergab sich bei der Argumentationsaufgabe, allerdings zeigte sich explorativ ein einfacher Haupteffekt der Bearbeitung in der Sensitivierungsbedingung ( $F[1,53]=6.08$ ,  $p=.017$ ,  $\eta^2=.10$ ), bei schematischer Bearbeitung erfolgte eine Zunahme evaluatistischen Argumentierens.

Die nur bei freier Bearbeitung zu findende Reduktion von Absolutismus und Multiplizismus kann aus einer tieferen Elaboration im Gegensatz zu einer weniger tiefen Elaboration bei schemageleiteter Bearbeitung resultieren. Hinsichtlich des Effekts der Sensitivierung bei Multiplizismus ist von einem „backfire“-Effekt (Trevers et al, 2016) auszugehen, die Diskussion der Gründe wissenschaftlicher Kontroversen kann die subjektive Beliebigkeit wissenschaftlicher Erkenntnisse nahelegen. Das Ausbleiben eines Effekts bei Evaluatismus bleibt offen, allerdings deutet der einfache Haupteffekt bei der Argumentationsaufgabe auf die implizite Handlungsrelevanz im Sinn gezeigter epistemologischer Überzeugung (vgl. Limon, 2006) beim Argumentieren hin. Die Ergebnisse legen nahe, epistemologische Überzeugungen eher gezielt kontextspezifisch statt kontextübergreifend zu fördern.

## **Wie genau schätzen angehende Lehrkräfte ihr Professionswissen ein? Eine differenzierte Betrachtung der Selbsteinschätzung von Lehrkräften mithilfe metakognitiver Maße**

**Helen Ernst, Jörg Wittwer, Thamar Voss**

Albert-Ludwigs-Universität Freiburg, Deutschland; [helen.ernst@ezw.uni-freiburg.de](mailto:helen.ernst@ezw.uni-freiburg.de)

### Theoretischer Hintergrund

Fachliches, fachdidaktisches und pädagogisch-psychologisches Wissen sind wichtig, um als Lehrkraft erfolgreich zu unterrichten (z. B. Kunter, Baumert & Blum, 2011). Um ein möglichst hohes Wissen in diesen drei Bereichen zu erlangen, ist es notwendig, dass Lehrkräfte akkurat einschätzen können, wie viel Wissen sie bereits besitzen und welches Wissen ihnen noch fehlt. Nur dann können sie systematisch weiteres Wissen aufbauen. In den letzten Jahren mehrte sich Evidenz, dass Selbsteinschätzungen von Lehrkräften kaum mit ihrem tatsächlichen über Tests erfassten Wissen zusammenhängen, woraus geschlossen werden könnte, dass Lehrkräfte ihr Wissen nicht akkurat beurteilen können (z.B. König, Kaiser & Felbrich, 2012). Allerdings beruhen die Befunde gewöhnlich auf interindividuellen Korrelationen, die streng genommen keinen Rückschluss auf intraindividuelle Ebene zulassen, beispielsweise ob eine Lehrkraft ihr Wissen überschätzt. Hierzu bedarf es metakognitiver Maße, mit denen Selbsteinschätzungen auf intraindividuelle Ebene betrachtet werden können. Ungeklärt ist auch, inwieweit andere Aspekte der Professionskompetenz von Lehrkräften wie etwa ihre Selbstwirksamkeit mit ihren Selbsteinschätzungen zusammenhängen.

### Fragestellung

Wir untersuchten mithilfe metakognitiver Maße, wie genau Lehrkräfte ihr fachliches, fachdidaktisches und pädagogisch-psychologisches Wissen einschätzten und inwieweit Persönlichkeitsmerkmale einen Beitrag zur Erklärung von möglichen Verschätzungen leisteten.

### Methode

Wir analysierten Daten aus der Studie COACTIV-R mit  $N = 628$  Referendar\_innen. Hierzu wurden das fachliche, fachdidaktische und pädagogisch-psychologische Wissen (Baumert et al., 2010; Krauss et al., 2008; Voss, Kunter, Seitz, Hoehne & Baumert, 2014), Persönlichkeitsmerkmale (NEO-FFI, Borkenau & Ostendorf, 1991), Selbstwirksamkeit (Schwarzer & Jerusalem, 1999) und kognitive Grundfähigkeiten (KFT 4-12+ R; Heller & Perleth, 2000) herangezogen. Zur Untersuchung der Fragen verwendeten wir zwei unterschiedliche metakognitive Maße (Schraw, 2009; Nelson, 1984): Der Bias wurde aus der Differenz zwischen Selbsteinschätzung und Testleistung ermittelt und bezog sich darauf, inwieweit Referendar\_innen ihr Wissen in jedem Bereich über- oder unterschätzten. Die relative Veridikalität, erfasst über Gamma-Korrelationen, gab Aufschluss darüber, wie gut Referendar\_innen diskriminieren konnten, in welchen Wissensbereichen sie bessere oder schlechtere Leistungen erbrachten.

### Ergebnisse

Auf interindividueller Ebene zeigten die Ergebnisse im Einklang mit bisheriger Forschung, dass Selbsteinschätzungen und Wissen in den drei Bereichen gering bis nicht signifikant korrelierten ( $r = -.04$  bis  $r = .26$ ). In Bezug auf den Bias gab es eine ungefähr gleiche Anzahl an Referendar\_innen, die ihr Wissen in den drei Bereichen über- und unterschätzten. Während höhere Selbstwirksamkeit statistisch signifikant mit einer größeren Überschätzung einherging ( $r = .15$  bis  $r = .22$ ,  $p < .001$ ), standen höhere kognitive Fähigkeiten signifikant in Zusammenhang mit einer geringeren Neigung zur Überschätzung ( $r = -.15$  bis  $r = -.26$ ,  $p < .001$ ). Es gab keine systematischen Zusammenhänge zwischen Persönlichkeitsmerkmalen und Verschätzung, die sich für alle Wissensarten zeigten. Für die relative Veridikalität zeigte sich, dass die Referendar\_innen ihre Stärken und Schwächen



in den drei Wissensbereichen teils ungenau einschätzten (Median:  $\gamma = .33$ ). Verschätzungen über alle Wissensarten hinweg hingen jedoch nicht signifikant mit Persönlichkeitsmerkmalen und kognitiven Grundfähigkeiten zusammen.

#### Diskussion

Insgesamt verdeutlichen die Ergebnisse, dass Referendar\_innen ihr Professionswissen teils ungenau einschätzen, wobei einige der Genauigkeitsmaße differentiell mit Selbstwirksamkeitseinschätzungen und kognitiven Grundfähigkeiten zusammenhängen. Dies steht in Einklang mit anderen Ergebnissen der Metakognitionforschung, die verdeutlichen, dass Menschen generell ihre Fähigkeiten nicht besonders akkurat einschätzen können. Eine methodische Einschränkung unserer Studie ist, dass sich die Selbsteinschätzungen allgemein auf die drei Wissensbereiche bezogen und damit nicht dieselbe Metrik wie die Wissenstests aufwiesen. Bei Nutzung derselben Metrik wäre es möglicherweise einfacher für Lehrkräfte, ihre Selbsteinschätzung auf die einzelnen Aufgaben der Wissenstests zu beziehen.

### **Einflüsse kognitiver Fähigkeiten auf die professionelle Wahrnehmung – Unterschiede zwischen Bachelor- und Masterstudierenden des Grundschullehramtes**

**Christina Gippert<sup>1</sup>, Bernadette Gold<sup>2</sup>, Manfred Holodynski<sup>1</sup>**

<sup>1</sup>Westfälische Wilhelms-Universität Münster, Deutschland; <sup>2</sup>Universität Erfurt, Deutschland; [christina.gippert@uni-muenster.de](mailto:christina.gippert@uni-muenster.de)

Der Begriff der professionellen Unterrichtswahrnehmung bezeichnet „the teacher’s ability to notice and interpret relevant features of classroom events for student learning“ (Stürmer, Könings & Seidel, 2013, S. 1), wobei es sich um eine wechselseitige Interaktion von selektiver Wahrnehmung (noticing) und wissensbasiertem Schlussfolgern bzw. Interpretieren (knowledge-based reasoning) handelt (Sherin, 2007). Die Fähigkeit zur professionellen Unterrichtswahrnehmung basiert dabei auf elaboriertem und flexibel strukturiertem Wissen und gilt als bedeutsamer Aspekt der professionellen Kompetenz von Lehrkräften (ebd.).

Zur Erfassung dieser situationsgebundenen Kompetenz werden meist videobasierte Tests zur Analyse realer Unterrichtssituationen eingesetzt, da diese die Komplexität, Simultanität und Unvorhersehbarkeit tatsächlicher Geschehnisse im Klassenraum repräsentieren (Seidel & Stürmer, 2014; Sherin, 2007; Gold & Holodynski, 2017). Bei einer solchen Analyse von Unterrichtssituationen spielen, wie auch beim tatsächlichen Handeln im unterrichtlichen Kontext, die Aufmerksamkeitsfähigkeit und das Arbeitsgedächtnis aufgrund der Fülle an sensorischen Informationen, die es zu verarbeiten gilt, eine entscheidende Rolle (Feldon, 2007). ExpertInnen verfügen dabei im Gegensatz zu NovizInnen über vernetztes Wissen, weshalb sie relevante Unterrichtssituationen sowohl schneller erkennen und adäquater interpretieren können (Kunter et al., 2013) als auch bei der Analyse von Unterricht weniger kognitive Ressourcen benötigen (Feldon, 2007), die dann ggf. für das konkrete Handeln zur Verfügung stehen.

Weiterhin zeigen Studienergebnisse zu einer vergleichbaren Modellierung der professionellen Wahrnehmung (Situation Awareness; Endsley, 1995), dass die Expertise der wahrnehmenden Person den Zusammenhang zwischen der Kapazität des Arbeitsgedächtnisses und der Fähigkeit zur professionellen Wahrnehmung moderiert, sodass sich bei NovizInnen ein höherer Zusammenhang zeigt als bei ExpertInnen (Sohn & Doane, 2004). Zudem ist die Rezeption von Filmmaterial, so auch bei der videobasierten Analyse von Unterrichtssituationen, im Sinne der Kognitionspsychologie als Informationsverarbeitungsprozess zu verstehen, der durch hohe kognitive Fähigkeiten in den Bereichen der Intelligenz, der Kapazität des Arbeitsgedächtnisses und der Aufmerksamkeitsfähigkeit positiv beeinflusst wird (Schwan & Buder, 2007), sodass es zu überprüfen gilt, ob die Validität videobasierter Tests zur Erfassung der Fähigkeit zur professionellen Wahrnehmung durch den Einfluss kognitiver Fähigkeiten eingeschränkt sein könnte.

Der Beitrag befasst sich daher mit der Fragestellung, ob und inwiefern kognitive Fähigkeiten mit der mittels eines videobasierten Tests erfassten Fähigkeit zur professionellen Unterrichtswahrnehmung zusammenhängen und ob dieser Zusammenhang von dem Expertisestatus der wahrnehmenden Personen (NovizInnen vs. Fortgeschrittene) moderiert wird.

An der aktuell noch laufenden Studie nahmen bisher 50 Bachelorstudierende (M-Alter = 21.31 [SD = 1.54]) sowie 25 Masterstudierende (M-Alter = 25.16 [SD = 1.49]) des Grundschullehramtes teil, wobei insgesamt eine Stichprobengröße von jeweils 50 Personen angestrebt wird. In Gruppenerhebungen bearbeiteten die Probanden Instrumente zur Erfassung des pädagogisch-psychologischen Professionswissens und allgemeiner kognitiver Fähigkeiten (fluider Intelligenz, Arbeitsgedächtniskapazität und Aufmerksamkeitsfähigkeit) sowie einen videobasierten Test zur Erfassung der professionellen Unterrichtswahrnehmung, der sowohl die Einschätzung geschlossener Items als auch die offene Kommentierung eines Unterrichtsclips umfasst.

Erste Auswertungsergebnisse zeigen sowohl in der Gruppe der Bachelor- als auch der Masterstudierenden keinen signifikanten Einfluss der kognitiven Fähigkeiten, aber einen signifikanten Effekt der prozeduralen Facette des Professionswissens auf die Fähigkeit zur professionellen Unterrichtswahrnehmung ( $\beta = .246^*$ ). Dementsprechend konnte kein signifikanter Moderationseffekt der Gruppenzugehörigkeit gefunden werden. Weiterhin soll anhand der Auswertung der Studierendenantworten aus der offenen Analyse geprüft werden, ob sich Unterschiede im Zusammenhang der kognitiven Fähigkeiten mit den einzelnen Anforderungsbereichen der professionellen Unterrichtswahrnehmung (noticing vs. knowledge-based reasoning) ausmachen lassen. Die eingangs dargestellten bisherigen Forschungserkenntnisse lassen vermuten, dass sich ein Einfluss der kognitiven Fähigkeiten auf das noticing, d. h. das selektive Wahrnehmen relevanter Unterrichtsereignisse, zeigen könnte.

Diese ersten Ergebnisse unterstreichen die Relevanz des Professionswissens bei der Analyse komplexer Unterrichtssituationen und geben Hinweise auf die Validität des eingesetzten videobasierten Tests.

## G04

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Rosshofgasse, S02

### Veränderungen der Selbstwirksamkeitserwartungen von Lehramtsstudierenden während des Praxissemesters

**Kathrin Ding<sup>1</sup>, Carsten Rohlf<sup>2</sup>, Birgit Spinath<sup>3</sup>**

<sup>1</sup>Heidelberg School of Education; <sup>2</sup>Pädagogische Hochschule Heidelberg; <sup>3</sup>Universität Heidelberg; [ding@heiedu.ph-heidelberg.de](mailto:ding@heiedu.ph-heidelberg.de)

#### Theoretischer Hintergrund

Die Lehrer-Selbstwirksamkeitserwartung (Bandura, 1977, 2012) bezeichnet die subjektive Überzeugung einer Lehrkraft, berufsspezifische Herausforderungen aufgrund der eigenen Kompetenzen bewältigen zu können. Im Rahmen der motivationalen Orientierungen und der Selbstregulation stellt die Selbstwirksamkeitserwartung eine wichtige Komponente der Lehrerprofessionalität dar (Baumert & Kunter, 2006). Lehrkräfte mit einer hohen Erwartung an ihre eigene Wirksamkeit zeigten etwa eine größere Berufszufriedenheit (Schmitz, 1999), eine bessere psychische und physische Gesundheit (Wang, Hall & Rahimi, 2015) und eine höhere Lehrleistung (Klassen & Tze, 2014).

Da von einer zunehmenden Veränderungsresistenz der Selbstwirksamkeitserwartung im Laufe des Berufslebens ausgegangen wird (Tschannen-Moran, Woolfolk Hoy & Hoy, 1998), ist vor allem die anfängliche Entwicklung der Selbstwirksamkeitserwartung während des Lehramtsstudiums von Bedeutung. Insbesondere interessiert dabei eine mögliche Veränderung der Selbstwirksamkeitserwartung während des Semesterpraktikums, wenn die ersten Unterrichtserfolge und -misserfolge erlebt werden; denn nach Bandura (1977, 2012) stellen eigene Erfahrungen die Hauptquelle der Selbstwirksamkeitserwartungen dar. Es scheint plausibel, dass der Einfluss dieser Erfahrungen auf die Selbstwirksamkeitserwartung unter anderem davon abhängt, mit welchen Ursachen ihre Entstehung erklärt wird (Weiner, 1985, 1994); hierbei wird vom Vorliegen bestimmter Attributionsmuster ausgegangen, also individueller Präferenzen bestimmter Ursachentypen (Peterson et al., 1982).

#### Fragestellung

Für diese Studie ergeben sich damit zwei zentrale Fragestellungen:

Wie verändert sich die Selbstwirksamkeitserwartung der Lehramtsstudierenden in Bezug auf ihren späteren Lehrerberuf während des Semesterpraktikums? Zeigen sich Unterschiede in der Entwicklung zwischen den Geschlechtern?

Wie ist der Zusammenhang der Selbstwirksamkeitserwartung der Studierenden mit ihrem Attributionsverhalten? Kann das Attributionsverhalten die Veränderung der Selbstwirksamkeitserwartung über das Praxissemester vorhersagen?

#### Methode

Nach einer Pilotierung der Skalen wurden im Wintersemester 2016/17 N = 162 Heidelberger Lehramtsstudierende der Primar- und Sekundarstufe in einer Paper-Pencil-Erhebung zu Beginn und gegen Ende ihres Praxissemesters nach ihrer Selbstwirksamkeitserwartung in Bezug auf ihren späteren Lehrerberuf und nach ihrem Attributionsverhalten befragt. Die Selbstwirksamkeitserwartung wurde anhand der Skala von Schwarzer und Schmitz (1999) erhoben; das Attributionsverhalten mit Hilfe des modifizierten Attributional Style Questionnaires von Peterson et al. (1982).

#### Ergebnisse

Die Veränderung der Selbstwirksamkeitserwartung während des Praktikums sowie Unterschiede zwischen den Geschlechtern wurden anhand einer zweifaktoriellen Varianzanalyse berechnet. Dabei zeigten sich ein signifikanter Anstieg der Selbstwirksamkeitserwartung während des Praktikums,  $F(1,160) = 75.21$ ,  $p < .001$ ,  $\eta^2p = .320$ , und eine signifikant höhere Selbstwirksamkeitserwartung unter den Frauen,  $F(1,160) = 9.52$ ,  $p = .002$ ,  $\eta^2p = .056$ , aber keine Interaktion zwischen den beiden Faktoren Messzeitpunkt und Geschlecht,  $F(1,160) < 1$ .

In einer hierarchischen Regressionsanalyse zur Vorhersage der Selbstwirksamkeitserwartung am Praktikumsende wurde zunächst die Selbstwirksamkeitserwartung zu Praktikumsbeginn in das Modell aufgenommen, die 30 % der Varianz erklärte. Im zweiten Schritt wurde das Attributionsverhalten zu Praktikumsbeginn hinzugefügt, das zusätzliche 5 % Varianz aufklären konnte; als signifikant erwies sich dabei die Attribuierung von Erfolgen auf kontrollierbare Ursachen.

Insgesamt stützen diese Ergebnisse die Annahme, dass das Semesterpraktikum einen wichtigen Beitrag zur Lehrerprofessionalisierung im Bereich der motivationalen Orientierungen leistet. Auch die Vermutung eines günstigen Einflusses bestimmter Attributionsmuster auf die Selbstwirksamkeitserwartung wurde durch die Ergebnisse bekräftigt. Dieser Zusammenhang soll in einer Folgestudie auf die Attribuierung von Misserfolgen übertragen werden und könnte längerfristig als potentielle Unterstützungsmöglichkeit der Selbstwirksamkeitserwartung in der Praktikumsbegleitung berücksichtigt werden.

### Das Zusammenspiel von klassenführungsspezifischen Selbstwirksamkeits- und Handlungswirksamkeitserwartungen im Rahmen des Beanspruchungserlebens von angehenden Lehrpersonen im Praxissemester

**Christoph Schüle<sup>1</sup>, Andreas Bach<sup>2</sup>**

<sup>1</sup>Technische Universität Braunschweig, Deutschland; <sup>2</sup>Paris Lodron Universität Salzburg; [c.schuele@tu-braunschweig.de](mailto:c.schuele@tu-braunschweig.de)

Theoretischer Hintergrund: Das Praxissemester stellt im Rahmen der Professionalisierung angehender Lehrpersonen eine besonders sensible Phase ihrer Beanspruchung dar (Rothland & Boecker, 2015; Schüle et al., 2017), in deren Rahmen vor allem der Umgang mit problematischem Schülerverhalten und Unterrichtsstörungen als besonders stresshaft erlebt wird (Chaplain, 2008; Danyluk, 2013; Jones, 2006; Kokkinos & Stavropoulos, 2016; Kyriacou & Stephens, 1999; Morton, Vesco, Williams & Awender, 1997; Rieg, Paquette & Chen, 2007; Veenman, 1984; Zhai, Raver & Li-Grining, 2011). In diesem Zusammenhang gilt insbesondere die klassenführungsspezifische Selbstwirksamkeitserwartung angehender Lehrpersonen als bedeutsame Ressource gegenüber einem erhöhtem Stress- und Beanspruchungserleben (Brouwers & Tomic, 2000; Dicke et al., 2015a; Fives, Hamman & Olivarez, 2007; Zee & Koomen, 2016). Diese kennzeichnet dabei die individuelle Einschätzung der eigenen Fähigkeiten von (angehenden) Lehrpersonen, mit denen sie innerhalb ihres Unterrichtes nicht lernzielbezogene Aktivitäten der Schüler sowie zeitverschwendende Störungen der Ordnungsstruktur des Klassenzimmers bewältigen können (Brouwers & Tomic, 2000; Doyle, 1986, 2006). Diesbezüglich erleben (angehende) Lehrpersonen mit einer höheren klassenführungsbezogenen Selbstwirksamkeitserwartung eine deutlich geringere Anzahl von Unterrichtsstörungen und

problematischen Schülerverhalten und können diese schnell sowie effektiv abhandeln, wodurch sie sich einerseits weniger beansprucht fühlen (Dicke et al., 2014; Schwarzer & Hallum, 2008) und andererseits wiederum ihre klassenführungsspezifische Selbstwirksamkeitserwartung stärken können (Dicke, et al., 2015b; Holzberger, Philipp & Kunter, 2013). Unklar in diesem Wirkmechanismus bleibt trotz der Ergebnisse Skaalvik & Skaalvik (2007, 2010) dennoch die Bedeutung der von Bandura (1997) konzeptualisierten Handlungs-Ergebnis-Erwartung, die in Anlehnung an Gibson & Dembo (1984) auch als klassenführungsspezifische Handlungswirksamkeitserwartung bezeichnet werden kann. Diese umfasst dabei die Einschätzung der Effektivität von unterrichtlichen Handlungsweisen in Bezug auf die Ordnungs- und Aktivitätsstruktur sowie auf die aktive Lernzeitnutzung. Vor diesem Hintergrund erscheint insbesondere in Anlehnung an Emmer & Stough (2001) sowie Kounin (1970) und Doyle (1986) eine präventive Handlungswirksamkeitserwartung zentral, die sich vor allem auf den störungsvorbeugenden Charakter einer effektiven Klassenführung bezieht (Müller, Felbrich & Blömeke, 2008). Hierbei kann in Anlehnung an Bandura (1997) davon ausgegangen werden, dass eine derartige klassenführungsbezogene Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen den Zusammenhang zwischen der klassenführungsspezifischen Selbstwirksamkeitserwartung und der erlebten Beanspruchung vermitteln kann.

**Fragestellung & Hypothesen:** Vor dem Hintergrund dieser Annahme untersucht die vorliegende Studie das Zusammenspiel der klassenführungsspezifischen Selbstwirksamkeits- und Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen in Bezug auf das individuelle Beanspruchungserleben von angehenden Lehrpersonen im Rahmen ihres Praxissemesters und prüft, inwiefern die klassenführungsspezifische Selbstwirksamkeits- und Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen die erlebte Beanspruchung angehender Lehrpersonen im Rahmen ihres Praxissemesters beeinflussen und inwiefern der Zusammenhang zwischen der klassenführungsspezifischen Selbstwirksamkeitserwartung und der erlebten Beanspruchung durch die klassenführungsbezogene Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen mediiert wird.

**Methoden:** Zur Untersuchung dieser Fragestellungen nahmen 137 Studierende an einer Längsschnittstudie mit zwei Messzeitpunkten vor und innerhalb ihres Praxissemesters teil. Zur Operationalisierung des Beanspruchungserlebens wurde die Subskala der Emotionalen Erschöpfung des Maslach Burnout Inventorys (Maslach, Jackson & Leiter, 1996) verwendet; zur Erfassung der klassenführungsspezifischen Selbstwirksamkeitserwartung wurde die Subskala des Classroom-Management des Lehrerselbstwirksamkeitsinventariums von Tschannen-Moran & Woolfolk Hoy (2001) genutzt. Zudem wurden die klassenführungsbezogene Handlungswirksamkeit präventiver Unterrichtsmaßnahmen durch ein von Müller et al. (2008) vorgestelltes Messinstrument erfasst.

**Ergebnisse:** Die Ergebnisse einzelner Latent-Change-Analysen zeigen einen mittelstarken Anstieg des Beanspruchungserlebens sowie der klassenführungsspezifischen Selbstwirksamkeitserwartung. Demgegenüber verändert sich die klassenführungsbezogene Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen im Mittel nicht. Allerdings zeigen sich über alle Variablen hinweg bedeutsame Unterschiede in der Veränderung über die Zeit zwischen den angehenden Lehrpersonen, auf dessen Basis zusätzlich ein multivariates Latent Change Modell spezifiziert wurde. Dieses zeigt einen bedeutsamen Einfluss der klassenführungsspezifischen Selbstwirksamkeitserwartung auf die Veränderung der klassenführungsbezogenen Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen, die ihrerseits mit dem Anstieg der Emotionalen Erschöpfung negativ korreliert ist. Der damit einhergehende indirekt Effekt zeigt sich dabei insofern signifikant, als dass der Zusammenhang zwischen der klassenführungsspezifischen Selbstwirksamkeitserwartung und der Veränderung des Beanspruchungserlebens, durch die Veränderung der klassenführungsbezogenen Handlungswirksamkeitserwartung präventiver Unterrichtsmaßnahmen vermittelt wird.

## **Familiärer Hintergrund und schulische Leistungsentwicklung: Der indirekte Einfluss von elterlichen Selbstwirksamkeitsüberzeugungen und Unterstützungshandlungen**

**Edith Niederbacher, Markus P. Neuenschwander**  
PH FHNW, Schweiz; [edith.niederbacher@fhnw.ch](mailto:edith.niederbacher@fhnw.ch)

Forschungsergebnisse zu Schulleistungen zeigen konsistent, dass strukturelle familiäre Faktoren, wie das Ausbildungsniveau und der sozioökonomische Status der Eltern, wichtige Prädiktoren der schulischen Leistungen von Schülerinnen und Schülern sind (Lee & Bowen, 2006; Sirin, 2005). Dies ist eine der zentralen Erklärungen für Bildungsdisparitäten zwischen Kindern mit und ohne Migrationshintergrund (Bonsen, Kummer, & Bos, 2008; Konsortium PISA.ch, 2010). Um besser zu verstehen, wie die familiäre Herkunft auf die Leistungen wirkt, ist es notwendig, neben Strukturmerkmalen auch Prozessmerkmale familiärer Lebensverhältnisse in den Blick zu nehmen. Gemäss dem Modell von Eccles (1993) zur familiären Sozialisation entwickeln Eltern geprägt von strukturellen Merkmalen - Überzeugungen, die ihre Handlungen prägen, welche wiederum die Schulleistungen der Kinder beeinflussen. Schulbezogene Überzeugungen und Handlungen von Eltern können allgemein als Parental Involvement (Elternbeteiligung) bezeichnet werden. Die Forschungslage zu Parental Involvement und Schulerfolg ist aufgrund einer grossen Bandbreite an Definitionen und unterschiedlichen Operationalisierungen widersprüchlich. In mehreren Meta-Analysen (Castro et al., 2015; Fan & Chen, 2001; Porumbu & Necşoi, 2013) wurde festgestellt, dass insbesondere Überzeugungen von Eltern, wie beispielsweise Selbstwirksamkeitsüberzeugungen oder Bildungsaspirationen, die schulischen Leistungen der Kinder wesentlich beeinflussen. Hinsichtlich der elterlichen Handlungen, wie beispielweise Hausaufgabenunterstützung, finden sich eher schwache Zusammenhänge mit Schulleistungen. Insgesamt zeigt sich jedoch, dass strukturelle familiäre Faktoren über Prozessmerkmale der Familie vermittelt werden (Baumert, Watermann, & Schümer, 2003; Davis-Kean, 2005). Einige Studien weisen dabei darauf hin, dass hinsichtlich dieser Prozesse Unterschiede zwischen Familien mit und ohne Migrationshintergrund bestehen (Okagaki & Frensch, 1998; Yamamoto & Holloway, 2010). So haben Eltern mit Migrationshintergrund beispielsweise tiefere Selbstwirksamkeitsüberzeugungen bezüglich ihrer Lernunterstützung (Suárez-Orozco & Suárez-Orozco, 2001) und der Prozess, wie die Überzeugungen und Handlungen der Eltern auf die Schulleistungen der Kinder wirken, verläuft bei Eltern mit Migrationshintergrund unterschiedlich (Davis-Kean, 2005). Für den Schweizer Kontext liegen zu diesem Thema kaum Befunde vor.

Im vorliegenden Beitrag wurden folgende Fragen überprüft: Wie sehr wird der Effekt von strukturellen familiären Merkmalen (sozioökonomischer Status und Ausbildungsniveau der Eltern) auf die Leistungsentwicklung im Fach Deutsch über die fachspezifische elterliche Selbstwirksamkeitsüberzeugung bezüglich der Lernunterstützung erklärt? Wird der Effekt der elterlichen Selbstwirksamkeitsüberzeugung auf die Leistungsentwicklung im Fach Deutsch durch elterliche Unterstützungshandlungen (z.B. Kommunikation mit dem Kind zu Schulthemen oder Hausaufgabenkontrolle) vermittelt? Inwiefern bestehen hinsichtlich dieser Wirkprozesse Unterschiede zwischen Familien mit und ohne Migrationshintergrund (Moderatorhypothese)?

Die Daten zur Prüfung der Fragestellungen stammen aus der SNF-Längsschnittstudie «Wirkungen der Selektion WiSel». Gearbeitet wurde mit einer Teilstichprobe von 815 Schülerinnen und Schülern aus drei Deutschschweizer Kantonen, die je in der 5. und 6. Primarschulklasse Fragebogen und standardisierte Leistungstests im Fach Deutsch ausfüllten. Weiter liegen Daten aus Fragebogen von 762 Bezugspersonen/Eltern vor.

Die Ergebnisse von linearen Mehrebenenstrukturgleichungsmodellen im Gruppenvergleich zeigten, dass der Effekt des sozioökonomischen Status und des Ausbildungsniveaus der Eltern auf die Leistungsentwicklung im Fach Deutsch durch die fachspezifische elterliche Selbstwirksamkeitsüberzeugung hinsichtlich ihrer Lernunterstützung vermittelt wurden. Der Effekt der elterlichen Selbstwirksamkeitsüberzeugung auf die Leistungsentwicklung wurde jedoch nicht durch Unterstützungshandlungen vermittelt. Der Zusammenhang zwischen der elterlichen Selbstwirksamkeitsüberzeugung und der Unterstützungshandlungen war signifikant, aber bei der Gruppe der Eltern mit Migrationshintergrund signifikant stärker als in der anderen Gruppe. Eltern mit Migrationshintergrund hatten eine signifikant tiefere Selbstwirksamkeitsüberzeugung als Eltern ohne Migrationshintergrund.

Die vorliegenden Ergebnisse zeigen, dass das theoretische Modell von Eccles (1993) insgesamt für beide Gruppen bestätigt wird, allerdings konnte kein Effekt der Unterstützungshandlungen auf die Leistungsentwicklung gefunden werden. Hingegen erweisen sich die Überzeugungen der Eltern einmal mehr als sehr wirksam für die schulische Leistungsentwicklung. Die Ergebnisse weisen somit darauf hin, dass die in zahlreichen Studien festgestellten tieferen Schulleistungen von Schülerinnen und Schülern mit Migrationshintergrund nicht auf fehlendes schulisches Involvement ihrer Eltern zurückgehen. Vielmehr können die Leistungsdisparitäten zum Teil durch tiefere Selbstwirksamkeitsüberzeugungen von Eltern mit Migrationshintergrund erklärt werden.

## **Veränderung der Einstellungen Lehramtsstudierender zu Schülerheterogenität durch Lehrveranstaltungen**

**Sabine K. Lehmann-Grube, Andreas Hartinger, Markus Dresel**

Universität Augsburg, Deutschland; [sabine.lehmann-grube@phil.uni-augsburg.de](mailto:sabine.lehmann-grube@phil.uni-augsburg.de)

Einstellungen von Lehrkräften gegenüber Schülerheterogenität gelten als direkt und indirekt wirkende Einflussgröße auf ihren professionellen Umgang mit dieser Heterogenität (z.B. Glock & Kovacs, 2013; Kumar et al., 2015). Angenommen wird, dass Lehrkräfte, die Heterogenität als positiv, lernförderlich für die Schülerinnen und Schüler und durch die Lehrkraft auch zu bewältigen auffassen, eher bereit sind, sich mit den Möglichkeiten der Differenzierung (Schüle et al., 2016) oder eigenen Voreingenommenheiten (Kärner et al., 2015) auseinanderzusetzen und Integrationsansätze zu implementieren (Abegglen et al., 2015). Für die Lehrerbildung wird daraus die Aufgabe abgeleitet, Einstellungen gegenüber Schülerheterogenität direkt oder indirekt zu adressieren (Gao & Mager, 2011; Glock & Kovacs, 2013). Die Annahme, dass positive Einstellungen einen professionellen Umgang mit Heterogenität begünstigen, legt eine uni-direktionale Förderung positiver Einstellungen nahe. Andererseits kann argumentiert werden, dass überhöht positive Einstellungen mit unrealistischen Erwartungen in der Praxis zu Frustration führen, eher ungünstiges pädagogisches Handeln nach sich ziehen und deshalb relativiert werden sollten. Eine Relativierung durch vertiefte Auseinandersetzung kann dann erwartet werden, wenn Lehramtsanwärter(innen) mit hoch positiven Einstellungen einsteigen, bislang wenig persönliche Erfahrungen mit dem Heterogenitätsmerkmal gemacht haben und in Lerngelegenheiten eine (erste) Konfrontation mit authentischer Praxis erfahren und reflektieren. Eine unidirektionale positive Wirkung praxisnaher Lern- und Reflexionsgelegenheiten auf die Bewertung von Schülerheterogenität kann dagegen für solche Heterogenitätsaspekte erwartet werden, mit denen Lehramtsstudierende persönliche Erfahrungen haben und die im Studium durch vielfältige professionsspezifische Lernangebote angesprochen werden, z.B. für die Leistungsheterogenität.

Im Kontext eines Projekts zur Entwicklung interdisziplinärer, mit authentischen Praxisbeispielen arbeitender Lehr-Lernangebote wurde deshalb untersucht, inwiefern sich Einstellungen Lehramtsstudierender gegenüber verschiedenen Heterogenitätsaspekten im Laufe eines Semesters verändern, ob diese Einstellungen durch eine vertiefte Auseinandersetzung mit Schülerheterogenität in Lehrveranstaltungen gezielt beeinflusst werden und von welcher Richtung diese Einflüsse sind. 568 Studierende (Altersdurchschnitt 22.7 Jahre; 77.8 % Frauen; 48 % Grundschule; 25 % Haupt/Mittelschule, 9 % Realschule, 15 % Gymnasium; 3 % BA) füllten in 53 Lehrveranstaltungen (fachwissenschaftliche, fachdidaktische, erziehungswissenschaftliche Module) zu Beginn und am Ende einen in früheren Studien validierten Fragebogen aus (Lehmann-Grube et al., 2017), mit dem ihre Einstellungen zur Leistungsheterogenität und ethnisch-kulturellen Heterogenität von Schülerinnen und Schülern erhoben wurden (Cronbachs Alpha > .81). Von den Lehrenden wurde erfasst, ob sie in ihrer Veranstaltung einen der beiden Heterogenitätsaspekte spezifisch adressiert und bearbeitet haben (gar nicht vs. am Rande vs. zentraler Fokus). In den Analysen wurden die Veränderungen in Lehrveranstaltungen mit zentralem Fokus auf den jeweiligen Heterogenitätsaspekt mit den Veränderungen in den übrigen Lehrveranstaltungen kontrastiert.

Hierarchische Mehrebenenanalysen mit zwei Ebenen ergaben für die Einstellungen zu Leistungsheterogenität zunächst eine schwach positive Entwicklung in Veranstaltungen ohne zentrale Fokussierung auf diesen Heterogenitätsaspekt ( $\beta = .12, p < .05$ ). Die zentrale Fokussierung von Leistungsheterogenität hatte, wie erwartet, einen darüber hinausgehenden positiven Effekt auf die Einstellungen zu Leistungsheterogenität ( $\beta = .16, p < .05$ ). Die Einstellungen Lehramtsstudierender gegenüber ethnisch-kultureller Heterogenität zeigten in Veranstaltungen ohne zentrale Fokussierung auf diesen Heterogenitätsaspekt eine schwach negative Entwicklung ( $\beta = -.09, p = .05$ ). In Veranstaltungen mit zentralem Fokus auf ethnisch-kultureller Heterogenität von Schülerinnen und Schülern entwickelten sich die Einstellungen der Lehramtsstudierenden stärker negativ ( $\beta = -.19, p = .05$ ).

Die Ergebnisse zeigen, dass die vertiefte Auseinandersetzung mit einer konkreten Form der Schülerheterogenität in Lehrveranstaltungen die Einstellungen Lehramtsstudierender gegenüber genau diesem Heterogenitätsaspekt in substanziellem Ausmaß beeinflusst und der Einfluss in verschiedene Richtungen gehen kann. Durch vertiefte Auseinandersetzung mit Heterogenität können also nicht nur Einstellungen gegenüber Maßnahmen, z.B. multikulturelle Ansätze (Grant & Secada, 1990) oder Inklusion (Gao & Mager, 2011), sondern gegenüber Heterogenität als Phänomen und Anforderung beeinflusst werden. Dass Lehrangebote ohne zentrale Fokussierung des jeweiligen Heterogenitätsaspekts Effekte in gleicher Richtung (allerdings in deskriptiv geringerem Ausmaß) hatten, spricht dafür, dass die Einstellungen Lehramtsstudierender im Verlauf des Studiums Veränderungen unterliegen und durch weitere Faktoren beeinflusst werden.

### Validly Measuring Domain-specific Knowledge in Higher Education: Examining Test Fairness Through Analysis of Non-Responses

**Jasmin Schlux<sup>1</sup>, Olga Zlatkin-Troitschanskaia<sup>1</sup>, Carla Kühling-Thees<sup>1</sup>, Judith Jitomirski<sup>2</sup>, Roland Happ<sup>1</sup>, Sebastian Brückner<sup>1</sup>, Manuel Förster<sup>1</sup>, Hans Anand Pant<sup>2</sup>**

<sup>1</sup>Johannes Gutenberg-Universität Mainz, Deutschland; <sup>2</sup>Humboldt Universität Berlin; [jasmin.schlax@uni-mainz.de](mailto:jasmin.schlax@uni-mainz.de)

Development of models and instruments to assess students' knowledge and skills in higher education in an objective, valid and reliable way has become an important research area over the years (e.g., Zlatkin-Troitschanskaia & Pant, 2016c). Especially in Germany, domain-specific knowledge tests for university students were rare (Zlatkin-Troitschanskaia et al. 2016b). Following the Standards for Educational and Psychological Testing (AERA et al., 2014), we first validated the German adaptation of the U.S. Test of Understanding in College Economics (TUCE IV; Council for Economic Education, Walstad et al., 2007) and the U.S. Test of Economic Literacy (TEL IV; Walstad et al., 2013) in a German large-scale study in higher education (Zlatkin-Troitschanskaia et al., 2014). The particular focus of the current follow-up study is on in-depth evaluation of test fairness. Our validation analyses revealed there might be a number of fairness issues, e.g., significant language-related effects on student test performance (Brückner et al., 2015). For instance, controlled for language skills, German native speakers showed higher performance on the economic knowledge tests than non-native speakers. To explain this finding, we conducted a follow-up nation-wide assessment study in the winter term 2016 including N=9,055 students from 54 universities (Zlatkin-Troitschanskaia et al., 2016a).

Analyzing the missing data, we found that 2,456 students (out of 9,055) showed at least one non-response in the economics knowledge test, 527 students showed non-response in more than 25% of the test items. Regression analyses revealed that the non-responses can be predicted by native language: non-native speakers showed more missing values in the test than German native speakers. The non-native speakers also showed lower total knowledge values overall, when missing values were coded as false answers (mean (native speakers) = 13.33±4.39 (N=8,702) vs. mean (non-native speakers) = 8.36±4.20 (N=286),  $t(305.77)=-19.63$ ,  $p<0.0001$ ). The group with more than 14 non-responses in the test items comprised 31.96% (31 out of 97) participants who grew up in another country compared to 3.37% (302 out of 8958) in the group with less than 14 non-responses.

The results suggest that non-response might also be due to a language-related bias. If these non-responses are partly due to language misunderstanding, construct-irrelevant factors have an influence on the item response behavior of test takers. The established practice for analyzing respective knowledge test items is to regard missing values as false answers, interpreting them as a low knowledge test score (Schnell et al., 2011). Consequently, test takers with many non-responses have lower overall test scores. If missing values in test items were influenced by language, they are not missing completely at random (Rubin, 1987). In this case, test score interpretations may be invalid due to possible bias through unequal distribution of missing values. To minimize problems in test fairness and validity issues, imputation procedures were applied to prevent non-native speakers from being disadvantaged and misjudged. We used multiple imputation procedures (Rubin, 1987) for knowledge items. Regarding the 67 (0.74%) non-responses in the item assessing language skills, we conducted single imputations (Rubin, 1987). After imputation, the unjustified gap (likely due to language skills) in economic knowledge between native and non-native speakers was reduced: non-native speakers still showed significantly lower knowledge total values, but the difference was reduced (mean (non-native speakers) = 9.88±0.37 (n=286), mean (native-speakers) = 13.90±0.05 (n=8,702),  $t(16.01)=-10.66$ ,  $p<0.0001$ ).

Overall, if non-response on a domain-specific knowledge test is predictable through language skills, the assessment is not fair and it systematically puts non-native speakers at a disadvantage. According to the consequential validity (AERA et al., 2014), the in-depth analyses regarding implications for test development, validity, fairness, and the consequences of testing will be critically discussed.

### Validierung geeigneter Stichprobengewichte der PISA Studie im hierarchischen Modell

**Julia Mang<sup>1,2,3</sup>, Helmut Küchenhoff<sup>4</sup>, Manfred Prenzel<sup>1,2</sup>, Sabine Meinck<sup>5</sup>**

<sup>1</sup>TU München, Deutschland; <sup>2</sup>TUM School of Education, Deutschland; <sup>3</sup>Zentrum für internationale Bildungsvergleichsstudien (ZIB), Deutschland; <sup>4</sup>Institut für Statistik, Ludwig-Maximilians-Universität München, Deutschland; <sup>5</sup>International Association for the Evaluation of Educational Achievement (IEA), Deutschland; [julia.mang@tum.de](mailto:julia.mang@tum.de)

#### Theoretischer Hintergrund

Das Programme for International Student Assessment (PISA) unterliegt einem komplexen Stichprobenziehungsdesign. Um die Ziel-Population aller 15-Jährigen Schülerinnen und Schülern in Deutschland adäquat zu repräsentieren, kommt ein zweistufiges Stichprobenziehungsverfahren zum Einsatz (Heine et al., 2016). In einem ersten Schritt werden Schulen proportional zu ihrer tatsächlichen Verteilung in den 16 Bundesländern und ihrer Schularzt (sog. Stratifizierung) mit dem sog. PPS-Verfahren (Probabilities Proportional to Size) gezogen. Der zweite Schritt beinhaltet die Zufallsziehung einer festen Anzahl von 15-jährigen Schülerinnen und Schüler in diesen Schulen. Man spricht in diesem Fall von informativer Stichprobenziehung auf der Ebene der Schulen (Pfeffermann et al., 1998).

Die Verwendung hierarchischer Modelle zur Analyse von Daten aus Schulleistungsstudien sind in mehrerer Hinsicht vorteilhaft. Einerseits bilden sie die real existierenden Mehrebenenstrukturen des Bildungssystems mit Schülern innerhalb des Klassen- bzw. Schulverbandes besser ab. Weiterhin werden die Clustereffekte auf Stichprobenfehler im Modell berücksichtigt, denen sonst durch die Verwendung spezieller aufwändiger Schätzverfahren (bspw. Balanced Repeated Replication in PISA, Heine et al., 2016) Rechnung getragen werden muss. PISA Daten reflektieren innerhalb eines Landes zwei Hierarchieebenen. Eine Ebene wird dabei durch die Schule bzw. die Klasse und ihre Eigenschaften dargestellt, die andere Hierarchieebene umfasst die Eigenschaften der Schülerinnen und Schüler innerhalb dieser Schulen.

Die üblichen Verfahren zur Inferenzschätzung (sog. Horvitz-Thompson Schätzer, vgl. Horvitz & Thompson, 1952) können im Falle von hierarchischen Modellen nicht einfach übernommen werden (Bertollet, 2008). Es liegen verschiedene Ansätze für die Schätzverfahren in hierarchischen Modellen vor (Asparouhov, 2006; Pfeffermann et al., 1998; Rabe-Hesketh & Skrondal, 2006; Korn & Graubard, 2003). All diese Ansätze beruhen auf einem designbasiertem Modell der Inferenzschätzung mit Hilfe der Pseudo Maximum Likelihood Methode (Snijders & Bosker, 2012; Pfeffermann, 1993) und verwenden adjustierte Gewichte für die entsprechenden Hierarchien.

In dieser hier vorgelegten Untersuchung werden zwei Adjustierungsverfahren für die Stichprobengewichte verglichen. Im ersten Verfahren werden die Gewichte der Schülerinnen und Schüler so angepasst, dass sie sich innerhalb der Schulen auf die effektive Stichprobengröße summieren. Im zweiten Verfahren werden diese Gewichte auf die eigentliche Stichprobengröße der Schule normiert (Pfeffermann et al., 1998).

Aus mehreren Simulationsstudien (Asparouhov, 2004; Bertolet, 2008; Rabe-Hesketh & Skrondal, 2006; Korn & Graubard, 2003) geht hervor, dass es kein eindeutiges Schätzverfahren bzw. keine eindeutig zu präferierende Adjustierung der Gewichtung gibt, obwohl eine Tendenz in Richtung der zweiten Adjustierung geht (Rabe-Hesketh & Skrondal, 2006). Vielmehr ist das Stichprobendesign an sich für die Wahl des Schätzverfahrens ausschlaggebend. Des Weiteren führen unterschiedliche Softwareprogramme zu unterschiedlichen Ergebnissen (Chantala & Suchidran, 2006).

#### Fragestellung

Anhand einer Simulationsstudie soll das Schätzverfahren mit entsprechenden Stichprobengewichten herausgearbeitet werden, welches die geringste Verzerrung zu den tatsächlichen Populationskennwerten aufweist. Dabei wird der Fokus insbesondere auf das Stichprobendesign der PISA Studie gelegt. Die Analysen werden mit zwei in Deutschland oft verwendeten Software Programmen für hierarchische Modellierung von Large Scale Assessment Daten MPlus (Muthén & Muthén, 2017) und HLM (Raudenbush, 2007) durchgeführt.

#### Methode

Die PISA Population der 15-Jährigen Schülerinnen und Schüler wird analog dem Vorgehen von Meinck & Vandenplas (2012) mit Hilfe einer Monte Carlo Simulation generiert. Dabei werden die Populationsparameter so gewählt, dass sie einer realistischen PISA-Zielgruppe entsprechen. Leistungsvariablen, eine sozioökonomische Verteilung sowie ein hierarchisches Design mit zwei Ebenen und entsprechenden Korrelationen zwischen diesen Ebenen sind hierbei ausschlaggebend. Auf Basis dieser Simulation können entsprechende Gewichte direkt abgeleitet werden und somit geschätzte Populationswerte mit den tatsächlichen Größen ihrer Zielgruppe in Deutschland verglichen werden.

#### Ergebnisse

Ziel der Untersuchung ist ein dezidiertes Vorschlag für die gängige Schätzung hierarchischer Modelle mit den Daten der PISA Studie. Zum jetzigen Zeitpunkt liegen noch keine Simulationsergebnisse vor.

## **Mehrebenenanalysen mit Bayes-Verfahren: Ein Kommentar zur Verwendung des potential scale reduction factors zur Feststellung von Konvergenz**

**Steffen Zitzmann**

Institut für pädagogisch-psychologische Lehr- und Lernforschung, Deutschland; [szitzmann@ipl.uni-kiel.de](mailto:szitzmann@ipl.uni-kiel.de)

#### Theoretischer Hintergrund

Die empirische Bildungsforschung versucht die Zusammenhänge und Wirkmechanismen zwischen Determinanten erfolgreichen Lernens und dem Lernerfolg zu identifizieren und basierend darauf Lernprozesse zu optimieren. Die zu diesem Zweck erhobenen Datensätze weisen nicht selten eine Mehrebenenstruktur auf, in der Schüler in Schulklassen geschachtelt sind. Unter den Determinanten erfolgreichen Lernens spielen Kontextmerkmale (z.B. Unterrichtsqualität) eine wichtige Rolle. Eine Strategie zur Erfassung von Kontextmerkmalen besteht darin, Variablen einer unteren Analyseebene auf eine höhere Analyseebene zu aggregieren. Beispielsweise wird zunächst die von den Schülern eingeschätzte Unterrichtsqualität auf Klassenebene aggregiert und dann der Effekt der Unterrichtsqualität auf den Lernerfolg mit Hilfe von Mehrebenenanalysen (z.B. Raudenbush & Bryk, 2002) geschätzt.

Mehrebenenmodellen erfreuen sich in der empirischen Bildungsforschung großer Beliebtheit. Es wurde allerdings an vielen Stellen darauf hingewiesen, dass diese Modelle vergleichsweise hohe Anforderungen an die Daten stellen, und dass Schätzprobleme (z.B. keine Konvergenz, degenerierte Lösung) auftreten, wenn diese Anforderungen nicht erfüllt sind. Lüdtke et al. (2008; vgl. auch Marsh et al., 2012) diskutieren zum Beispiel ein Mehrebenenmodell, das den Gruppenmittelwert als latente Variable modelliert und die Schätzung von Kontexteffekten sowohl für den Messfehler als auch den durch die Aggregation entstehenden Stichprobenfehler korrigiert. Simulationsstudien haben allerdings gezeigt, dass dieses „doppelt-latente“ Mehrebenenmodell selbst bei größerer Anzahl (z.B. 100) Schulklassen noch zu Schätzproblemen führen kann (Lüdtke, Marsh, Robitzsch, & Trautwein, 2011).

Aufgrund der hohen Anforderungen an die Daten und der Schätzprobleme werden in der methodischen Literatur als Alternative Bayes-Verfahren empfohlen (Gelman & Hill, 2007; Hamaker & Klugkist, 2011; Kaplan, 2014; Swaminathan & Rogers, 2008). Bayes-Verfahren stehen mit der prominenten latenten Variablen-Software Mplus (Asparouhov & Muthén, 2012) zur Verfügung.

#### Fragestellung

Eine Herausforderung, die mit der Verwendung von Bayes-Verfahren einhergeht, besteht darin festzustellen, ob diese Verfahren konvergieren. Die Software Mplus nimmt diese Entscheidung den Anwendern ab, indem es automatisch überprüft, ob der potential scale reduction factor (PSR, Gelman & Rubin, 1992) kleiner 1.05 ist. Im vorliegenden Beitrag wird dieses Kriterium kritisch evaluiert. Es wurde untersucht, ob der PSR in für die empirische Bildungsforschung typischen Datenkonstellationen (z.B. 100 Schulklassen) hinsichtlich der Schätzung der Effekte von Kontextmerkmalen ein zuverlässiges Kriterium darstellt. Es wurde insbesondere untersucht, ob der PSR zu stabilen Schätzergebnissen führt.

#### Methode

In einer Simulationsstudie untersuchte ich das statistische Verhalten (insbesondere RMSE) des PSRs und verglich die Ergebnisse mit denen alternativer (automatisierter) Ansätze zur Feststellung von Konvergenz. Das datengenerierende Modell ist ein Mehrebenenmodell mit einem Prädiktor auf Schülerebene und einem weiteren Prädiktor auf Klassenebene. Es wurden die Klassenanzahl ( $J = 50, 100, 200, 300$ ) sowie die Klassengröße ( $n = 5, 10, 20, 30$ ) variiert. Insgesamt umfasste das Simulationsdesign 16 Zellen. Für jede Zelle wurden 1000 Datensätze generiert und mit den verschiedenen Ansätzen zur Feststellung von Konvergenz analysiert.

#### Ergebnisse

Der PSR lieferte vor allem bei einer kleinen Klassenanzahl wenig stabile Schätzungen (großer RMSE) des Effekts des Prädiktors auf Klassenebene. Er lieferte außerdem weniger stabile Schätzungen als einige der alternativen Ansätze (z.B. Hoff-Kriterium: Verhältnis aus time series error zur Standardabweichung der Markov-Chain-Monte-Carlo-Kette).

#### Diskussion

Während in der bisherigen Literatur zur Schätzung von Mehrebenenmodellen mit Bayes-Verfahren die Prior-Verteilung im Fokus der Diskussion stand (z.B. Zitzmann, Lüdtke, & Robitzsch, 2015), diskutiert dieser Beitrag erstmals das Problem der Feststellung von Konvergenz. Es konnte gezeigt werden, dass das von Mplus standardgemäß genutzte Kriterium „PSR < 1.05“ in für die empirische Bildungsforschung typischen Datenkonstellationen (z.B. 100 Schulklassen) kein besonders zuverlässiges Kriterium darstellt.

## **Die Einführung eines IRT-basierten Tests in eine bildungswissenschaftliche Vorlesung – erste Ergebnisse zu Skalierung und Effekten des Darbietungsmediums**

**S. Franziska C. Wenzel, Sabine Fabriz, Horz Horger**

Goethe Universität Frankfurt am Main, Deutschland; [wenzel@psych.uni-frankfurt.de](mailto:wenzel@psych.uni-frankfurt.de)

Die Forderung nach einem häufigeren Einsatz von IRT-basierten Tests in der Hochschullehre ist nicht neu (z.B., Weiss & Kingsbury, 1984), jedoch ist der zusätzliche Aufwand der Entwicklung solcher Instrumente verhältnismäßig hoch. Die zusätzliche Investition lohnt insbesondere dann, wenn es sich um große, sich jährlich oder semesterweise wiederholende Veranstaltungen mit klarer curricularer Verankerung handelt, (für Bildungswissenschaften: Standards für Lehrerbildung, KMK, 2004).

Die Erfassung akademischen Lernerfolgs zählt zu den Kernelementen von Hochschullehre und nimmt innerhalb der Triangulation mit Lernzielen und Lehr-Lernaktivitäten eine wichtige Rolle ein (Biggs & Tang, 2011). Assessments liefern Informationen über den Lehr-Lernprozess, haben das Potential die Lernentwicklung von Studierenden zu verbessern (Kingston & Nash, 2011) und liefern relevante Hinweise für Lehrende.. Trotz ihrer wichtigen Rolle unterliegen Prüfungen an Hochschulen jedoch weitestgehend keiner psychometrischen Prüfung und weisen bekannte inhaltliche Schwächen auf (z.B. Testknacker-Phänomen; geringe Veranstaltungsbezogenheit etc.). Dabei ist diagnostische Präzision unabdingbar für die valide Interpretation von formativen oder summativen Prüfungsergebnissen. Den bekannten Schwächen kann durch das Potential von IRT-basierten Tests begegnet werden (Item Response Theory, z.B. van der Linden & Hambleton, 1997). Auch bieten IRT-skalierte Aufgabenpools die Grundlage für computerbasiertes adaptives Testen (CAT) bei dem ein, auf die Fähigkeiten des Prüflings zugeschnittener Test präsentiert wird, abhängig vom gezeigten Antwortverhalten (Bjorner, Chang, Thissen & Reeve, 2007). Zudem liegen Evidenzen vor, dass CATs, verglichen mit traditionellen paper-pencil Prüfungen, die Motivation von Studierenden steigern können (Maravić Čisar, Čisar & Pinter, 2016) und als fair wahrgenommen werden (McCoubrie, 2004).

Fragestellung:

In unserer Studie stellen wir die Entwicklung eines IRT -basierten Tests in einer bildungswissenschaftlichen Vorlesung auf Basis existierender Prüfungsaufgaben vor. Dies ist ein Schritt auf dem Weg zur Entwicklung und Implementierung eines CAT für diesen spezifischen Inhaltsbereich als formativen oder summativen Test. Für eine systematische Prüfung des vorliegenden Aufgabenpools verfolgten wir folgende Teilziele:

(1) Prüfung der Itemeigenschaften, Itemselektion und Schätzung der Itemparameter: Für einen Aufgabenpool mit hoher psychometrischer Güte sollen bestehende Aufgaben basierend auf dem Rasch-Modell skaliert werden. Aufgaben mit ungünstigen Eigenschaften sollen eliminiert oder zur Überarbeitung identifiziert werden. Die Passung zwischen den Verteilungen von Aufgabenschwierigkeit und Personenfähigkeit soll darüber hinaus geprüft werden, um Hinweise für zukünftige Aufgabenentwicklungen abzuleiten.

(2) Effekte der Darbietungsform: Mit der geplanten Implementation eines CAT in die Vorlesung geht ein Wechsel der Darbietungsform einher. Um potenzielle Gefahren für die Testfairness in den Blick zu nehmen, prüfen wir mit dieser Studie, inwiefern Studierende von der Art der Darbietung profitieren.

Methode:

Für eine Stichprobe von N = 673 Studierenden und einen Pool von 40 MC Aufgaben lagen Daten für eine traditionelle Papierklausur und eine Onlineversion vor.

Neben der Itemschwierigkeit wurden die Passung der einzelnen Aufgaben mit dem Rasch-Modell anhand der gewichteten Abweichungsquadrate (weighted mean square; WMNSQ) und des entsprechenden t-Wertes, sowie die Korrelation des Aufgabentestwertes mit dem Testwert der Gesamtskala als Kriterien für die psychometrische Güte berücksichtigt. Die Personenparameter wurden über einen Maximum-Likelihood-Ansatz geschätzt. Der Effekt des Darbietungsmediums auf die Lösungswahrscheinlichkeit einzelner Aufgaben wurde über Analysen zum differentiellen Itemfunktionieren untersucht (DIF; Osterlind & Everson, 2009).

Ergebnisse:

Insgesamt wurden 29 Aufgaben ausgewählt. Das geschätzte Mittel der Schwierigkeitsverteilung lag bei -1.02 (SD = 1.18) und die mittlere punktbiseriale Korrelation zwischen den Ergebnissen einzelner Items und dem Gesamtwert bei .36 (range: .16 - .51). Die Prüfung eines möglichen Darbietungsformeffekts ergab einen leichten aber nicht signifikanten Vorteil für die Papierversion ( $\chi^2 = 1.78$ ,  $df = 1$ ). Die Gesamtinteraktion zwischen Items und Medium war ebenfalls nicht bedeutsam ( $\chi^2 = 39.20$ ,  $df = 29$ ).

Mit den Ergebnissen liegen günstige Voraussetzungen für die Weiterentwicklung des Tests als formativen und summativen CAT vor. Für die Weiterentwicklung des Aufgabenpools konnten Erkenntnisse aus dem Itemprüfungs- und selektionsprozess gewonnen werden.

## G06

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus S212

### Der Einfluss nicht-kognitiver Faktoren auf die Kompetenzentwicklung im Übergang vom Kindergarten in die Grundschule

**Daniel Schmerse<sup>1</sup>, Steffen Zitzmann<sup>2</sup>**

<sup>1</sup>Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland; <sup>2</sup>Christian-Albrechts-Universität zu Kiel, Institut für Pädagogisch-Psychologische Lehr- und Lernforschung; [schmerse@ipn.uni-kiel.de](mailto:schmerse@ipn.uni-kiel.de)

#### Theoretischer Hintergrund

Der Übergang vom Kindergarten in die Grundschule markiert eine bedeutende Phase in der Bildungsbiografie von Kindern. Die erfolgreiche Bewältigung des Übergangs zeigt sich u.a. in der Entwicklung positiver Beziehungen zu Peers und Erwachsenen, einer positiven Einstellung zum Lernen sowie an Lernfortschritten in schulleistungsbezogenen Kompetenzen. Zahlreiche Befunde aus internationalen und nationalen Studien heben dabei die Bedeutung kognitiver Faktoren und domänen-spezifischer Vorläuferfähigkeiten insbesondere für die mathematische und schriftsprachliche Kompetenzentwicklung hervor (Anders et al., 2012; Duncan et al., 2007; Hasselhorn et al., 2012; Sammons et al., 2008). Auf Seiten der nicht-kognitiven Faktoren standen bisher vorwiegend sozial-emotionale Fähigkeiten sowie die Akzeptanz bzw. Zurückweisung durch Peers als Prädiktoren schulischer Leistungsentwicklung im Fokus (Flook et al., 2005; Ladd et al., 1999; Valiente et al., 2008). Gleichzeitig ist der Einfluss zentraler nicht-kognitiver Faktoren wie Anstrengung und Beharrlichkeit bei schulischen Herausforderungen (academic perseverance, Farrington et al., 2012) in der Transition zwischen den frühen Bildungsetappen bislang kaum untersucht – trotz zahlreicher Hinweise für deren Zusammenhänge mit schulischen Leistungsmaßen bei SuS in höheren Bildungsstufen (Credé et al., 2017). Da die Bewältigung des Übergangs in die Grundschule von Kindern nicht nur die Integration in ein neues soziales Umfeld, sondern auch die Anpassung an schulische Lernanforderungen verlangt, kann davon ausgegangen werden, dass Indikatoren der Anstrengungsbereitschaft ebenso wie die soziale Eingebundenheit eine bedeutsame Rolle für die frühe Entwicklung schulleistungsbezogener Kompetenzen spielen.

#### Fragestellungen

Vor diesem Hintergrund adressiert der vorliegende Beitrag folgende Forschungsfragen:

- (1) Welche Bedeutung haben (a) soziale Eingebundenheit und (b) Anstrengung und Beharrlichkeit bei Herausforderungen für die Vorhersage schulleistungsbezogener Kompetenzen in der zweiten und dritten Klasse?
- (2) Stellen die nicht-kognitiven Faktoren vermittelnde Einflussgrößen zwischen vorschulischen kognitiven sowie sozial-emotionalen Fähigkeiten und domänen-spezifischer Kompetenzentwicklung dar?

#### Methode

Als Datengrundlage diente die Startkohorte 2 des Nationalen Bildungspanels (NEPS, Blossfeld, Roßbach & von Maurice, 2011) mit 2706 Kindern und vorliegenden Informationen aus fünf Erhebungswellen zwischen dem 6. und 10. Lebensjahr. Anhand von Lehrkräfteeinschätzungen wurden die soziale Eingebundenheit in die Klasse (2 Items,  $\alpha = .80$ ) sowie Anstrengungsbereitschaft und Beharrlichkeit (2 Items,  $\alpha = .81$ ) in der 1. Klasse erfasst. Die mathematische und naturwissenschaftliche Kompetenz wurde mittels speziell für NEPS entwickelter Tests (Hahn et al., 2013; Neumann et al., 2013) im Kindergarten und in der 2. (Mathematik) bzw. 3. Klassenstufe (Naturwissenschaft) erhoben. Die Erfassung der Leseleistung erfolgte durch den ELFE-Test (Lenhard & Schneider, 2006) in der 2. Klasse sowie der frühen Buchstabenkenntnis im Kindergarten. Zusätzlich wurden zahlreiche vorschulische kognitive und sozial-emotionale Prädiktoren in die Analysen einbezogen. Die Auswertung der Daten erfolgte anhand von Strukturgleichungsmodellen. Die Clusterstruktur der Daten wurde durch Standardfehler-Korrektur berücksichtigt.

#### Ergebnisse

Nach Kontrolle des elterlichen Bildungsstandes sowie Alter und Geschlecht ergaben sich drei wesentliche Befunde. Erstens, je höher die Anstrengungsbereitschaft und Beharrlichkeit bei Herausforderungen eines Kindes eingeschätzt wird, desto höher fallen die Leistungen im Leseverständnis und im naturwissenschaftlichen, nicht aber im mathematischen Kompetenztest unter Kontrolle des vorschulischen Ausgangsniveaus aus. Für die eingeschätzte soziale Eingebundenheit zeigen sich keine Effekte auf die Leistungen in den Kompetenztests. Zweitens, für alle drei Kompetenzdomänen zeigen sich Effekte der jeweiligen vorschulischen Testleistung. Zusätzlich erweisen sich frühe mathematische Fähigkeiten als Prädiktor für die spätere naturwissenschaftliche Kompetenz. Drittens, der nicht-kognitive Faktor Anstrengung und Beharrlichkeit mediiert Zusammenhänge der Aufmerksamkeitsfähigkeit, des Problemverhaltens und der mathematischen Kompetenz mit der schulischen Leseleistung, stellte in den anderen Kompetenzdomänen jedoch keine vermittelnde Einflussgröße dar. Die Ergebnisse bestätigen Befunde internationaler Längsschnittstudien sowohl im Hinblick auf die Bedeutung früher mathematischer Fähigkeiten als auch im Hinblick auf die eher geringen Zusammenhänge sozial-emotionaler Faktoren mit der Leistungsentwicklung (Duncan et al., 2007; Pagani et al., 2010; Romano et al., 2010). Gleichzeitig erweitern die Befunde die Erkenntnisse zu diesen Einflussmustern, indem sie die Bedeutung nicht-kognitiver Faktoren wie Anstrengung und Beharrlichkeit beim Übergang in die Grundschule herausstellen.

### Selbst- und fremdbezogene Kausalattributionen bei Kindern im Mathematikunterricht der Grundschule

**Margarita Knickenberg, Frank Hellmich**

Universität Paderborn, Deutschland; [margarita.ort@uni-paderborn.de](mailto:margarita.ort@uni-paderborn.de)

Kausalattributionen gelten als bedeutsame Determinanten der Leistungsentwicklung von Schülerinnen und Schülern, die als subjektive Ursachenerklärungen von Lernerfolgen und -misserfolgen verstanden werden. Attributionen sind entweder internal (Anstrengungen oder Fähigkeiten) oder external (zufällige ursächliche Ereignisse oder Aufgabenschwierigkeiten) verankert (Weiner, 1986). Wahrgenommene Ursachen von Erfolgen oder Misserfolgen können als stabil im Sinne von Talenten oder vorhandenen Fähigkeiten oder als variabel wie das Anstrengungsverhalten klassifiziert werden. Gegenwärtig ist noch weitgehend ungeklärt, ob und inwiefern selbst- und fremdbezogene Kausalattributionen von Kindern übereinstimmen oder sich voneinander unterscheiden. So liegen Befunde vor, die belegen, dass Kinder im Grundschulalter ihre eigenen Erfolge und Misserfolge auf ähnliche Ursachen zurückführen wie diejenigen ihrer Mitschülerinnen und -schüler (Folmer et al., 2008; Nicholls & Miller, 1984). In wiederum anderen Untersuchungen konnte nachgewiesen werden, dass selbst- und fremdbezogene Kausalattributionen bei Schülerinnen und Schülern im Sekundarstufenalter stark voneinander abweichen (Buff, 2004; Malle, 2006). Auf der Grundlage der vorliegenden Studien werden entwicklungsbedingte Unterschiede in Bezug auf Zusammenhänge



zwischen selbst- und fremdbezogenen Kausalattributionen angenommen (Folmer et al., 2008). Möglicherweise ist die Diskrepanz zwischen selbst- und fremdbezogenen Attributionen bei Schülerinnen und Schülern aber auch in ihrem schulischen Leistungsvermögen begründet. Unterschiede im Attributionsverhalten liegen zwischen leistungsstärkeren und -schwächeren Kindern vor (Jacobsen, Lowery & DuCette, 1986; Short, 1992).

Für den deutschsprachigen Raum sind keine Studien vorhanden, bei denen Zusammenhänge zwischen selbst- und fremdbezogenen Kausalattributionen bei Kindern im Grundschulalter betrachtet worden sind. Mitunter ist dies auf methodische Schwierigkeiten bei der Erfassung von Kausalattributionen bei sehr jungen Kindern zurückzuführen. Um Kinder zu ihren Kausalattributionen bei Lernerfolgen und -misserfolgen befragen zu können, ist das Visualisieren realer Lern- und Leistungssituationen, an denen sie emotional beteiligt sind, vonnöten (Köller & Möller, 1996).

Vor diesem Hintergrund wird in unserer Studie am Beispiel des Mathematikunterrichts der Grundschule der Frage nachgegangen, ob und inwiefern Unterschiede in den Kausalattributionen von eigenen versus fremden Lernerfolgen und -misserfolgen bei Kindern im Grundschulalter in Abhängigkeit ihrer selbsteingeschätzten Mathematikleistungen zu beobachten sind.

In unserer Studie wurden N=587 Grundschul Kinder der dritten und vierten Jahrgangsstufe mithilfe eines 'Paper and Pencil'-Fragebogens gebeten, Auskünfte über ihre Kausalattributionen im Unterrichtsfach Mathematik zu geben. Auf der Grundlage von Selbstauskünften der Kinder wurde die Gesamtstichprobe in zwei Gruppen unterteilt: Diejenigen Kinder (n=483), die ihre Mathematikleistungen als positiv bewerteten, bearbeiteten einen Fragebogenteil zu den Kausalattributionen ihrer Lernerfolge in Mathematik. Diejenigen Kinder (n=104), die ihre Leistungen als weniger positiv beurteilten, gaben Auskunft über ihre Kausalattributionen bei Lernmisserfolgen. Beide Gruppen wurden gebeten, Fragebogenaussagen zu ihren anstrengungs- und fähigkeitsbezogenen Kausalattributionen im Mathematikunterricht zu geben. Zur Untersuchung fremdbezogener Kausalattributionen nahmen die Kinder jeweils Stellung zu den Ursachen der Lernerfolge bzw. -misserfolge ihrer Mitschülerinnen und -schüler.

Die Befunde verdeutlichen, dass die leistungsschwächeren Kinder für die Erklärung eigener und fremder Lernmisserfolge signifikant häufiger mangelnde Fähigkeiten im Vergleich zu geringen Anstrengungen in Betracht ziehen. Demgegenüber führen leistungstärkere Kinder eigene und fremde Lernerfolge im Mathematikunterricht signifikant häufiger auf das Anstrengungsverhalten als auf mathematische Fähigkeiten zurück. Zur Überprüfung von Zusammenhängen zwischen selbst- und fremdbezogenen Kausalattributionen wurden Strukturgleichungsmodelle in Mplus für die beiden in unserer Studie betrachteten Leistungsgruppen unter der Kontrolle des Geschlechts und der Klassenstufe berechnet. Die selbstbezogenen Kausalattributionen auf Anstrengungen und Fähigkeiten wurden dabei als exogene Variablen, die fremdbezogenen Kausalattributionen als endogene Variablen modelliert.

Die Ergebnisse der jeweils für die Leistungsgruppen getrennt spezifizierten Strukturgleichungsmodelle verdeutlichen, dass fremdbezogene durch selbstbezogene Kausalattributionen von Erfolgen bzw. Misserfolgen erklärt werden können. Weiterhin zeigt sich, dass der Zusammenhang zwischen selbst- und fremdbezogenen Kausalattributionen bei den leistungsschwächeren weitaus stärker ausfällt als bei den leistungstärkeren Kindern.

Die Ergebnisse geben zusammenfassend Hinweise darauf, dass leistungsschwächere und -stärkere Kinder in unterschiedlicher Weise attribuieren und fremdbezogene Kausalattributionen bei Kindern im Grundschulalter durch selbstbezogene Ursachenzuschreibungen eigener Erfolge und Misserfolge erklärt werden können.

## **Bildungsaspirationen von Schulkindern im Grundschulalter**

**Nicole Gözl, Florian Wohlkinger**

LMU München, Deutschland; [nicole.goelz@edu.lmu.de](mailto:nicole.goelz@edu.lmu.de)

Theoretischer Hintergrund. Bildungsaspirationen stellen eine wichtige Größe für die Untersuchung von Mechanismen bildungsbezogener Disparitäten dar, ihr Einfluss auf Bildungsentscheidungen wurde bereits in zahlreichen Studien nachgewiesen. In der Literatur findet sich häufig die Unterscheidung zwischen realistischen und idealistischen Bildungsaspirationen, womit einerseits ideelle Bildungswünsche unabhängig von situativen Gegebenheiten sowie andererseits tatsächlich erreichbare Bildungsziele gemeint sind. Theoretisch wird davon ausgegangen, dass idealistische Aspirationen die realistischen übersteigen, da sie nicht von situationsabhängigen Bedingungen (wie etwa der Einschätzung der Erfolgswahrscheinlichkeit) beschränkt werden (vgl. Becker & Gresch, 2016).

Empirische Befunde zur Formation von Bildungsaspirationen liegen überwiegend für die Perspektive der Eltern (z.B. Kurz & Paulus, 2008) oder für die Perspektive von Jugendlichen im fortgeschrittenen Verlauf der Sekundarstufe (z.B. Trebbels, 2014) vor. Doch schon im Grundschulalter verfügen Kinder über Bildungsaspirationen, wie sowohl qualitative Studien (z.B. Kramer et al., 2009) als auch quantitative Untersuchungen (z.B. Wohlkinger, 2014) belegen. Darüber, wie sich solche Bildungsaspirationen bei Kindern in jüngeren Altersbereichen entwickeln, liegen bislang keine Erkenntnisse vor. Der Beitrag widmet sich dieser Forschungslücke, indem er die Bildungsziele von Schulkindern im Grundschulalter ins Zentrum der Untersuchung stellt.

Fragestellung. Das Ziel des Beitrages liegt darin herauszustellen, wie sich differenzielle Bildungsaspirationen in der frühen Schulkarriere entwickeln. Hierfür wird der Frage nachgegangen, inwieweit auch bei Kindern im Grundschulalter die theoretisch angenommenen Unterschiede zwischen idealistischen und realistischen Aspirationen empirisch vorgefunden werden können. Trennen auch Grundschul Kinder schon Wunsch und Wirklichkeit? Gibt es Gruppenunterschiede in Abhängigkeit von typischen ungleichheitsrelevanten Merkmalen (z.B. Alter, Geschlecht, Schulleistungen der Kinder; Migrationsstatus, soziale Lage, Ressourcenausstattung des Elternhauses)? Lassen sich weitere Einflussfaktoren für die idealistischen und realistischen Aspirationen der Grundschul Kinder identifizieren?

Methoden. Als Datengrundlage für die Untersuchung dient die Startkohorte 2 des Nationalen Bildungspanels (NEPS). Die fünfte Welle enthält Angaben von n = 5799 Schulkindern sowie deren Eltern. Die Kinder füllten in der 3. Klasse erstmals selbst einen Fragebogen aus, in dem neben Angaben zu Bereichen wie der schulischen Anstrengungsbereitschaft oder der häuslichen Computernutzung auch idealistische und realistische Bildungsaspirationen der Kinder erfasst wurden. Diese stellen die zentralen abhängigen Größen der Untersuchung dar.

Das methodische Vorgehen ist durch einen explorativen Zugang geprägt. Neben deskriptiven Kennwerten werden für ausgewählte Merkmale (s.o.) Mittelwertvergleiche und Varianzanalysen berechnet, um relevante Beziehungen zu Bildungsaspirationen zu bestimmen. Die dabei erszielten Ergebnisse stellen die Grundlage für vertiefende Regressionsanalysen dar, die Aufschluss darüber geben sollen, wie Bildungsaspirationen sich bei heranwachsenden Kindern entwickeln.

Ergebnisse. Erste Analysen zeigen, dass es Drittklässlern unterschiedlich gut gelingt, idealistische und realistische Bildungsaspirationen einzuschätzen: fast drei Viertel der Kinder (72,9%) geben eine Wunschschulform im Anschluss an die Grundschule an, aber weniger als die Hälfte (43,8%) haben eine Vorstellung davon, welche Schule es wohl tatsächlich wird. Unter den genannten Schulformen überwiegt das Gymnasium deutlich, und idealistische Aspirationen fallen erwartungsgemäß etwas größer aus als realistische. Desweiteren sind der Bildungsstatus der Eltern, die Schulleistungen, die Zahl der Bücher und andere kulturelle Ressourcen im Haushalt mit den Aspirationen der Kinder assoziiert. Die Befunde werden mit Blick auf die Entstehung von Bildungsungleichheiten im Zuge der familialen Sozialisation diskutiert.

## **Schulformspezifische Trajektorien im allgemeinen und schulischen Wohlbefinden von Schülerinnen und Schülern: Sekundarstufenschock oder Novitätseffekt nach dem Übergang zur Sekundarstufe?**

**Katharina Rathmann<sup>1</sup>, Max Herke<sup>2</sup>, Klaus Hurrelmann<sup>3</sup>, Matthias Richter<sup>2</sup>**

<sup>1</sup>Fachgebiet Rehabilitationssoziologie, Fakultät Rehabilitationswissenschaften, Technische Universität Dortmund, Deutschland;

<sup>2</sup>Institut für Medizinische Soziologie (IMS), Medizinische Fakultät, MLU Halle-Wittenberg, Deutschland; <sup>3</sup>Hertie School of Governance gGmbH, Berlin, Deutschland; [katharina.rathmann@tu-dortmund.de](mailto:katharina.rathmann@tu-dortmund.de)

### Theoretischer Hintergrund

Der schulischen Umwelt kommt in der Jugendphase eine besondere Bedeutung zu (Hascher & Hagenauer 2011a, 2011b). Gerade der frühe Übergang an die weiterführenden Schulen im deutschen Bildungssystem stellt ein kritisches Lebensereignis und einen Einschnitt im Verlauf der Kindheit und Jugend dar (Sirsch 2000, Van Ophuysen 2008), der mit Veränderungen der schulischen Lernumwelt durch veränderte Rahmen- und Lernbedingungen, neue Anforderungen und Interaktionsformen und ein verändertes soziales Umfeld einhergeht. Studien weisen auf einen Abfall des Wohlbefindens im Zusammenhang mit diesem kritischen Lebensereignis hin (Danielsen et al. 2009, Knoppick et al. 2015, Wigfield & Eccles 1994), dem sogenannten Sekundarstufenschock (Weißbach 1985). Andere Studien berichten keinen oder sogar einen positiven Effekt des Übergangs für das Wohlbefinden (Van Ophuysen 2008); der sogenannte Novitätseffekt (Van Ophuysen 2008). Damit ist ein grundlegend negativer Effekt des Übergangs auf die psychosoziale Entwicklung nicht haltbar, sondern es zeigen sich vielmehr differenzielle Verläufe. Diese sind auch zwischen unterschiedlichen Lernmilieus vorzufinden: Während für Hauptschülerinnen und -schüler Zuwächse im Wohlbefinden nach dem Übergang zur Sekundarstufe aufgezeigt wurden (Valentin & Wagner 2004, Knoppick et al. 2016), und diese damit als „selbstwertschützende Nischen“ fungieren (Aust et al. 2010, Baumert et al. 2006), zeigte sich für die Schülerschaft an Gymnasien ein gegenteiliger Effekt (Knoppick et al. 2015). Bisher mangelt es allerdings an Studien, die Trajektorien im Wohlbefinden von Heranwachsenden während der Schullaufbahn nicht nur zu zwei Messzeitpunkten (vor und nach dem Übergang) untersuchen, sondern das Wohlbefinden während der gesamten Sekundarstufenzeit mit mehreren Messzeitpunkten und differenziert nach der besuchten Schulform betrachten, um mögliche Novitäts- oder Sekundarstufenschockeffekte nachzeichnen zu können.

### Fragestellung

Der Beitrag untersucht Verläufe im Wohlbefinden während der Sekundarstufe 1 und prüft eine differenzielle Entwicklung ausgewählter Facetten des Wohlbefindens nach dem Übergang zu Schulformen des gegliederten Schulsystems. Folgende Fragestellungen stehen im Fokus:

- 1) Wie entwickelt das Wohlbefinden der Heranwachsenden während der Sekundarstufe?
- 2) Lassen sich Novitäts- oder Sekundarstufenschockeffekte für das Wohlbefinden der Schülerschaft aufzeigen?
- 3) Gestalten sich die Trajektorien im Wohlbefinden differenziell zwischen schulischen Lernmilieus?

### Methode

Die Datenbasis bildet das Nationale Bildungspanel (NEPS). In die Analysen flossen circa 29.200 Beobachtungen über fünf Beobachtungszeitpunkte von rund 8.130 Schülerinnen und Schüler von der 5. bis 9. Klasse (2010-2014) ein (Blossfeld et al. 2011). Unter Kontrolle von soziodemografischen und –ökonomischen Merkmalen wurden lineare, quadratische und kubische Wachstumskurvenmodelle für das allgemeine Wohlbefinden (Lebenszufriedenheit) und schulische Wohlbefinden (Schulzufriedenheit) berechnet. Zudem wurden differenzielle Verläufe für die beiden Facetten des Wohlbefindens, differenziert nach der besuchten Schulform (Gymnasium versus Real-, Gesamt- und Hauptschulen), analysiert.

### Ergebnisse

Im Verlauf der Sekundarstufe 1 verschlechtern sich das allgemeine und schulische Wohlbefinden aller Schülerinnen und Schüler von Klasse 5 bis 9. Insbesondere nach dem Übergang zur Sekundarstufe in verschlechtert sich das Wohlbefinden der Schülerschaft und weist auf den Sekundarstufeneffekt in Klasse 6 hin. Bis Klassenstufe 8 konsolidieren sich beide Wohlbefindensmerkmale für Heranwachsende, die ein Gymnasium besuchen, und weisen sogar einen leichten Anstieg für Schülerinnen und Schüler anderer Schulformen auf, der auf einen verspäteten Novitäts- bzw. Gewöhnungseffekt hindeutet. Die schulformspezifischen Verläufe in beiden Facetten des Wohlbefindens zeigen, dass Schülerinnen und Schüler in der 5. Klasse, die nicht das Gymnasium besuchen, signifikant niedrigere Mittelwerte in beiden Wohlbefindensfacetten im Vergleich zu Gymnasiastinnen und Gymnasiasten aufweisen. Während der gesamten Sekundarstufenzeit zeigen Heranwachsende, die eine andere Schulform als das Gymnasium besuchen, ein niedrigeres Niveau im Wohlbefinden, das sich zum Ende der Sekundarstufe 1 allerdings zwischen beiden Schulformen angleicht. Für alle Schülerinnen und Schüler wird von Klasse 8 bis 9 zudem ein erneuter Rückgang im Wohlbefinden deutlich, der auf einen „zweiten Schock“ bzw. „Coming-Of-Age“-Effekt hindeutet. Insgesamt gibt der Beitrag erste Hinweise darauf, dass insbesondere der normative Übergang und das Ende der Sekundarstufe eine kritische Lebensphase für Heranwachsende darstellt.

### Quality-aspects of all-day schools in Switzerland: What affects academic and socio-emotional skills?

**Marianne Schüpbach<sup>1</sup>, Lukas Frei<sup>1</sup>, Benjamin von Allmen<sup>1</sup>, Wim Nieuwenboom<sup>2</sup>**

<sup>1</sup>Universität Bamberg, Deutschland; <sup>2</sup>HSA FHNW; [marianne.schuepbach@uni-bamberg.de](mailto:marianne.schuepbach@uni-bamberg.de)

In the past decade, societal changes and political efforts have led to an increase in all-day schools across Switzerland (Stern et al., 2013), which are defined as schools not only comprising regular school hours but also offering education and care during the rest of the day (EDK, 2013). Those services (referred to as extended education) generally include lunch, a supervised program in the afternoon, and, although less frequently, before-school care. Most all-day schools implemented an open-attendance model (voluntary all-day schools).

In current debates on education in Switzerland, all-day schools are granted a central role in view of school achievement and socio-emotional development (Coradi Vellacott et al., 2003). Initial research supported a positive effect of an intensive utilization of extended education on the development of mathematical and reading skills (Schüpbach, 2014; Schüpbach et al., 2013) but not on socio-emotional behavior (Schüpbach et al., 2014). However, our follow-up study on voluntary all-day schools failed to confirm those promising results on school achievement and even showed detrimental effects of extended education utilization on externalizing behaviour (Frei et al., 2016).

While the quality of extended education has long been proposed to moderate the impact of all-day schools on educational and socio-emotional outcomes (e.g. Stecher, Radisch et al., 2007), it is unclear which aspects of quality are most relevant with regard to current all-day schools in Switzerland. Therefore, the aim of this contribution is (1) to identify distinctive components of quality in extended education and (2) to assess the impact of those components with regard to educational and socio-emotional outcomes.

Data are drawn from the longitudinal research project EduCare-TaSe, which is funded by the Swiss National Science Foundation. EduCare-TaSe is studying 1'990 students from 53 voluntary all-day schools in the German-speaking part of Switzerland, with some students utilizing extended education, whereas others do not. Educational skills and socio-emotional behavior are measured at the end of first grade (T1) and again at the end of second grade (T2): Mathematical skills are assessed with the HRT (Haffner et al., 2005), reading skills with the ELFE (Lenhard & Schneider, 2006), and externalizing/internalizing behavior is evaluated by class teachers, using the Strengths and Difficulties Questionnaire (Goodman, 1997). During first grade, the quality of the extended education services is rated through observation, using the Hort- und Ganztagskala (HUGS; Tietze et al., 2005). Since the originally proposed 6-factor-structure of the HUGS did not prove valid in our sample, we conducted exploratory factor analyses to identify distinctive quality-components. Subsequently, we calculated multilevel-regression-models to assess the impact of those components on externalizing behavior, internalizing behavior, mathematical skills and reading skills.

Our analyses of the HUGS-items identified three distinctive quality-components with four items each: (1) Space and Equipment, (2) Caregiver-Student Interactions, and (3) Collaboration. Subsequent multilevel-regression-analyses identified caregiver-student interactions and collaboration as important aspects of quality with regard to academic and socio-emotional outcomes: While caregiver-student interactions moderated the effect of extended education on externalizing behaviour, the development of mathematical skills in students, regardless of their utilization of extended education, was positively influenced by the level of collaboration at the extended education setting.

Our analyses suggest three distinctive quality-components with regard to all-day schools in Switzerland: While Space and Equipment and Caregiver-Student Interactions correspond with quality-aspects as proposed by Tietze et al. (2005), collaboration emerged as a new aspect of quality. On the other hand, several other original subscales did not surface in our analyses, pointing to confounding characteristics within several quality-aspects or to limitations regarding our observational instrument. Interestingly, different quality-components predicted different aspects of student-development, highlighting their importance with respect to academic and socio-emotional development, as well as the normative nature of all-day school quality.

### Responsivität der Eltern und Berufswahlprozesse von Jugendlichen mit tiefen und hohen schulischen Leistungen

**Markus P. Neuenschwander<sup>1</sup>, Stephan Schumann<sup>2</sup>, Andreas Jüttler<sup>2</sup>**

<sup>1</sup>Pädagogische Hochschule Nordwestschweiz, Schweiz; <sup>2</sup>Universität Konstanz; [markus.neuenschwander@fhnw.ch](mailto:markus.neuenschwander@fhnw.ch)

Theoretischer Hintergrund

Am Ende der obligatorischen Schule müssen Jugendliche eine erste Laufbahntscheidung treffen. Jugendliche wählen die berufliche Laufbahn oft so, dass sie möglichst gut zu ihrer Person passt (Holland, 1997) und dass sie in ihrer Entscheidung sicher sind (Hirschi, 2011). Im Jugendalter haben Schüler/innen mit besseren schulischen Leistungen mehr Ausbildungsmöglichkeiten als solche mit schlechteren schulischen Leistungen und haben daher eher die Möglichkeit, eine passende Ausbildung zu finden (Hupka-Brunner et al., 2010). Dennoch korrelieren die schulischen Leistungen nicht mit der wahrgenommenen Passung der beruflichen Ausbildung und der Entscheidungssicherheit der Berufswahl (Rogers & Creed, 2011). Zur Erklärung der Berufswahlprozesse nehmen wir im Anschluss an Lent et al. (1994) an, dass die Responsivität von Eltern die Selbstwirksamkeitsüberzeugung und die berufliche Exploration beeinflusst, welche wiederum die Entscheidungssicherheit der Berufswahl und die wahrgenommene Passung zwischen Person und Beruf vorhersagen.

Fragestellung

Auf dieser theoretischen Grundlage gehen wir den Fragen nach, ob die schulischen Leistungen den Effekt der elterlichen Responsivität auf die Selbstwirksamkeit und die berufliche Exploration moderieren. Und weiter: Moderieren die schulischen Leistungen den Effekt der Exploration und der Selbstwirksamkeit auf die Entscheidungssicherheit und die wahrgenommene Passung zwischen der Person und der Laufbahntscheidung?

Methode

Diese Fragen wurden mit Längsschnittdaten aus dem Deutschschweizer Projekt „Wirkungen der Selektion (WiSel)“ überprüft. Im fünften Schuljahr umfasste die Stichprobe 897 weibliche und 938 männliche Schüler/innen (Durchschnittsalter 11.8 Jahre) aus 126 Klassen, die zufällig aus Schulen in fünf Schweizer Kantonen gewählt wurden. Von diesen Jugendlichen füllten 629

Jugendliche den standardisierten Fragebogen auch im 7. Schuljahr aus. Im 7. Schuljahr wurde die Stichprobe mit den Schüler/innen ergänzt, die in den gleichen Klassen wie die Kinder in der Längsschnittstichprobe unterrichtet wurden (7. Klasse: w: N=716, m: N=768). Von dieser Stichprobe wurden 667 Schüler/innen im 9. Schuljahr wieder befragt. Ihre neuen Mitschüler/innen wurden zur Teilnahme an der Studie angefragt (Stichprobe im 9. Schuljahr: w: N=716, m: N=768). Wir fanden keine systematischen Rücklaufverzerrungen zwischen den Wellen für die Variablen, die in dieser Studie verwendet wurden.

Die Jugendlichen füllten reliable standardisierte Fragebogen und Leistungstests in Deutsch und Mathematik aus. Für die Analysen wurde die Stichprobe aufgrund der Leistungen in Deutsch und Mathematik im 7. Schuljahr mit Medianhalbierung in zwei Gruppen geteilt. Die Fragen wurden unter Beizug von Strukturgleichungsmodellen mit Gruppenvergleich bearbeitet. Fehlende Werte wurden mit der FIML Prozedur in Mplus behandelt. Invarianzanalysen der Messmodelle zeigten vollständige oder partielle metrische Invarianz (Schoot et al. 2012).

#### Ergebnisse

Die Responsivität der Eltern im 5. Schuljahr sagte die berufliche Exploration und die Selbstwirksamkeitserwartung im 7. Schuljahr vorher. Die berufliche Exploration im 7. Schuljahr sagte wiederum die berufliche Exploration im 9. Schuljahr vorher, die Selbstwirksamkeitserwartung im 7. Schuljahr die Selbstwirksamkeitserwartung im 9. Schuljahr. Berufliche Exploration und Selbstwirksamkeitserwartung im 9. Schuljahr sagte die Entscheidungssicherheit vorher. Vier Pfade unterschieden sich zwischen den Gruppen signifikant.

Danach wurde das analoge Strukturgleichungsmodell mit Gruppenvergleich überprüft, wobei die Entscheidungssicherheit mit der Passungswahrnehmung ersetzt wurde. Drei Pfade unterschieden sich signifikant zwischen den Gruppen.

Die Ergebnisse bestätigen Aspekte des erweiterten und angepassten Modells von Lent et al. (1994) und belegen die moderierende Wirkung der Leistungen auf den Effekt der elterlichen Responsivität auf die berufliche Exploration und die Selbstwirksamkeit. Jugendliche mit geringen Schulleistungen explorieren vor Laufbahnentscheidungen intensiver, wenn ihre Eltern auf ihre Bedürfnisse responsiv reagieren (Dietrich & Kracke, 2009). Die Ergebnisse zeigen überdies die moderierende Wirkung der Schulleistungen auf den Effekt der Selbstwirksamkeit und der beruflichen Exploration auf die Entscheidungssicherheit. Überdies wird deutlich, dass Jugendliche mit geringeren Leistungen bei hoher beruflicher Exploration eher eine passende Laufbahn wählen. D.h., eine intensive berufliche Exploration kann Leistungsnachteile bei Laufbahnentscheidungen kompensieren. Die Studie trägt zur bisher vernachlässigten Forschung über die moderierende Wirkung von Schulleistungen bei Laufbahnentscheidungen bei.

## Prädiktoren der elterlichen Zeitverwendung für Kinderbetreuung

**Ramona Schlupf, Michaela Gläser-Zikuda**

Universität Erlangen-Nürnberg, Deutschland; [ramona.schlupf@fau.de](mailto:ramona.schlupf@fau.de)

Kinder werden zunehmend als „Investitionsgut“ angesehen, in welches zeitliche, finanzielle und soziale Ressourcen investiert werden sollen (Berghammer, 2013; Gimenez-Nadal & Molina, 2012; Wild & Lorenz, 2010). Die Dauer, welche Eltern in die Kinderbetreuung investieren, hat sich in den vergangenen Jahrzehnten kontinuierlich erhöht (vgl. z.B. Berghammer, 2013). Bei bewusster Entscheidung für die Familiengründung stellen Kinder eine wichtige Ressource für die elterliche Zufriedenheit und deren Lebensglück dar. Eine optimale Nutzung der häufig nur knapp zur Verfügung stehenden zeitlichen Ressourcen ist für Erziehung und Bildung des/r Kindes/r bedeutsam. Neben der allgemeinen Zunahme der für Kinderbetreuung (Boll et al., 2013) investierten Zeit, ist speziell ein Anstieg der Betreuungszeit durch die Väter beobachtbar (Flaake, 2011). Ob und wieviel an zeitlichen Ressourcen für die Kinderbetreuung und –erziehung zur Verfügung stehen, ist allerdings von einer Vielzahl an Einflussgrößen abhängig, welche in der vorzustellenden Untersuchung analysiert wurden.

#### Forschungsfragen

Die Analysen basierend auf Daten des Statistischen Bundesamtes 2012/13 geht folgenden Forschungsfragen nach: Welche prädiktive Wirkung auf die Zeitverwendung der Eltern für ihre Kinder kommen Geschlecht, Bildungsstand und Erwerbstätigkeit der Eltern zu? Wie verhält sich der Einfluss dieser Prädiktoren unter Berücksichtigung eines Interaktionseffektes mit dem Geschlecht?

#### Methode

Insgesamt konnten Daten von N Eltern = 3204 Elternteilen genutzt werden, von denen 43.7% (N männlich = 1401) männlich und 56.3% (N weiblich = 1803) weiblich sind. Das durchschnittliche Alter liegt bei  $M = 43.24$  ( $SD = 7.19$ ). 47.6% der Befragten geht einer Vollzeitbeschäftigung nach, während 28.8% teilzeitbeschäftigt und 23.6% nicht erwerbstätig sind. Der Großteil der befragten Elternteile (49.8%) verfügt über eine mittlere Bildung, während 46.6% über eine hohe Bildung und 3.6% über keine bzw. niedrige Bildung verfügen. Die Dauer, in der die Kinderbetreuung als Haupttätigkeit ausgeführt wurde, betrug täglich knapp eine Stunde und sieben Minuten ( $M = 67.15$ ;  $SD = 89.45$ ) pro Person.

Die abhängige Variable der Zeitverwendung für die Kinderbetreuung als Hauptaktivität wurde mittels eines Tagebuchs erfasst, während Geschlecht, Erwerbstätigkeit sowie das Alter und die Anzahl der Kinder direkt in den Personen-/Haushaltsfragebögen erfasst wurden. Der Bildungsstand wurde mittels der Angaben zum Schul- und Berufsabschluss der Befragten durch das Statistische Bundesamt entsprechend der verschiedenen ISCED-97-Levels kodiert und zu den Variablen keine bzw. niedrige Bildung, mittlere Bildung und hohe Bildung zusammengefasst. Aufgrund des nominalen bzw. ordinalen Skalenniveaus erfolgte eine Dichotomisierung.

#### Ergebnisse

Ergebnisse der linearen Regression zeigen einen starken Effekt des Einflussfaktors Geschlechts: Mütter wendeten täglich durchschnittlich  $M = 24.79$  ( $B = 1497.25$ ;  $S.E. = .19$ ;  $p < .001$ ) Minuten mehr als Väter auf. Die Erwerbstätigkeit in Vollzeit stellt sich als stärkster negativer Prädiktor heraus ( $B = -2049.32$ ;  $S.E. = .37$ ;  $p < .001$ ). Auch die Berufstätigkeit in Teilzeit hat einen relativ starken negativen Einfluss ( $B = -1184.28$ ;  $S.E. = .52$ ;  $p < .001$ ). Die Erwerbstätigkeit wirkt sich über das Geschlecht vermittelt stark auf die Zeitverwendung zur Kinderbetreuung aus: Mütter in Vollzeit- oder Teilzeitbeschäftigung wenden wesentlich weniger Zeit auf als die Elternteile der Referenzgruppen ( $B$  InteraktionVZ =  $-1257.29$ ;  $S.E. = .46$ ;  $p < .001$ ;  $B$  InteraktionTZ =  $-1628.05$ ;  $S.E. = .56$ ;  $p < .001$ ). Obwohl eine niedrige Bildung kaum Einfluss auf die Zeitverwendung der Eltern hat ( $B = -135.23$ ;  $S.E. = .55$ ;  $p < .001$ ), besteht unter Berücksichtigung des Geschlechts jedoch ein deutlicher Effekt ( $B = -696.39$ ;  $S.E. = .70$ ;  $p < .001$ ).

Es kann belegt werden, dass (1) Tätigkeiten der elementaren Kinderbetreuung den größten Zeitfaktor einnehmen, (2) sowohl das Geschlecht, als auch die Erwerbstätigkeit auf Seiten der Eltern und die Anzahl sowie das Alter der Kinder starke

Prädiktoren darstellen, und sich (3) vor allem die Erwerbstätigkeit und der Bildungshintergrund über eine Interaktion mit dem Geschlecht auf die Zeitverwendung zur Kinderbetreuung auswirken. Die Bedeutung der Ergebnisse werden mit Blick auf den Forschungsstand abschließend diskutiert.

## **Der Einfluss mehrfacher Betreuungswechsel auf die sozial-emotionalen Fähigkeiten von Kindern im Alter von 5 Jahren**

**Franziska Cohen, Yvonne Anders, Eric Vogel**

Freie Universität Berlin, Deutschland; [franziska.cohen@fu-berlin.de](mailto:franziska.cohen@fu-berlin.de)

Nationale sowie internationale Studienergebnisse belegen die positiven Effekte frühkindlicher Betreuung auf die kognitive als auch sozial-emotionale Entwicklung von Kindern. Insbesondere von einer qualitativ guten Betreuung profitieren gerade Kinder aus Familien mit geringerer Anregungsqualität (Camilli et al., 2010; Nores & Barnett, 2010; Melhuish 2004; Ulferts & Anders, 2015). Die Auswirkungen frühkindlicher Betreuung auf die kognitiven bzw. sprachlichen Outcomes von Kindern sind zahlreich belegt, aber nur wenige deutschsprachige Studien beschäftigen sich mit den Auswirkungen auf die sozial-emotionale Entwicklung von Kindern. Sozial-emotionale Fähigkeiten, und im Speziellen die Selbstregulation des Kindes, gelten auch langfristig als ein wichtiger Indikator für den späteren Schulerfolg des Kindes.

Es werden jedoch immer wieder Bedenken bezüglich negativer Auswirkungen der frühkindlichen außerfamilialen Betreuung auf die sozial-emotionale Entwicklung von Kindern geäußert. Neben einem intensiven Betreuungsumfang und einem frühen Eintrittsalter in die Kindertagesstätte, haben verschiedene internationale Studien gezeigt, dass instabile Betreuungsmuster, in Form von mehrfachen Betreuungswechseln, und damit mehreren verschiedenen Betreuungspersonen im Alltag des Kindes, negative Auswirkungen auf das sozial-emotionale Wohlbefinden des Kindes haben können (Anders, 2013; Claessens, 2012; DeSchipper, Tavecchio, van Ijzendoorn & van Zeijl, 2004; Morrissey, 2009, Claessens & Chen, 2013). Neuere und nationale Studien finden diese Zusammenhänge jedoch nicht. (z.B. Tietze et al., 2013). Betreuungswechsel können dabei zwischen formalen (KiTa, Tagespflege) als auch informellen Betreuungssettings (z.B. Großeltern, Babysitter) stattfinden.

Vor diesem Hintergrund wird der Frage nachgegangen, inwiefern die Anzahl an Betreuungswechseln in den ersten drei Lebensjahren Auswirkungen auf die sozial-emotionale Entwicklung und die selbstregulatorischen Fähigkeiten von Kindern im Alter von 5 Jahren hat.

Die Datenbasis bilden 162 Familien, die an einer längsschnittlich angelegten Evaluationsstudie mit zwei Messzeitpunkten teilnahmen. Die Kinder waren zum ersten Messzeitpunkt durchschnittlich drei Jahre alt und zum zweiten Messzeitpunkt fünf Jahre alt. Die Eltern wurden durch geschulte Datenerheber/-innen mittels Fragebögen zur Anzahl an Betreuungspersonen und dem Betreuungsumfang in informellen und formalen Settings befragt. Ebenso wurden Informationen zum sozialen Hintergrund und der familialen Anregungsqualität erhoben. Die sozial-emotionalen Fähigkeiten des Kindes wurden mithilfe des Strength-Difficulties-Questionnaire (SDQ: Goodman, 1997) erfasst. Zusätzlich wurden die pädagogischen Fachkräfte in den Kindertagesstätten gebeten, die Fähigkeiten der Kinder auf der SDQ-Skala einzuschätzen. Die selbstregulatorischen Fähigkeiten wurden mithilfe des Impuls-Management-Fragebogens (IMMA) (Pauen et al. 2014) durch die Eltern eingeschätzt.

Erste deskriptive Ergebnisse zeigen, dass das durchschnittliche Eintrittsalter in die Kindertagesstätte bei zwei Jahren liegt. In den ersten drei Lebensjahren der Kinder nehmen Eltern sowohl informelle, als auch formale Betreuungsmöglichkeiten in Anspruch. Gleichzeitig zeigt sich eine große Varianz im Betreuungsumfang sowie in der Anzahl der Betreuungspersonen. Ein höherer Bildungshintergrund der Mutter geht mit einem frühen Eintritt in die institutionelle formale Betreuung einher. Familien, die im Alltag mit ihrem Kind nicht deutsch sprechen, nehmen wiederum seltener formale Betreuung in den ersten drei Lebensjahren wahr. Bivariate Korrelationsanalysen zeigen ebenfalls, dass mit einer höheren Anzahl an Betreuungswechseln zwischen formalen Betreuungssettings im Alter bis zu drei Jahren, die pädagogischen Fachkräfte in den Kindertageseinrichtungen das Problemverhalten der Kinder im Alter von 5 Jahren geringer einschätzen und gleichzeitig das prosoziale Verhalten als stärker ausgeprägt berichten. Ebenfalls zeigt sich, dass mit höherer Anzahl an Betreuungspersonen in den ersten drei Lebensjahren die selbstregulatorischen Fähigkeiten (Zielorientierung) des Kindes im Alter von fünf Jahren von den Eltern höher eingeschätzt werden. Detaillierte Einblicke in die Zusammenhänge zwischen Betreuungswechseln und den sozial-emotionalen Fähigkeiten von Kindern werden mithilfe multivariater Regressionsanalysen ermöglicht.

## G09

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Bernoullianum 32, H223

### Soziale Stereotype und Peer Review: Der Einfluss von Ethnizität und institutioneller Affiliation

Charlotte Hufeisen, Florian Klapproth

Medical School Berlin, Deutschland; [florian.klapproth@medicalschooll-berlin.de](mailto:florian.klapproth@medicalschooll-berlin.de)

#### Theoretischer Hintergrund

Die Beurteilung wissenschaftlicher Arbeiten im Rahmen des Peer Review gehört zu den Aufgaben des wissenschaftlichen Personals an Hochschulen und Universitäten. Üblicherweise werden diese Arbeiten anonymisiert zur Begutachtung eingereicht. Jedoch gibt es immer wieder Ausnahmen, so dass zum Beispiel dem Herausgeber einer Fachzeitschrift oder - bei der Einreichung von Beiträgen für ein Symposium auf einem Kongress - dem jeweiligen Gutachter der Name und die Affiliation der Autoren bekannt sind (Fröhlich, 2002). Dabei kann der Peer-Review-Prozess beeinflusst werden von Informationen über den Beurteilten (Quinlan, 2002). Als Ursache dieses Effektes wird angenommen, dass Name und Herkunft einer Person stereotype Einstellungen über diese Person aktivieren und in Folge die Bewertung ihrer Leistung beeinflussen können (Bormann & Daniel, 2004; Klapproth, Kärchner & Glock, 2017; Wheeler & Petty, 2001).

#### Fragestellung

Mit einem Experiment wurde die Hypothese geprüft, dass die Beurteilung eines Beitrags für eine wissenschaftliche Tagung abhängig ist von (a) der Ethnizität des Beitragautors sowie (b) von dem Renommee seiner institutionellen Affiliation.

#### Methode

Grundlage der Studie war eine deutschsprachige Zusammenfassung einer psychologischen Studie (319 Wörter), die von insgesamt N = 105 Studierenden (71.4 % weiblich) des Bachelorstudiengangs Psychologie im Hinblick darauf beurteilt wurde, ob sie als Vortrag auf einer wissenschaftlichen Tagung angenommen werden kann. Die Zusammenfassung der Studie wurde ergänzt durch den Namen des Autors und seine Affiliation. Die Versuchspersonen wurden zufällig vier Bedingungen zugewiesen. Unabhängige Variablen waren (a) die Ethnizität des Autors (operationalisiert durch einen türkischen oder englischen männlichen Namen) und (b) das Renommee seiner institutionellen Affiliation (gering vs. hoch). Abhängige Variablen waren (a) die Empfehlung der Versuchspersonen (für vs. gegen die Veröffentlichung auf einer wissenschaftlichen Tagung) und (b) die Bewertung des Beitrags auf einer 5-stufigen Likert-Skala.

#### Ergebnisse und Diskussion

Der Beitrag wurde signifikant positiver bewertet, wenn der Autor einen englischen statt eines türkischen Namens hatte,  $F(1, 101) = 122.74$ ,  $p < .001$ ,  $\eta^2 = .549$ , und seine Affiliation hohes statt geringes Renommee hatte,  $F(1, 101) = 43.10$ ,  $p < .001$ ,  $\eta^2 = .299$ . Die Interaktion war nicht signifikant,  $F(1, 101) = 0.06$ ,  $p = .807$ ,  $\eta^2 = .001$ . Die Ergebnisse einer logistischen Regressionsanalyse zeigten ferner, dass die Wahrscheinlichkeit für die Annahme des Beitrags signifikant reduziert war, wenn der Autor einen türkischen Namen hatte,  $B = -5.29$ ,  $p < .001$ , Odds Ratio = 0.005. Allerdings war weder der Haupteffekt des Renommee der Affiliation noch der Interaktionseffekt signifikant. Die Ergebnisse dieser Studie legen nahe, dass soziale Stereotype zumindest bei Studierenden der Psychologie einen großen Einfluss auf die Beurteilung wissenschaftlicher Arbeiten ausüben. Das Ausbleiben des Interaktionseffektes deutet ferner darauf hin, dass der Effekt ethnischer Stereotype auf die Urteilsbildung nicht nivelliert wird durch die Affiliation des Autors.

### Politische Partizipation und Wissen

Natalie Grobshäuser, Georg Weißeno

PH Karlsruhe, Deutschland; [koehle@ph-karlsruhe.de](mailto:koehle@ph-karlsruhe.de)

Der Politikunterricht soll auf die Bürgerrolle vorbereiten, die sowohl Wissen als auch politische Partizipation umfasst. Im Politikunterricht finden Informationsverarbeitungsprozesse von politischen Inhalten statt. Die nach dem Modell von Detjen et al. (2012) zu entwickelnde und operationalisierbare Politikkompetenz der Schüler/-innen besteht aus drei Dimensionen: Fachwissen, politischer Urteilsfähigkeit, politischer Handlungsfähigkeit. 30 Fachkonzepte mit ihren konstituierenden Begriffen entfalten das im Unterricht zu erwerbende Fachwissen. Mit der politischen Urteils- und Handlungsfähigkeit wird die Argumentationsfähigkeit, nicht das tatsächliche Handeln im Sinne von aktiver politischer Partizipation beschrieben (Manzel & Weißeno, 2017). Die Partizipation wird in der vorliegenden Studie als reales und als beabsichtigtes politisches Handeln operationalisiert. Als theoretischer Rahmen für die Partizipationserfahrung und -bereitschaft dient die Theorie der geplanten Handlung (Ajzen, 1991). Ausgangspunkt ist die Annahme, dass Menschen subjektiv rational sind. Sie nutzen aber die ihnen zur Verfügung stehenden Informationen nicht immer korrekt oder vollständig. Sie umfassen aufgrund dessen diesbezügliche individuelle Überzeugungen. Eine Handlung wird ausgeübt in Bezug auf ein Ziel, in einem bestimmten Kontext und zu einem bestimmten Zeitpunkt. Die Theorie geht davon aus, dass sich Partizipationserfahrung auf Partizipationsbereitschaft auswirkt. Nach Gabriel (2013) wird politische Partizipation durch bestimmte Determinanten der kognitiven Involvierung wie Interesse, Selbstkonzept sowie Wissen beeinflusst. Das Modell sieht die Wirkrichtung von Wissen auf Partizipation. Die Wirkrichtungsfrage gilt aber als nicht entschieden (Delli Carpini & Keeter, 1996; Solhaug, 2006). Oft werden in den Studien nur Korrelationen berichtet. Offen ist zudem, ob ein Beitrag der Partizipation zur Aufklärung des Fachwissens angenommen werden kann. Es gibt bisher kaum empirische Belege zu dieser Frage im schulischen Kontext (Biedermann, 2006; Landwehr, 2017).

Hier wird angenommen, dass sich 1. vorangegangene Partizipationserfahrung positiv auf schulisches politisches Wissen sowie Partizipationsbereitschaft und 2. das Wissen ebenfalls auf die Partizipationsbereitschaft auswirken.

Es wurden 1324 Schüler/-innen der neunten Klassenstufe an Realschulen in Baden-Württemberg befragt. In einem standardisierten Fragebogen wurde zum einen das schulische politische Wissen basierend auf den Fachkonzepten für den Politikunterricht erhoben. Zum anderen wurden erprobte Likert-Skalen zur politischen Partizipation aus der ICCS-Studie (2011) in deutscher Version übernommen. Das schulische politische Wissen wurde mit dem dichotomen Raschmodell skaliert. Als individuelle Personalfähigkeit wurden WLEs (Weighted Likelihood Estimate) zur weiteren Analyse verwendet. Die Skalen zur politischen Partizipation wurden konfirmatorischen Faktorenanalysen unterzogen. Anschließend wurden die Zusammenhänge und Effekte zwischen politischer Partizipation und dem schulischen politischen Wissen latent modelliert. Hierbei wurde die hierarchische Struktur der Daten berücksichtigt und auf Hintergrundvariablen wie Geschlecht kontrolliert. Ein KFT wurde als unabhängiger Indikator für Fachwissen zusätzlich durchgeführt.

Die Ergebnisse zeigen den erwarteten positiven Effekt des schulischen politischen Wissens auf die Partizipationsbereitschaft. Wer mehr über Politik weiß, äußert eher die Bereitschaft zu partizipieren. Die Partizipationserfahrung hat einen positiven Einfluss auf die geäußerte Partizipationsbereitschaft, jedoch entgegen der Erwartung nicht auf das Wissen. Wer angibt, bereits politisch aktiv gewesen zu sein, will auch zukünftig partizipieren, verfügt allerdings nicht über mehr Wissen. Die Note in Gemeinschaftskunde wirkt sich günstig für die Partizipationsbereitschaft aus. Mädchen verfügen über etwas weniger schulisches politisches Wissen und äußern auch weniger Bereitschaft, sich politisch zu beteiligen. Der Effekt des Geschlechts auf die Partizipationsbereitschaft ist allerdings nur schwach signifikant. Schüler/-innen mit Migrationshintergrund sind unter Kontrolle des Wissens weniger bereit, sich politisch zu beteiligen und wissen weniger über Politik als Schüler/-innen ohne Migrationshintergrund. Überraschend ist, dass es keine Rolle für den Wissensstand spielt, ob zu Hause nur Deutsch gesprochen wird. Auf die Partizipationsbereitschaft wirkt sich Deutsch als zu Hause gesprochene Sprache hingegen positiv aus.

## **Zum Einfluss von Lernereigenschaften auf das Verständnis der Struktur der Materie**

**Andreas Jackowski, Stefan Rumann**

Universität Duisburg-Essen, Deutschland; [andreas.jackowski@uni-due.de](mailto:andreas.jackowski@uni-due.de)

Neben der fachdidaktischen und der allgemeinpädagogischen Ausbildung ist die Ausbildung in den jeweiligen Fachwissenschaften zentraler Bestandteil des Lehramtsstudiums. In der Chemie werden verschiedene Teildisziplinen wie die organische oder die physikalische Chemie studiert. Dabei ist allen Fachbereichen gemein, dass sich die jeweiligen Inhalte zu den Basiskonzepten der Chemie zuordnen lassen. Eines dieser Basiskonzepte ist die Struktur der Materie, deren Fachinhalte sich mit der Entstehung, der Zusammensetzung und den Charakteristiken der Materie beschäftigen. Die Entwicklung eines konzeptuellen Verständnisses der Struktur der Materie ist daher eine der zentralen Anforderungen der Studieneingangsphase in der Chemie.

Die Vermittlung von Informationen ohne Einsatz von Abbildungen ist praktisch nicht realisierbar (Gilbert, 2007). Dies gilt, aufgrund des Aufbaus der beinhalteten Fachinhalte, insbesondere für das Konzept der Struktur der Materie. Da Fachinhalte in diesem Bereich durch eine Vielzahl unterschiedlicher Repräsentationsformen vermittelt werden, muss die spezifische Bildsprache der jeweiligen Repräsentationstypen den Lernenden bekannt sein. Sonst kann es zu Fehlern und Fehlinterpretationen bei der Informationsaufnahme kommen (Weidenmann, 1994). Die verschiedenen Arten der Repräsentationen können hierbei nach ihrem Abstraktionsgrad, in symbolisch und ikonisch, unterschieden werden (Schnotz, 2005).

Ikonische Abbildungen teilen strukturelle Gemeinsamkeiten mit realen Bezugsobjekten und sind depiktive Abbildungen. Symbolische Abbildungen hingegen haben keine räumliche Ähnlichkeit mit dem Realobjekt und präsentieren Informationen in mathematisierter, deskriptiver Form (Schnotz, 2005). Bei der Arbeit mit unterschiedlichen Repräsentationstypen haben dabei individuelle Lernereigenschaften einen direkten Einfluss auf die Lernwirksamkeit

(Höfler, Schmeck & Opfermann, 2013).

Als Forschungsfrage ergibt sich daher: Wie wirken sich die mathematische Fähigkeit und das räumliche Vorstellungsvermögen von Lernenden auf ihren Umgang mit verschiedenartigen Repräsentationen im Bereich der Struktur der Materie aus?

In dieser Studie wird überprüft, ob das mathematische Vorwissen einen Einfluss auf das konzeptuelle Verständnis der Struktur der Materie im symbolischen Bereich hat. Dies wird als Arbeitshypothese vermutet, da symbolische Repräsentationen oftmals in mathematisierten Aufgabenkontexten Einsatz finden.

Analog wird untersucht, ob das räumliche Vorstellungsvermögen das konzeptuelle Verständnis der Struktur der Materie im ikonischen Bereich beeinflusst. Dies wird angenommen, da sich Repräsentationen dieser Art oftmals in Aufgaben finden, in denen mentale Rotationsbewegungen erdacht werden sollen.

Die Untersuchung wurde in einem Prä-Post-Testdesign mit N=85 Erstsemester-Bachelor-Lehramtsstudierenden der Universität Duisburg-Essen durchgeführt. Der Prätest wurde vor Beginn der Fachvorlesungen, der Posttest am Ende des ersten Fachsemesters durchgeführt.

Hierbei wurden, neben dem konzeptuellen Verständnis der Struktur der Materie im symbolischen und ikonischen Teilbereich, auch die mathematischen Fähigkeiten, sowie das räumliche Vorstellungsvermögen der Studierenden erfasst. Zusätzlich wurden noch weitere Kontrollvariablen, wie das Vor- und Fachwissen in Chemie, oder die kognitiven Fähigkeiten erhoben.

Um einen vertieften Einblick in die kognitive Verarbeitung von Repräsentationen variierenden Abstraktionsgrades zu gewinnen, wurden mit einer Teilstichprobe qualitative Prozessdaten in einer Smartpen Thinking-Aloud-Untersuchung und einer Eye-Tracking Nebenstudie erhoben.

Bei der Analyse der Testergebnisse konnte herausgestellt werden, dass Studierende sowohl im ikonischen ( $p < .05$ ) als auch im symbolischen Teilbereich ( $p < .05$ ) im Laufe des ersten Semesters signifikant hinzulernen. Hierbei konnte ein mittlerer Effekt ( $d = 0,37$ ) für den ikonischen Teilbereich und ein großer Effekt ( $d = 0,64$ ) für den Symbolischen ausgemacht werden. Auch konnte aufgezeigt werden, dass es zwischen den verschiedenen Teilbereichen der Struktur der Materie und der Zeit eine signifikante Interaktion gibt ( $F[1, 166] = 5,6, p < .05$ ), die Teilbereiche sich also unterschiedlich entwickeln.

Eine Identifizierung von möglichen Einflussvariablen auf die jeweiligen Teilbereiche mittels Korrelationsanalyse ergab, dass die Leistung im symbolischen Bereich mit der gemessenen mathematischen Fähigkeit im mittleren Bereich manifest korreliert ( $r = .374$  mit  $p < .05$ ). Die Leistung im ikonischen Bereich hingegen zeigt eine mittlere, manifeste Korrelation mit dem räumlichen Vorstellungsvermögen ( $r = .282$  mit  $p < .05$ ).

Es wird daher davon ausgegangen, dass die untersuchten Lernereigenschaften einen Einfluss auf den Wissenserwerb in den jeweiligen Teilbereichen haben. Dieser Einfluss wird zurzeit weiter analysiert und soll im Rahmen des Vortrags vorgestellt werden.

## G10

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus S107

### Ein Rahmenmodell zu "Scientific Reasoning and Argumentation" als methodisches Werkzeug im Biologieunterricht

**Tobias Dorfner, Christian Förtsch, Michael Germ, Birgit Jana Neuhaus**

Ludwig-Maximilians-Universität München, Deutschland; [tobias.dorfner@bio.lmu.de](mailto:tobias.dorfner@bio.lmu.de)

Die geringen Leistungen deutscher Schülerinnen und Schüler bezüglich ihrer naturwissenschaftlichen Fähigkeit („scientific literacy“; Bybee, 2002) in internationalen Vergleichsstudien (z. B. TIMSS 1995: Baumert, Bos, & Lehmann, 2000) führten 2005 zur Einführung der Bildungsstandards für die Fächer Biologie, Chemie und Physik. Die Kompetenzbereiche der Bildungsstandards gliedern sich in Inhalts- (Fachwissen) und Prozessdimensionen (Erkenntnisgewinnung, Kommunikation, Bewertung). Neben den Bildungsstandards gibt es nationale und internationale Rahmenmodelle zur Beschreibung naturwissenschaftlicher Fähigkeit, wobei Kampa und Köller (2016) ein wiederholtes Vorkommen von Fachwissen und Erkenntnisgewinnung in diesen Rahmenmodellen zeigten. Des Weiteren konnten die Autoren mit einer Pilotstudie zeigen, dass Schülerinnen und Schüler im Problemlöseprozess sowohl Facetten des Kompetenzbereichs Fachwissens als auch des Kompetenzbereichs Erkenntnisgewinnung nutzen. Allerdings liegen kaum Erkenntnisse über eine gezielte Förderung dieser beiden Kompetenzbereiche im Biologieunterricht vor (Kampa & Köller, 2016).

Mayer (2007) definiert im Forschungskontext vier Prozessvariablen für den Prozess der Erkenntnisgewinnung: (1) naturwissenschaftliche Fragen formulieren, (2) Hypothesen generieren, (3) Untersuchungen planen und (4) Daten analysieren/Schlussfolgerungen ziehen. Zusätzlich setzt Mayer (2007) diese vier Prozessvariablen mit wissenschaftlichem Denken („scientific reasoning“) als kognitionspsychologisches Konstrukt in einem Strukturmodell in Beziehung. Auch in Bereichen der Kognitions- und Entwicklungspsychologie werden die Begriffe „scientific reasoning“ und „scientific argumentation“ in den letzten Jahren vermehrt beschrieben (Zimmerman, 2000). Forschung in diesen Bereichen hat die Intention das Wissen über die natürliche Welt zu fördern und Prozesse zu unterstützen, dieses Wissen zu explorieren, zu evaluieren, zu überarbeiten und zu kommunizieren (Klahr et al., 2011). Fischer et al. (2014) haben anhand psychologischer Theorien ein allgemein und interdisziplinär gültiges Rahmenmodell entwickelt, um Fähigkeiten im Rahmen der Konstrukte „scientific reasoning“ und „scientific argumentation“ (SRA; Fischer et al. 2014) zu fördern. Das Rahmenmodell zu SRA enthält dabei die folgenden acht epistemischen Aktivitäten (EPA): (1) Problem identification, (2) Questioning, (3) Hypothesis generation, (4) Construction and redesign of artefacts, (5) Evidence generation, (6) Evidence evaluation, (7) Drawing conclusions, (8) Communicating and scrutinising.

Diese Studie greift, begründet durch eine Verzahnung von „scientific inquiry“ mit „scientific reasoning“ (Mayer 2007), das Rahmenmodell zu SRA auf, um Anregungen für eine gezielte Förderung der Inhalts- und Prozessdimension im Biologieunterricht zu erhalten. Die Fragestellungen sind: (a) Welche und wie oft kommen die EPA des SRA in regulärem Biologieunterricht vor? (b) Welchen Einfluss hat das vermehrte Vorkommen der EPA des SRA im Biologieunterricht auf die Schülerleistung? Zusätzlich zur Überprüfung dieser beiden Fragestellungen wurde Biologieunterricht anhand des SRA-Rahmenmodells geplant. Die Studie ist in das kooperative, BMBF-geförderte Projekt ProWiN eingebettet. Im biologischen Teilprojekt wurden je zwei Unterrichtsstunden (N=85 Videos) von 43 Lehrkräften in der 9. Jahrgangsstufe zum Thema Neurobiologie videographiert. Zusätzlich wurde jeweils vor und nach der Videographie ein Schülerleistungstest mit 18 bzw. 22 Items erhoben und mittels Partial-credit-model Rasch-skaliert (Item-Infit/Outfit $\leq$ 1,5; Personenreliabilitäten $>$ 0,63; Itemreliabilitäten=1,0). Die videographierten Biologiestunden wurden bezüglich der EPA des Rahmenmodells nach Fischer et al. (2014) analysiert und deren Auftreten in den einzelnen Unterrichtsstunden kodiert. Außerdem wurde der Einfluss des Vorkommens von mehr EPA im Biologieunterricht des Rahmenmodells auf die Schülerleistung mittels Mehrebenenpfadmodell berechnet. Die Analysen der videographierten Biologieunterrichtsstunden zeigten einerseits ein unterschiedlich häufiges Auftreten der einzelnen EPA, andererseits zeigte das Vorkommen von mehr EPA positive signifikante Einflüsse auf die Schülerleistung ( $\beta=0,30$ ,  $SE=0,15$ ,  $p<0,05$ ,  $R^2=0,09$ ). Das bedeutet, dass sich in unserer Teilstudie bereits lediglich das Vorkommen von mehr EPA positiv auf die Schülerleistung auswirkt. Basierend auf den Ergebnissen planen wir drei Biologiestunden unter expliziter Berücksichtigung aller acht EPA. Ziel war es Ergebnisse über eine fachspezifische Ausschärfung des Rahmenmodells zu erlangen sowie weitere Anregungen zur Förderung des SRA im Biologieunterricht zu erhalten. Die Beispielstunden können als „best practice“ Beispiele dienen und können somit auch als Maß, um die Qualität der EPA im Biologieunterricht zu analysieren, herangezogen werden. Die Stunden werden detailliert auf der Tagung vorgestellt.

### Lassen sich Gestaltgesetze zum Aufbau von chemischen Demonstrationsexperimenten bestätigen? Evidenz einer experimentellen Eye-Tracking-Studie

**Andreas Nehring, Sebastian Busch**

Leibniz Universität Hannover, Deutschland; [nehring@idn.uni-hannover.de](mailto:nehring@idn.uni-hannover.de)

#### Einleitung

Die von dem deutschen Chemiedidaktiker Heinz Schmidkunz auf den Chemieunterricht angewendeten Gestaltgesetze der Wahrnehmungspsychologie stellen Regeln dar, die Lehrkräfte beim Aufbau von Demonstrationsexperimenten im Unterricht beachten sollen. Danach sollen Gestaltgesetze die visuelle Aufmerksamkeit auf inhaltlich zentralen Geräten einer Experimentalapparatur erhöhen (Schmidkunz, 1983, 1992). Obwohl seit den Neunzehnjahrig Jahren zahlreiche angehende Chemielehrkräfte an Universitäten und Seminarstandorten die Anwendung der Gestaltgesetze erlernen, wurden die Gesetze bisher keiner direkten systematischen, empirischen Überprüfung unterzogen.

#### Theoretischer Hintergrund □

Mit Orquin und Mueller Loose (2013, S. 191) lässt sich visuelle Aufmerksamkeit als „selectivity in perception“ definieren, wobei zwischen bottom-up- und top-down-gesteuerten Prozessen unterschieden wird. Während bottom-up-Prozesse stimulusgesteuert sind und vom jeweils betrachteten Material abhängen, werden top-down-Prozesse von den Intentionen, den Vorerfahrungen bzw. dem Wissen des oder der Betrachtenden beeinflusst (Stigchel et al., 2009). Bei der Wahrnehmung greifen beide Prozesse ineinander. Mit dem Fokus auf den Experimentalaufbau postulierte Schmidkunz in dieser Hinsicht einen expliziten Einfluss auf die bottom-up-gesteuerte visuelle Aufmerksamkeit.

Zu den hier untersuchten Gestaltgesetzen gehören das Gesetz des Figur-Grund-Kontrastes, das Gesetz der glatt durchlaufenden Linie, das Gesetz der Dynamik von links nach rechts, das Gesetz der Nähe, das Gesetz der Einfachheit sowie das Gesetz der Symmetrie.



## Fragestellung

Die folgende Fragestellung war Gegenstand der vorliegenden Studie: Inwiefern lässt sich ein Einfluss der Gestaltgesetze auf die bottom-up-gesteuerte visuelle Aufmerksamkeit nachweisen, wenn Schülerinnen und Schüler Demonstrationsexperimente betrachten?

## Methode

Zur Beantwortung der Forschungsfrage wurde ein experimentelles Studiendesign gewählt, in dem Demonstrationsexperimente (Fachkontexte: Eigenschaften von Gasen und Stofftrennung) gegeneinander getestet wurden, in denen je eines von sechs Gestaltgesetzen beachtet („Gestaltexperimente“) oder gezielt manipuliert wurde („Kontrollexperimente“). Die Experimente wurden abfotografiert und, zum Konstanthalten von Reihenfolgeeffekten, in Form von zwölf verschiedenen Folgen in einen Eye-Tracker integriert (Eyegaze Edge Eye-Follower 2.0, 120 Hz sampling rate, Fixation: 100 ms als default). 143 Schülerinnen und Schüler der Sekundarstufen I und II wurden randomisiert auf je eine Serie zugeteilt und betrachteten die Demonstrationsexperimente. Um eine Steuerung der top-down-kontrollierten Aufmerksamkeit zu vermeiden, wurde ein allgemein gehaltener Beobachtungsauftrag gegeben. Da weder den Probanden noch der durchführenden Hilfskraft mitgeteilt wurde, inwiefern welche Experimente als Gestalt- oder Kontrollexperimente erscheinen würden, wurde ein Doppel-Blind-Design umgesetzt.

Zur Analyse der Daten wurden inhaltlich relevante Geräte mit areas of interest markiert und Gestalt- und Kontrollexperimente anhand von Indikatoren für visuelle Aufmerksamkeit mittels Mann-Whitney-U-Tests gegen einander getestet. Dabei wurden u. a. die Anzahl der Fixationen, die Gesamtblickdauer, die mittlere Fixationsdauer sowie Indikatoren für den ersten, zweiten und dritten Blickdurchlauf analysiert.

## Ergebnisse

Die Daten zeigen signifikante positive Effekte zugunsten der Gestaltexperimente, die von starken (z. B.  $r = 0.60$ ; Gesamtblickdauer beim Gesetz der Einfachheit) bis zu kleinen Effekten reichen (z. B.  $r = 0.14$ ; Gesamtblickdauer beim Gesetz der Nähe). Dabei existieren auch Indikatoren für visuelle Aufmerksamkeit, die sich nicht signifikant voneinander unterscheiden. So lassen sich z. B. beim Gesetz der Dynamik von links nach rechts zwar nur Hinweise für kürzere Dauern von Fixationen nachweisen, jedoch keine Unterschiede in weiteren Indikatoren. Entsprechend können sich – je nach Gesetz, Experiment und analysiertem Gerät – Effektstärken und Alphafehlerwahrscheinlichkeiten unterscheiden, so dass sich ein differentielles und kein pauschales Befundmuster ergibt. Hinweise für einen negativen Einfluss auf die visuelle Aufmerksamkeit konnten nicht gefunden werden.

## Diskussion

Insgesamt verdeutlichen die Ergebnisse dieser experimentellen Doppel-Blind-Studie, dass die Gestaltgesetze einen differentiell-positiven, jedoch keinen pauschalen Einfluss auf die bottom-up kontrollierte visuelle Aufmerksamkeit haben könnten. Gleichzeitig eröffnet die Studie ein neues Feld naturwissenschaftsdidaktischer Forschung, in dem das Zusammenspiel von bottom-up- und top-down-Prozessen, der Einfluss des bestehenden Schülerwissens bzw. des Verhaltens von Lehrkräften untersucht werden kann. Ebenso könnte die Untersuchung von Interaktionseffekten zwischen Gestaltgesetzen vielversprechend sein. Der Vortrag präsentiert die Daten im Detail wie auch Praxisimplikationen und Möglichkeiten für weitere Studien zur visuellen Aufmerksamkeit im Bereich der Naturwissenschaftsdidaktiken.

## Wie entscheidend ist die Situation? Fachspezifität des Schülererlebens bei der Vorhersage von Wissenserwerb

**Kristina Kögler<sup>1</sup>, Richard Göllner<sup>2</sup>**

<sup>1</sup>Universität Hohenheim; <sup>2</sup>Universität Tübingen; [koeqler@econ.uni-frankfurt.de](mailto:koeqler@econ.uni-frankfurt.de)

### Theoretischer Hintergrund

Selbstauskünfte von Lernenden zu ihrem subjektiven Erlebens des Unterrichtsgeschehens spielen bei der Erforschung von lernbezogenen Emotions- und Motivationskonstrukten eine zentrale Rolle und sind überdies wichtige Prädiktoren bei der Aufklärung der Variabilität von Schulleistungen (Debus 2000; Rausch et al. 2010). Insbesondere situationsnahe Experience-Sampling-Verfahren die die möglichen Verzerrungseffekte retrospektiver Erhebungen vermeiden, führen in diesem Zusammenhang zu umfangreichen Befunden (vgl. Scollon et al. 2003; Sembill et al. 2008). Mit der Erhebung des Schülererlebens im Unterrichtsprozess ist die Erwartung verknüpft, die Situationsspezifität valide abzubilden und zur differenzierten Vorhersage von Outcomes zu nutzen (Csikszentmihalyi & Larson 1987). Offen ist allerdings, inwiefern die Selbstauskünfte der Lernenden ausschließlich die spezifische Situation erfassen oder darüber hinaus stabile personengebundene Anteile enthalten.

### Fragestellung

Davon ausgehend stellen sich die Fragen, (1.) ob das Schülererleben in unterschiedlichen Fächern prädiktiv für den Wissenserwerb ist und (2.) inwiefern die Effekte des Erlebens auf den Wissenserwerb unter Kontrolle der Fachspezifität des Schülererlebens stabil bleiben. Die Beantwortung dieser Fragen erfolgt im Rahmen eines prozessorientierten Forschungszugangs mittels Continuous-State-Sampling in unterschiedlichen Fächern.

### Methode

Die Datengrundlage entstammt einer Prozessuntersuchung in den Fächern Rechnungswesen und Betriebswirtschaft, in der das Schülererleben mittels Continuous-State-Sampling während des Unterrichts in insgesamt 56 Unterrichtseinheiten mit äquidistanten Messintervallen erhoben wurden. Es nahmen  $N=96$  Lernende in vier Klassen an der Erhebung in einer kaufmännischen Schule in Nordbayern teil, in jeder Klasse gingen 12 Unterrichtseinheiten mit jeweils siebenminütig getakteten Messzeitpunkten in die Datenauswertung ein. Eine Klasse wurde durchgängig von einer Lehrperson unterrichtet und wird daher angesichts der vorliegenden Fragestellung von den Analysen ausgeschlossen. Damit resultieren für jeden Lernenden jeweils rund 70 Messzeitpunkte in Rechnungswesen und Betriebswirtschaft. Die eingesetzten Items bezogen sich auf die subjektive Verständlichkeit („Verstehe, worum es gerade geht“), das Interesse („Bin interessiert“) und die Kontrollierbarkeit der Unterrichtssituation („Nachdenkzeit reicht mir“) und waren auf einer stufenlosen Skala von 0 bis 100 zu beantworten. Vorwissen und Wissenserwerb wurden mittels identischer lernzielorientierter Tests vor und nach der Inhaltseinheit in beiden Fächern gemessen. Die Daten wurden auf zwei Ebenen modelliert, um die Variabilität innerhalb von Stunden (Ebene der Messzeitpunkte) und zwischen Lernenden (Ebene der Individuen) abzubilden.

### Ergebnisse

Die Ergebnisse eines ersten mehrbenenanalytischen Faktorenmodells zeigten sowohl auf der Ebene der einzelnen Messzeitpunkte ( $.30 \leq r \leq .58$ ) als auch auf der Ebene von Schülerinnen und Schülern ( $.20 \leq r \leq .93$ ) deutliche Interkorrelationen der drei Erlebensdimensionen in beiden Unterrichtseinheiten. Zudem fanden sich auf der Ebene von Lernenden konstruktsspezifische Erlebenskonsistenzen in Rechnungswesen und Betriebswirtschaft ( $.68 \leq r \leq .73$ ), welche auf der Ebene einzelner Messzeitpunkte nicht nachweisbar waren ( $.00 \leq r \leq .04$ ).

Ein weiterführendes Modell mit zwei globalen Erlebensfaktoren für die Unterrichtsfächer zeigte insgesamt eine gute Passung,  $\chi^2(24) = 58.03$ ,  $p < .001$ ; CFI = .96; TLI = .95; SRMRw = .03; und SRMRb = .05, wobei lediglich der Erlebensfaktor im Fach Rechnungswesen auch unter Kontrolle des Vorwissens mit dem Wissenserwerb von Schülerinnen und Schülern assoziiert war ( $\beta = .24$ ,  $p = .020$ ). Für die betriebswirtschaftlichen Unterrichtseinheiten zeigte sich kein Zusammenhang ( $\beta = .02$ ,  $p = .860$ ). Unter gegenseitiger Kontrolle des Unterrichtserlebens in beiden Unterrichtseinheiten fanden sich hingegen weder für das Fach Rechnungswesen noch für das Fach Betriebswirtschaft Zusammenhänge mit dem Wissenserwerb von Schülerinnen und Schülern.

Es lässt sich festhalten, dass das Unterrichtserleben bei der Vorhersage des Wissenserwerbs erwartungskonforme Relevanz offenbart, wobei dies nur für das Fach Rechnungswesen gilt, was zu diskutieren ist. Allerdings zeigen die Ergebnisse auch, dass die Fach- bzw. Situationsspezifität des Erlebens geringer ausfällt als erwartet. Schlussendlich entfaltet sich im situativen Erleben vermutlich eine Interaktion zwischen dem Lernenden und der Unterrichtsumwelt, die mit der Messung in situ erfasst wird und für die Stabilität des subjektiven Erlebens auch in wechselnden situativen Konstellationen verantwortlich ist.

### **Mehrdimensionales Interesse in Physik – Welche Rolle spielen Basisdimensionen der Unterrichtsqualität zur Förderung des Interesses?**

**Ariane S. Willems, Julia Rudnik**

Georg-August-Universität Göttingen, Deutschland; [awillem1@gwdg.de](mailto:awillem1@gwdg.de)

Schule und Unterricht haben den Anspruch, mehrdimensionale Bildungsziele zu erreichen. Neben der Vermittlung von Fachwissen zählt hierzu auch, fachspezifische Interessen von Schüler\*innen zu fördern. Studien zeigen jedoch, dass Schüler\*innen vor allem in den naturwissenschaftlichen Fächern – insbesondere in Physik – wenig Interesse haben und dieses im Laufe der Schulzeit absinkt (Daniels, 2008; Hoffmann et al., 1995). Eine Ursache hierfür wird in der erlebten Qualität des Unterrichts vermutet. In der empirischen Unterrichtsforschung werden drei Basisdimensionen der Unterrichtsqualität unterschieden, die als relevant für die Lernentwicklung von Schüler\*innen gelten: strukturierte und effiziente Klassenführung, konstruktive Schülerorientierung und Unterstützung sowie kognitives Aktivierungspotential (Klieme & Rakoczy, 2008). Während Studien auf die differenzielle Bedeutsamkeit der Dimensionen für den Leistungszuwachs hinweisen, liegen bisher insbesondere für das Fach Physik keine belastbaren Befunde zum Zusammenhang zwischen der Unterrichtsqualität und dem Interesse von Schüler\*innen vor. Basierend auf der Person-Gegenstands-Theorie des Interesses (Krapp, 2002) fokussieren wir in dem vorliegenden Beitrag das Konzept des individuellen, d.h. zeitlich relativ überdauernden und stabilen Interesses von Schüler\*innen und differenzieren hier zwischen drei Dimensionen: dem Fach-, Sach- und Inhaltsinteresse (vgl. auch Hofmann et al., 1995). Das Ziel der vorliegenden Studie ist es (i) ein Instrument zur differenzierten Erfassung des Fach-, Sach- und Inhaltsinteresses von Schüler\*innen in Physik zu entwickeln und zu validieren sowie (ii) Zusammenhänge zwischen der wahrgenommenen Unterrichtsqualität und den drei Interessendimensionen im Fach Physik zu untersuchen.

Analysegrundlage sind Daten der IPhy-Studie ('Fach-, Inhalts- und Sachinteresse in Physik'), in deren Zentrum die Wahrnehmung der Unterrichtsqualität aus Schülersicht sowie verschiedene Dimensionen des Physikinteresses stehen. Insgesamt wurden N=300 Schüler\*innen (54% weiblich) der 8. Klasse an Gymnasien (13 Schulklassen) im Anschluss an eine Unterrichtsreihe zum Themengebiet Elektrizität mittels eines standardisierten Fragebogens zu ihrem Interesse am Physikunterricht (Fachinteresse: 7 Items,  $=.87$ ), ihrem Sachinteresse an Physik als Naturwissenschaft (Sachinteresse: 5 Items,  $=.87$ ) sowie ihrem inhaltlichen Interesse am Unterrichtsthema Elektrizität (Inhaltsinteresse: 10 Items,  $=.90$ ) befragt. Die Dimensionalitätsanalysen mittels latenter konfirmatorischer Faktoranalysen unter Kontrolle der Mehrebenenstruktur der Daten (Mplus 8.0) zeigen, dass das postulierte 3-Faktormodell zur Beschreibung des Interesses von Schüler\*innen die empirischen Daten gut abbildet (CFI=.93, RMSEA=.04, adj.-BIC=12570.26) und erwartungsgemäß besser auf die Daten passt als ein 1-Faktormodell im Sinne eines 'Generalfaktor Interesse' (CFI=.82, RMSEA=.09, adj.-BIC=13050.23).

Die drei Basisdimensionen der Unterrichtsqualität wurden jeweils durch 6 Items operationalisiert. Auch hier weisen entsprechende konfirmatorische Faktoranalysen auf einen zufriedenstellenden Modellfit des 3-Faktormodells zur Beschreibung der Unterrichtsqualität hin (CFI = .91, RMSEA = .06). Die Zusammenhänge zwischen den Basisdimensionen der Unterrichtsqualität einerseits und den drei Dimensionen des Interesses wurden schließlich mittels Strukturgleichungsmodellen auf Individualebene – jedoch unter statistischer Kontrolle der Mehrebenenstruktur – untersucht. Die Befunde der Strukturgleichungsmodellierung zeigen, dass die drei Dimensionen der Unterrichtsqualität zunächst jeweils einen direkten positiven Effekt auf das individuelle Physik-Fachinteresse der Schüler\*innen haben. Dieses wiederum wirkt vermittelt über das Inhaltsinteresse positiv auf das Sachinteresse der Schüler\*innen. Direkte Effekte der Unterrichtsqualität auf das Sachinteresse an Physik können nicht nachgewiesen werden. Die Befunde unserer Analysen verdeutlichen zunächst die Mehrdimensionalität (Fach-, Sach- und Inhaltsbezogenheit) des Physikinteresses von Schüler\*innen und unterstreichen somit auch aus theoretisch-konzeptueller Sicht die Notwendigkeit und den Mehrwert einer fachspezifisch differenzierten Erfassung von Interessen. Schließlich weisen die Befunde auch darauf hin, dass die individuelle Ausprägung der Interessendimensionen in unterschiedlichem Maß durch die Qualität des Unterrichts gefördert werden kann. Aus schulpraktischer Sicht erscheint vor allem der Befund von Bedeutung, dass auch ein relativ dauerhaftes Inhalts- und Sachinteresse an physikalischen Themen über das unmittelbare, d.h. unterrichtsbezogene Fachinteresse gefördert werden kann.

# G11

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus, Aula 33

## **Berufsfachliche Kompetenzen von Elektronikern für Automatisierungstechnik – Kompetenzdimensionen, Messverfahren und erzielte Leistungen**

Chair(s): **Leo van Waveren** (Technische Universität Kaiserslautern, Deutschland)

DiskutantIn(nen): **Felix Walker** (Technischen Universität Kaiserslautern)

Nicht zuletzt Dank des DFG Schwerpunktprogrammes „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ konnte im gewerblich-technischen als auch im kaufmännischen Bereich die Entwicklung und die Modellierung der beruflichen Kompetenz in den Fokus der fachdidaktischen Forschung gerückt werden. Das Symposium gibt einen Überblick zu ausgewählten Ergebnissen des Projekts „Erstellung eines empirisch geprüften Fachkompetenzmodells für Elektroniker (KOKO EA)“, das im Rahmen des ASCOT Programms als Verbundprojekt der Universität Stuttgart, der Pädagogischen Hochschule Ludwigsburg und des Elektro Technologie Zentrums Stuttgart gefördert wurde.

Ein Hauptanliegen dieses Symposiums ist es, Erkenntnisse zur Messung und Modellierung berufsfachlicher Kompetenz bei Elektronikern für Automatisierungstechnik (EfA) von den theoretischen Überlegungen zu den Kompetenzkonstrukten über die entwickelten Messinstrumente und deren Güte bis hin zu möglichen Verwertungsperspektiven (z. B. in der Unterrichtspraxis, Prüfungen oder eines Large-Scale-Assessments) vorzustellen.

Die Beiträge des Symposiums greifen den aktuellen Forschungsstand zum Thema Messung und Modellierung berufsfachlicher Kompetenz auf und widmen sich einzelnen Aspekten wie z. B. der Strukturmodellierung, der Konstruktvalidität und den Zusammenhängen der einzelnen Facetten berufsfachlicher Kompetenz. Im Anschluss an vorliegende Arbeiten zur Kompetenzmodellierung in der beruflichen (Nickolaus und Seeber 2013) bzw. der gewerblich-technischen Berufsbildung (Nickolaus 2011) wurde zu Beginn des Projekts eine zweidimensionale Grundstruktur berufsfachlicher Kompetenz unterstellt, die durch das Fachwissen und die berufsfachliche Problemlösekompetenz aufgespannt wird.

Für das Fachwissen sind gegen Ende der Ausbildung bisher für alle untersuchten Berufe mehrdimensionale Modelle als adäquat bestätigt worden (Nickolaus und Seeber 2013) weshalb angenommen wurde, dass dies auch für den EfA unterstellt werden kann.

Für die berufsfachliche Problemlösekompetenz unterstellten wir eine zweidimensionale Struktur und entwickelten Tests um die analytische Problemlösekompetenz einerseits mit Hilfe von steuerungstechnischen realen Modellanlagen und andererseits mit hochauthentischen Simulationen der gleichen Systeme zu erfassen. Für die konstruktive Problemlösekompetenz wurde ein Test entwickelt, dessen Aufgaben a) in einer realen Programmierumgebung und b) mit einem Papier- und Bleistift zu bearbeiten waren, um zu prüfen, inwieweit die gleiche Fähigkeit zur Bearbeitung benötigt wird.

Der erste Beitrag von Leo van Waveren stellt unterschiedliche theoretische Ansätze zur Strukturmodellierung des Fachwissens ins Zentrum und welche Kompetenzstruktur bei EfA am Ende der Ausbildung beobachtet werden kann.

Der Beitrag von Felix Walker befasst sich mit der Entwicklung und Validierung von Tests zur Erfassung der analytischen Problemlösekompetenz in steuerungstechnischen Anlagen der Konstruktvalidität konstruktiver und analytischer Problemlösekompetenz.

Geprüft wird in dem Beitrag von Nico Link und Bernd Geißel die in den vorangegangenen Beiträgen analysierten Konstrukte in einer integrativen Perspektive. Im Mittelpunkt steht dabei die Frage, welche Zusammenhänge zwischen der fachspezifischen Problemlösefähigkeit in analytischen und konstruktiven Anforderungskontexten einerseits und dem Fachwissens andererseits bestehen.

Der Beitrag von Matthias Hedrich untersucht die Bedeutung schulischer, betrieblicher und privater Einflussfaktoren für die Fachkompetenzentwicklung und verwendet dafür ein Instrumentarium, welches systemische, soziale und bildungsbiografische Kontextfaktoren erfasst (Baethge und Baethge-Kinsky 2013).

Der Zuschnitt des Symposiums gibt die Möglichkeit, das Gesamtkonstrukt berufsfachlicher Kompetenz entlang der einzelnen Elemente des angenommenen Modells vorzustellen und die Bezüge zwischen unterschiedlichen Facetten berufsfachlicher Problemlösefähigkeit und berufsfachlichen Wissens nachzuzeichnen.

*Beiträge des Symposiums*

### **Struktur- und Niveaumodell des Fachwissens bei Elektronikern der Automatisierungstechnik**

**Leo van Waveren**

Technische Universität Kaiserslautern

\*Theoretischer Hintergrund\*

Die ASCOT-Initiative des Bundesministeriums für Bildung und Forschung (<http://www.ascot-vet.net/>) hatte zum Ziel für ausgewählte, anspruchsvolle Berufsbilder Aufschlüsse zur berufsfachlichen Handlungskompetenz zu gewinnen. An der Universität Stuttgart wurde dazu eine Studie für den Beruf des Elektrikers für Automatisierungstechnik am Ende der Ausbildung durchgeführt.

Das Fachwissen erweist sich sowohl im allgemeinbildenden (SCHERER 2012, S. 40ff. und S. 16) als auch im gewerblich-technischen Bereich (NICKOLAUS u.a. 2012, S. 265ff.) als wichtiger Prädiktor für die Problemlösekompetenz.

Im gewerblich-technischen Bereich wurden durchgängig mehrdimensionale Strukturen des Fachwissens bestätigt, unter anderem für Mechatroniker (GÖNNENWEIN/NITZSCHKE/SCHNITZLER 2011) die Grundstufe Bau (PETSCH/NORWIG/NICKOLAUS 2015) und Schreiner (PITTICH 2013).

Für den Kfz-Mechatroniker (GSCHWENDTNER 2008) und den Elektriker für Energie- und Gebäudetechnik (GEIßEL 2008) zeigten sich am Ende der Grundbildung noch eindimensionale Strukturen des Fachwissens, die sich zum Ende der Ausbildung zu mehrdimensionalen Konstrukten entlang von Inhaltsbereichen ausdifferenzieren (NICKOLAUS U.A. 2011, S.85f; SCHMIDT/WEBER/NICKOLAUS 2014, S.560). Zu den erreichten Niveaus wird durchgängig berichtet, dass in der Regel erhebliche Anteile der Auszubildenden den curricularen Anforderungen nicht gerecht werden (Nickolaus/Seeber 2013).

Für den Elektroniker für Automatisierungstechnik wurden für die Problemlösefähigkeit eine mehrdimensionale Struktur bestehend aus einer analytischen und einer konstruktiven Teildimension bestätigt (WALKER/LINK/NICKOLAUS 2015). Offen ist, ob sich für das Fachwissen bei Elektronikern für Automatisierung am Ende der Ausbildung eine mehrdimensionale Struktur ergibt.

Dies hat einerseits vor dem Hintergrund einer fairen Abschlussprüfung durch die IHK und andererseits vor dem Hintergrund einer möglichen Neuausrichtung industrieller Elektroberufen (ZINKE U.A. 2013) erhebliche Relevanz.

#### \*Fragestellung\*

Es wird geprüft, ob sich die in anderen Berufsfeldern nachgewiesene Mehrdimensionalität des Fachwissens auch bei Elektronikern für Automatisierungstechnik entlang inhaltlicher Strukturierungen bestätigen lässt oder ob alternative Modellierungen eine bessere Passung zu den Daten aufweisen. Daran anschließend stellt sich die Frage, welche Niveaus in den einzelnen Teildimensionen erreicht werden.

#### \*Methode\*

Ausgehend von den curricularen Analysen wurde von den Probanden ein Paper-Pencil-Test im Balanced Incomplete Block Design (FREY/HARTIG/RUPP 2009, S.45) bearbeitet, dessen Konzeptionierung von einer Expertengruppe aus beruflichen Schulen, Betrieben und Prüfungsausschuss begleitet wurde.

Auf der Grundlage von den Ergebnissen von 878 Probanden aus 9 Bundesländern wurden IRT-Skalierungen unterschiedlicher Strukturierungsansätze für das Fachwissen anhand von Informationskriterien hinsichtlich der Passung zu den erhobenen Daten miteinander verglichen (BUCKLAND/BURNHAM/AUGUSTIN 1997, S.606). Neben Strukturierungen ausgehend von curricularen Schwerpunktsetzungen werden ein Ansatz auf der Basis kognitionspsychologischer Anforderungen, wie sie für Elektroniker für Energie- und Gebäudetechnik (GEIßEL 2008) in Betracht gezogen wurden und eine Differenzierung nach mathematischen Anforderungen, die sich im ingenieurwissenschaftlichen Grundstudium als strukturbildend erwiesen (BEHRENDT U.A. in Vorbereitung) berücksichtigt. Bei rund der Hälfte der Stichprobe wurden mittels des CFT-20R die kognitiven Grundfähigkeiten zur Kontrolle mit erhoben.

Niveaumodellierungen wurden nach Beaton und Allen (1992) für die Subdimensionen vorgenommen.

#### \*Ergebnisse\*

Die zu meist (90%) männlichen Probanden waren im Schnitt 20,4 Jahre alt (SD = 2) und besaßen überwiegend einen Realschulabschluss (70,5 %; rund 25 % hatten eine Hochschulzugangsberechtigung). Der alterskorrigierte IQ von 108 (SD = 14) unterstreicht, dass die Stichprobe im Vergleich mit anderen Elektroberufen aus kognitiv starken Auszubildenden besteht.

Die angenommene Mehrdimensionalität des Fachwissens wird gestützt. Es zeigt sich, dass das Konstrukt aus den drei eng assoziierten, empirisch trennbaren Teildimensionen „Automatisierungstechnik/Speicherprogrammierbare Steuerungen“, „elektrotechnische Grundlagen“ und „elektrische Energietechnik“ aufgespannt wird mit latenten Korrelationen zwischen .83 und .89. Die EAP/PV-Reliabilitäten liegen alle > .7 und erreichen damit für die drei Dimensionen akzeptable Werte.

Die Modellierung nach Beaton und Allen offenbart deutliche Diskrepanzen zwischen curricular intendierten und den in der Praxis erreichten Kompetenzständen. In jedem der Teilbereiche verbleiben zwischen 50 und 60 % der Probanden (zum Teil deutlich) hinter den in den Rahmenlehrplänen formulierten Zielsetzungen zurück.

## **Zum Zusammenhang fachspezifischer Problemlösekompetenz in konstruktiven und analytischen Anforderungskontexten**

**Felix Walker**

Technische Universität Kaiserslautern

#### \*Theoretischer Hintergrund\*

Das Konstrukt der berufsfachlichen Kompetenz lässt sich nach den bislang vorliegenden Befunden domänenübergreifend mit dem Fachwissen und der fachspezifischen Problemlösekompetenz in zumindest zwei übergeordnete Subdimensionen ausdifferenzieren (vgl. Nickolaus, 2011). Die Fähigkeit das Fachwissen in problemhaltigen variierenden Anforderungskontexten anzuwenden wird im Allgemeinen als fachspezifischer Problemlösekompetenz bezeichnet. Konnte für das Fachwissen eine mehrdimensionale berufsspezifische Inhaltsstruktur (z. B. Gschwendtner, 2011; Nickolaus, Geißel, Abele & Nitzschke, 2011) festgestellt werden, welche vor allem curricular induziert scheint, ist die Frage zur Struktur der fachspezifischen Problemlösekompetenz in unterschiedlichen Anforderungskontexten im gewerblich-technischen Bereich noch ungeklärt.

In bisherigen Arbeiten zur fachspezifischen Problemlösekompetenz erfolgte deren Erfassung insbesondere in analytischen Anforderungskontexten in Form der analytischen Problemlöse- bzw. der Fehleranalysefähigkeit (vgl. Nickolaus, Abele, Gschwendtner, Nitschke & Greiff, 2012). Der Einbezug weiterer Anforderungskontexte zur Erfassung der fachspezifischen Problemlösekompetenz wurde im gewerblich-technischen Bereich bislang nicht unternommen. Ergebnisse von Leutner, Fleischer, Wirth, Greiff und Funke (2012) im allgemeinbildenden Bereich konnten nachweisen, dass das Konstrukt der Problemlösekompetenz analytische (Fehlersuche) und eher konstruktive (Entscheidungen treffen und Systeme entwerfen) Anforderungskontexte umfasst, welche sich empirisch separieren lassen (vgl. Leutner et al., 2012, S. 37).

#### \*Fragestellung\*

In wie fern die fachspezifische Problemlösekompetenz im gewerblich-technischen Bereich ebenfalls analytische und konstruktive Anforderungskontexte umfasst, und ob sich diese bezogen auf das Konstrukt der fachspezifischen Problemlösekompetenz als konvergent valide erweisen, soll im Rahmen dieses Beitrages beantwortet werden.

#### \*Methode\*

Nachdem das zu Grunde gelegte Verständnis der fachspezifischen Problemlösekompetenz in konstruktiven und analytischen Anforderungskontexten vorgestellt wurde, werden anschließend Teile der entwickelten Instrumente vorgestellt. Geprüft wird im Kern die Hypothese, ob sich, bezogen auf die fachspezifische Problemlösekompetenz, konstruktive und analytische Anforderungskontexte als divergent bzw. konvergent valide erweisen. Die Hypothesenprüfung wird anhand der Daten von Elektronikerinnen und Elektronikern der Automatisierungstechnik (n=150) am Ende ihrer Ausbildung vorgenommen.

#### \*Ergebnisse\*

Die Analysen sind momentan noch nicht ganz abgeschlossen, werden dies aber bis zur Tagung sein. Erwartet wird ein mittlerer Zusammenhang der fachspezifischen Problemlösefähigkeit in konstruktiven und analytischen Anforderungskontexten, welcher unter Berücksichtigung des Fachwissens geringer ausfällt.

#### \*Diskussion\*

Mit diesem Beitrag werden theoretische Annahmen zu analytischen und konstruktiven Anforderungen bei Elektronikerinnen und Elektronikern der Automatisierungstechnik erstmals empirisch überprüft, welche insbesondere in unterrichtspraktischer Perspektive gewinnbringend angewendet werden können.

### **Konstruktvalidität analytischer und konstruktiver Problemlösekompetenz bei Elektroniker/innen für Automatisierungstechnik**

**Nico Link<sup>1</sup>, Felix Walker<sup>1</sup>, Bernd Geißel<sup>2</sup>**

<sup>1</sup>Technische Universität Kaiserslautern, <sup>2</sup>Pädagogische Hochschule Ludwigsburg

#### \*Theoretischer Hintergrund\*

Elektroniker für Automatisierungstechnik arbeiten an komplexen automatisierten Industrieanlagen und sind für deren Programmierung, Inbetriebnahme und Instandhaltung verantwortlich (Internetauftritt der Arbeitsagentur, 2014). Der Fokus dieses Beitrags liegt auf der Konstruktvalidität analytischer und konstruktiver Problemlösekompetenz.

Aus curriculärer Sicht sollten Auszubildende am Ende ihrer Ausbildung in der Lage sein, Automatisierungssysteme zu programmieren sowie aus Fehlerdiagnosen Folgerungen für die Fehlerbeseitigung abzuleiten (KMK, 2003). Zu ähnlichen Ergebnissen führte die Analyse beruflicher Tätigkeiten. So stellen Programmänderungen bzw. die eigenständige Neuprogrammierungen von speicherprogrammierbaren Steuerungen (SPS) den Schwerpunkt konstruktiver Problemsituationen dar (Zinke, Schenk & Kröll, 2014). Das Analysieren und Prüfen von fehlerhaften Steuerungsprogrammen, wie es bei einer nicht problemlos verlaufenden Inbetriebnahme oder beim Ausfall eines Automatisierungssystems vorkommt, repräsentiert den Kern analytischer Problemsituationen (ebd.). Wird im Folgenden von analytischer Problemlösekompetenz gesprochen, so sind darunter die kognitiven Voraussetzungen dafür zu verstehen, dass eine Person die Ursache für ein fehlerhaftes Automatisierungssystem eindeutig identifizieren kann (Walker, Link & Nickolaus, 2015). Die kognitiven Voraussetzungen, die Auszubildende in die Lage versetzen, Programmänderungen bzw. Neuprogrammierungen von speicherprogrammierbaren Steuerungen vorzunehmen, werden dagegen als konstruktive Problemlösekompetenz bezeichnet (Link, 2016). In vorangegangenen Untersuchungen erwiesen sich Papier-und-Bleistift-Tests sowie auf realen Arbeitsprozessen basierende Computersimulationen als geeignete Messinstrumente zur Erfassung fachlicher Kompetenzen im Hinblick auf die Einhaltung wissenschaftlicher Gütekriterien, welche ergänzt werden durch die Kriterien der Praktikabilität (Testzeit, Aufwandsgrenzen) und Skalierbarkeit (Nickolaus, Gschwendtner & Abele, 2009).

Geeignete Testverfahren zur Messung analytischer und konstruktiver Problemlösekompetenz im Beruf Elektroniker für Automatisierungstechnik, die den zuvor genannten Kriterien genügen, liegen bislang noch nicht vor.

#### \*Fragestellung\*

Vor diesem Hintergrund wurde in diesen beiden Teilstudien den Fragen nachgegangen, ob es möglich ist, (1.) die analytische Problemlösekompetenz authentisch (valide) und zugleich mit geringem Aufwand (ökonomisch) mit einer Computersimulation zu erfassen und (2.) die konstruktive Problemlösekompetenz valide und zugleich mit geringem Aufwand (ökonomisch) mit einem Papier-und-Bleistifttest (PBT) zu messen.

#### \*Methode\*

Es wurden jeweils für die analytische als auch die konstruktive Problemlösekompetenz acht Problemstellungen, basierend auf curricularen Analysen und in Abstimmung mit Experten beruflicher Ausbildungspraxis, entwickelt. Zur Klärung der Fragen, ob Leistungen, die in der Computersimulation bzw. im PBT erbracht wurden, mit den „realen“ Leistungen am Automatisierungssystem und in der Entwicklungsumgebung als valide interpretiert werden dürfen, wurde ein diagnostischer Zugang gewählt, der sich bereits beim Beruf des Kfz-Mechatronikers als erfolgreich bewährt hat (Nickolaus, Gschwendtner & Abele, 2009). Bei diesem Ansatz werden die Auszubildenden zufällig (randomisiert) einer von zwei Gruppen zugeteilt und bekommen dieselben Problemstellungen in unterschiedlichen Reihenfolgen zur Bearbeitung vorgelegt (Cross-Over-Design).

#### \*Ergebnisse\*

Für die Beantwortung der Forschungsfragen wurden Elektroniker für Automatisierungstechnik (jeweils n=300) am Ende ihrer Ausbildung untersucht. Die Testwerte der analytischen und der konstruktiven Problemlösekompetenz wurden jeweils gruppenspezifisch (d.h. Gruppe 1 und 2 getrennt) mit komplexen statistischen Verfahren (Strukturgleichungsmodellen) untersucht.

Bezogen auf die analytische Problemlösekompetenz zeigt sich, dass in beiden Gruppen die mit der Computersimulation und die am „Realmodell“ erfassten Kompetenzen sehr hoch korrelieren ( $r \geq .85$ ) und es für die Erfassung der Fehlerdiagnosekompetenz empirisch gerechtfertigt ist anzunehmen, dass es (statistisch) keinen Unterschied macht, ob ein Problem an der realen oder simulierten Automatisierungsanlage bearbeitet wird (Messgenauigkeit (SEM-Reliabilität) = .75).

Für die konstruktive Problemlösekompetenz kann ebenfalls gezeigt werden, dass die erbrachten Testleistungen im PBT und in der Entwicklungsumgebung hoch miteinander korrelieren und als konvergent valide interpretiert werden können. D. h., dass auch für die konstruktive Problemlösekompetenz gezeigt werden kann, dass auf Basis eines PBTs die in der Entwicklungsumgebung erbrachten Leistungen valide und mit vergleichsweise geringem Aufwand und hoher Messgenauigkeit (SEM-Reliabilität = .83) erfasst werden können.

#### \*Diskussion\*

Perspektivisch können die entwickelten Instrumente in einem Large-Scale-Assessment beruflicher Bildung einmünden. Konzeptionelle Vorarbeiten hierzu wurden in einer Machbarkeitsstudie bereitgestellt (Beathge & Arends, 2009). Denkbar wäre es ebenso, den Papier-und-Bleistift-Test in Prüfungskontexten einzusetzen (z. B. anknüpfend an Abele, 2014).

### **Einflussfaktoren auf berufliches Fachwissen bei Elektronikern für Automatisierungstechnik am Ende der Ausbildung**

**Matthias Hedrich**  
Universität Stuttgart

Bezogen auf Untersuchungen zu (schulischen, betrieblichen und privaten) Qualitätsmerkmalen bzw. Einflussfaktoren auf Fachwissen im (elektrotechnischen) Bereich der beruflichen Bildung sind u.a. zwei zentrale Instrumente zu benennen. (1) das „Mannheimer Inventar zur Erfassung betrieblicher Ausgangssituationen“ (MIZEBA) (Zimmermann, Müller & Wild, 1994) bzw.

seine Überarbeitung für den gewerblich-technischen Bereich (Wosnitza & Eugster, 2001) sowie das (2) „Inventar zur betrieblichen Ausbildungsqualität“ (IBAQ) (Velten & Schnitzler, 2012).

Fokussiert das stark theoriegeleitet entwickelte Instrument MIZEBA ausschließlich den betrieblichen Bereich, so werden beim IBAQ, dessen Items vor allem aus Recherchearbeiten zu anderen Studien entstammen, auch schulische Qualitätsmerkmale erfasst. Dies wird unter Einbezug der intrinsischen Motivationsvariante geleistet.

Trotz der ausdifferenzierten und vor allem beim MIZEBA theoriegeleiteten Entwicklung der Instrumente konnten weder mit dem MIZEBA noch mit dem IBAQ die erhofften Einflüsse auf Fachwissen identifiziert werden.

Die Untersuchung folgt der Fragestellung, welche (Qualitäts-)Merkmale des schulischen, betrieblichen und privaten Bereiches noch Einfluss auf das Fachwissen von Elektronikern für Automatisierungstechnik am Ende der Ausbildung nehmen. Dabei widmet sich das Erkenntnisinteresse der Studie aus quantitativ-empirischer Sicht einer bis dato kaum beachteten Berufsgruppe der elektrotechnischen Ausbildung, in der bisher vor allem Untersuchungen zum Fachwissen oder zur fachspezifischen Problemlösefähigkeit bei Elektronikern für Energie- und Gebäudetechnik (vgl. Nickolaus, Geißel, Abele & Nitzschke, 2011), bei Kfz Mechatronikern (vgl. Nickolaus, Abele, Gschwendtner, Nitzschke & Greiff, 2012) oder in der elektrotechnischen Grundbildung im Blickpunkt standen (vgl. Gschwendtner, Geißel & Nickolaus, 2007; Nickolaus, Gschwendtner & Geißel, 2008; Gschwendtner, Geißel & Nickolaus, 2010). Zudem versucht die Untersuchung durch den Einbezug neuer Instrumente die Lücke zu schließen, die durch wenig aussagekräftige Ergebnisse vorangegangener Instrumente entstanden ist (vgl. dazu die Anmerkungen beim Theoriebezug, weiter oben in dieser Tabelle). Die Relevanz der Untersuchung legitimiert sich aus der Notwendigkeit Qualitätsmerkmale im Sinne von möglichen Stellschrauben aufzudecken, die einen positiven Einfluss auf das Ergebnis des Lehr-Lernprozesses im elektrotechnischen Bereich der beruflichen Bildung nehmen.

Die Untersuchung fokussiert den Beruf Elektroniker für Automatisierungstechnik am Ende der Ausbildung. Die Stichprobe ist eine Teilstichprobe aus dem Verbundprojekt ASCOT-KoKo EA dar, die sich wie folgt zusammensetzt: N = 167, weiblich: 10; männlich: 157; Durchschnittsalter: 20.6 (SD = 3.14). Die Probanden entstammen drei Bundesländern: Baden-Württemberg (NBW = 74), Bayern (NBY = 62) und Nordrhein-Westfalen (NNW = 31). Bei allen Probanden wurde erhoben: das Fachwissen mit einem im ASCOT-Teilprojekt KoKoEA entwickelten Test, kognitive Grundfähigkeiten mit Teil 1 des Instrumentariums CFT 20-R (Weiß, 2006), allgemeine Basiskompetenzen in den Bereichen Mathematik, Naturwissenschaft und Lesen (Bernhardt et al., 2013) sowie Einflussfaktoren aus Schule, Betrieb & privatem Umfeld mit dem ebenfalls innerhalb des ASCOT-Verbundprojektes entwickelten Instrument SiKoFak (Baethge & Baethge-Kinsky, 2013). Vor allem letztgenanntes Instrument SiKoFak bietet dabei die Möglichkeit, deutlich über die bisher mit MIZEBA und IBAQ erfassten Merkmale hinausgehend schulische, betriebliche und private Einflussfaktoren zu erfassen.

Die Ergebnisse gliedern sich in zwei Bereiche:

(1.) Mittels regressionsanalytischen Untersuchungen wird der Einfluss der erhobenen Qualitätsmerkmale auf die Ausprägung des Fachwissens (1D-WLE-Personenfähigkeitsschätzer) gezeigt (vgl. dazu auch Hedrich, 2015). (2.) In einem zweiten Bereich sollen Strukturgleichungsmodelle vorgestellt werden, die den Einfluss der Qualitätsmerkmale in einem Gesamtzusammenhang beschreiben.

## G12

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H119

### **Blickverhalten in der Lernforschung: Eye-Tracking Indikatoren für kognitive und affektive Prozesse**

Chair(s): **Babette Park** (Universität des Saarlandes, Deutschland), **Klaus-Peter Wild** (Universität Regensburg)

DiskutantIn(nen): **Alexander Eitel** (Universität Freiburg)

Blickverhaltensdaten werden schon seit mehreren Jahren auch in der Lernforschung zur Untersuchung von Lernprozessen genutzt. Es gibt mittlerweile zahlreiche forschungsmethodische Veröffentlichungen mit technischen Anleitungen zur Verwendung des Eye-Tracking-Messverfahrens und Diskussionen über die psychometrische Genauigkeit dieses Online-Messinstruments (z.B. Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka & van de Weijer, 2011). Darüber hinaus existieren aktuell zusammenfassende Überblicksartikel zu ausstehenden Forschungsfragen (z.B. Jarodzka, Holmqvist & Gruber, 2017). Die bisher überwiegend kognitive Perspektive auf die Nutzung von Blickverhaltensdaten vernachlässigt jedoch das Potential der Blickbewegungsmessung, mit unterschiedlichen Eye-Tracking-Indikatoren auch affektive Prozesse erfassen zu können. Das vorliegende Symposium wird daher neben der „cold cognition“-Perspektive auch eine affektive und motivationale Perspektive auf die Nutzung von Eye-Tracking-Daten in der Lernforschung aufgreifen. In drei Beiträgen werden sowohl motivationale als auch kognitive Prozesse beim Lernen unter Verwendung der Eye-Tracking-Methode untersucht. In einem ersten Beitrag von Park et al. (Universität des Saarlandes) werden verschiedene Blickbewegungsindikatoren zur Analyse des positiven Lerneffekts durch Hindeuten (Pointing) auf bzw. Nachzeichnen (Tracing) des Lernmaterials untersucht, um verschiedene Erklärungsansätze des Tracing-Effekts zu überprüfen. Die Analyse des Blickverhaltens zeigt, dass die Pointing-Gruppe im Vergleich mit der Tracing-Gruppe ihre visuelle Aufmerksamkeit früher auf die relevanten Strukturen der Abbildungen im Lernmaterial richtete. Ein weiterer Beitrag von Zander et al. (Bauhaus-Universität Weimar) beschäftigt sich mit der Wirksamkeit persönlicher Ansprache in neutralen und potentiell aversiven Lernmaterialien. Ein Ergebnis der Studie ist es, dass fokussierte und vermeidende Aufmerksamkeitsprozesse möglicherweise in besonderem Maße in Bezug auf Bildmaterialien auftreten. Die Studie zeigt insgesamt, dass Lerninhalte und deren emotionaler Gehalt eine der Grenzbedingungen für die Wirksamkeit des Personalisierungsprinzips sein können. Ein letzter Beitrag von Wild und Viehauser (Universität Regensburg) verfolgt eine eher motivationspsychologisch orientierte Fragestellung. Die vorgestellte Studie verweist auf negative Folgen, wenn motivational förderlich gemeinte spezifische Instruktionen zu einer instruktional unerwünschten Verschiebung kognitiver Ressourcen führen. Die Blickdaten werden hier auf den an der Zeichenzahl relativierten Fixationszeiten der als relevant vs. nicht-relevant deklarierten Textabschnitte in den Texten analysiert und die Ergebnisse in einem motivationspsychologischen Kontext kritisch diskutiert. Alle drei Beiträge präsentieren verschiedene Eye-Tracking-Indikatoren zur Prozessanalyse des Lernverhaltens, welche gerade im Quervergleich und im Überblick die notwendige Differenzierung in Abhängigkeit vom jeweiligen forschungsmethodischen und psychometrischen Ziel erkennen lassen. Im Rahmen einer abschließenden Diskussion durch Alexander Eitel (Universität Freiburg) werden die beiden angerissenen kognitiven und affektiven Perspektiven auf die Nutzung der Blickverhaltensmessung sowie die dahinter stehenden Konstrukte auch im Hinblick auf die Unterrichts- und Lehrpraxis diskutiert.

*Beiträge des Symposiums*

### **Haptik und Lernen: Eine Eye-Tracking Studie zur Prüfung verschiedener Erklärungsansätze des Effekts von Pointing- und Tracing-Gesten**

**Babette Park<sup>1</sup>, Andreas Korbach<sup>1</sup>, Paul Ginns<sup>2</sup>, Roland Brünken<sup>1</sup>**

<sup>1</sup>Universität des Saarlandes, <sup>2</sup>The University of Sydney

Die zu berichtende Studie wurde als Folgestudie konzipiert um den positiven Effekt durch die Verwendung der Finger beim Lernen genauer zu bestimmen (Park, Korbach, Ginns & Brünken, 2017). Beide Studien wurden in Anlehnung an die Studien von Hu, Ginns und Bobis (2015), sowie von Macken und Ginns (2014) entworfen und verwenden Pointing- und Tracing-Gesten die mit den Zeigefingern ausgeführt werden. Grundsätzlich geht es bei den instruierten Fingerbewegungen darum, bezugnehmende Text- und Bildinformation mit den Fingern zu markieren. Pointing bezeichnet dabei das Ablegen der Finger in unmittelbarer Nähe der bezugnehmenden Informationen und Tracing bezeichnet das zusätzliche, aktive Nachfahren der betreffenden Strukturen in der Abbildung. Die positiven Effekte durch die Verwendung der Finger beim Lernen können unter anderem mit Hilfe der Cognitive Load Theorie (Sweller, Ayres, & Kalyuga, 2011) und Theorien zu Embodied Cognition erklärt werden (Glenberg, Witt, & Metcalfe, 2013). Dabei wird angenommen, dass durch die Nutzung des haptischen Kanals und die Einbindung einer taktilen und kinästhetischen Modalität zusätzliche Ressourcen des Arbeitsgedächtnisses für den Lernprozess genutzt werden können (Kaas, Stoeckel, & Goebel, 2008). Fingerbewegungen, als Form biologischen Primärwissens, unterstützen demnach den Erwerb biologischen Sekundärwissens (Hu et al., 2015). Eine alternative Erklärung für den positiven Effekt der Verwendung der Finger beim Lernen bietet die Annahme, dass durch den Einsatz der Finger die Aufmerksamkeit der Lerner gezielt auf wichtige Informationen gelenkt wird (Cosman & Vecera, 2010), was zu einer längeren und tieferen Verarbeitung der Informationen führt (Reed, Grubb, & Steele, 2006). Park et al. (2017) konnten einen positiven Effekt für die Verwendung der Finger beim Lernen zeigen, der sich sehr gut durch eine Veränderung im Fokus der visuellen Aufmerksamkeit und eine erhöhte kognitive Aktivität für die Integration von Text und Bildinformationen erklären lässt. Um den potentiellen Einfluss der Bewegung genauer zu untersuchen, wurden in der aktuellen Studie die Verwendung von Pointing- und Tracing-Gesten getrennt instruiert. Eine Gruppe wurde angewiesen ihre Finger für Tracing-Gesten zu verwenden und relevante Strukturen der Abbildungen aktiv nachzufahren, wohingegen die andere Gruppe angewiesen wurde ihre Finger für Pointing-Gesten zu verwenden und die Finger lediglich in unmittelbarer Nähe der relevanten Text- und Bildinformationen abzulegen.

Für die aktuelle (noch laufende) Studie wurde eine ins Deutsche übersetzte Version des Lernmaterials und der Lerntests aus der Studie von Macken und Ginns (2014) verwendet und für die Präsentation auf einem 1920 x 1080 Touchscreen Monitor angepasst, um während der Lernphase die Augen- und Fingerbewegungen der Lernenden aufzeichnen zu können. Die kognitive Belastung wurde anhand einer Rating Skala erhoben (Leppink, Paas, Van Gog, Van der Vleuten, & Van Merriënboer, 2014). Die Versuchspersonen (N=27) wurden zufällig auf zwei Gruppen verteilt, wobei eine Gruppe instruiert wurde Tracing-Gesten auszuführen und eine Gruppe instruiert wurde nur Pointing-Gesten zu verwenden.

Die vorläufigen Ergebnisse zeigen in Bezug auf den Lernerfolg einen positiven Effekt für die Pointing-Gruppe (M = 11.31, SD = 2.53) im Vergleich zur Tracing-Gruppe (M = 7.86, SD = 19.30),  $F(1,25) = 5.893$ ,  $p = .023$ ,  $\eta^2 = .191$ . Darüber hinaus zeigt die Tracing-Gruppe eine insgesamt höhere kognitive Belastung (M = 5.44, SD = 2.22) im Vergleich zur Pointing-Gruppe (M = 3.87, SD = 1.61),  $F(1,25) = 4.387$ ,  $p = .047$ ,  $\eta^2 = .149$ . Die Analyse des Blickverhaltens zeigt, dass die Pointing-Gruppe ihre visuelle

Aufmerksamkeit früher auf die relevanten Strukturen der Abbildung richtete ( $M = 6.19$ ,  $SD = 3.40$ ) als die Tracing-Gruppe ( $M = 10.27$ ,  $SD = 6.09$ ),  $F(1,25) = 4.531$ ,  $p = .043$ ,  $\eta^2 = .153$ .

Insgesamt bestätigen diese Ergebnisse die Annahme, dass der positive Effekt von Pointing- und Tracing-Gesten eher auf Veränderungen im Fokus der visuellen Aufmerksamkeit zurückgeführt werden kann, als durch den Ansatz der Embodied Cognition Theorie.

## **Persönliche Ansprache? – Eyetrackingstudien zur Wirksamkeit persönlicher Ansprache in neutralen und potentiell aversiven Lernmaterialien**

**Steffi Zander<sup>1</sup>, Stefanie Wetzel<sup>2</sup>, Tim Kühl<sup>3</sup>, Sven Bertel<sup>2</sup>**

<sup>1</sup>Bauhaus-Universität Weimar, <sup>2</sup>Hochschule Flensburg, <sup>3</sup>Universität Mannheim

Eines der gut etablierten und empirisch beforschten Prinzipien zur Gestaltung medial vermittelter Lerninhalte ist das Personalisierungsprinzip (Mayer, 2014). Es besagt, dass eine persönliche Ansprache Lernender, zum Beispiel durch die Verwendung eines persönlichen Sprachstils mit Personal- und Possesivpronomen, im Vergleich zu einem formalen Sprachstil (z.B. „Du“ und „Dein“ statt „man“) zu höherem Lernerfolg führt und die kognitive Belastung senkt. Dies wurde in einer Reihe von Studien mit „neutralen“ Lerninhalten wie Botanik, Wetter oder Astronomie belegt (Ginns et al., 2013). Dennoch ist weitgehend ungeklärt ob die Wirksamkeit persönlicher Ansprache für verschiedene Lerninhalte gleichermaßen gilt.

Studien von Kühl & Zander (2017) haben gezeigt, dass die Wirksamkeit persönlicher Ansprache sich für potentiell aversive Inhalte (z. B. medizinische Inhalte zu unkontrollierbaren, gefährlichen Krankheiten) hinsichtlich Lernerfolg umdreht, und zu tendenziell höherer kognitiver Belastung führt. Die Prozesse, welche dem (umgekehrten) Personalisierungseffekt zu Grunde liegen sind noch ungeklärt. Ziel der hier präsentierten Studien ist es, anhand von Blickbewegungsmessung, zu untersuchen, ob unterschiedliche Aufmerksamkeitsprozesse bei der Verarbeitung von „neutralem“ und „aversivem“ / „personalisiertem“ und „formalem“ Lernmaterial diese Prozesse erklären können.

Angenommen wurde, dass bei der Verarbeitung von personalisiertem/neutralem Material Fixationsdauer, Lesetiefe und Transitionen zwischen Text und Bild auf eine intensivere, fokussiertere Auseinandersetzung mit dem Lernmaterial hinweisen. Dies sollte sich auch in höherem Lernerfolg und geringerer mentaler Belastung widerspiegeln.

Für potentiell aversives Lernmaterial wurde umgekehrt angenommen, dass in der personalisierten/ aversiven Variante ein vermeidendes Blickverhalten mit geringerer Fixationsdauer, Lesetiefe und weniger Transitionen zwischen Text und Bild auftreten sollte. In zwei Studien ( $N=37$ ,  $N=38$ ), die Lernmaterial zum Thema Wetter und zur Entstehung von Hirnblutungen verwendeten, wurden die Hypothesen untersucht.

Die Ergebnisse lassen darauf schließen, dass unterschiedliche Aufmerksamkeitsprozesse während des Lernens mit personalisiertem/neutralem und personalisiertem/aversivem Lernmaterial stattfinden. Für Lernerfolg ( $Mdnform = 16.11$ ,  $Mdnpers = 21.74$ ,  $U = -1.654$ ,  $p = .098^*$ ,  $r = -.271$ ), Motivation ( $Mdnform = 17.39$ ,  $Mdnpers = 20.53$ ,  $U = -.883$ ,  $p = .377$ ,  $r = -.145$ ) und mentale Belastung ( $Mdnform = 20.14$ ,  $Mdnpers = 17.92$ ,  $U = -.625$ ,  $p = .532$ ,  $r = -.103$ ) zeigen die Ergebnisse die angenommenen Trends mit einem Vorteil personalisierter Ansprache, in neutralem Lernmaterial. Die entsprechenden Blickbewegungen zeigen niedrigere Fixationszeiten auf Text ( $Mdnform = 176.01$ ,  $Mdnpers = 158.44$ ,  $U = -1.66$ ,  $p = .099^*$ ,  $r = -.27$ ), eine höhere Lesetiefe auf dem Bildmaterial ( $Mdnform = 24.11$ ,  $Mdnpers = 31.37$ ,  $U = -2.076$ ,  $p = .038^*$ ,  $r = -.34$ ) und eine deskriptiv höhere Anzahl an Transitionen zwischen Text und Bild ( $Mdnform = 0.102$ ,  $Mdnpers = 0.118$ ,  $U = -.936$ ,  $p = .361$ ,  $r = -.15$ ), wenn persönliche Ansprache verwendet wurde.

Für potentiell aversiven Lerninhalt, zeigen sich umgekehrte Tendenzen, mit weniger Lernerfolg ( $Mform = 7.82$ ,  $Mpers = 7.11$ ,  $F(1,35) = .065$ ,  $p = .08^*$ ,  $d = .08$ ), und deskriptiv höherer Belastung ( $Mform = 4.06$ ,  $Mpers = 4.17$ ,  $F(1,35) = .068$ ,  $p = .795$ ,  $d = .088$ ) unter der personalisierten Bedingung. Ähnlich wie in Experiment 1 sind die Unterschiede hauptsächlich für das Bildmaterial in der erwarteten Richtung: Fixationszeiten ( $Mform = 80.87$ ,  $Mpers = 66.32$ ,  $F(1,35) = 8.323$ ,  $p = .007^*$ ,  $d = .98$ ), Lesetiefe ( $Mform = 91.19$ ,  $Mpers = 69.05$ ,  $F(1,35) = 3.972$ ,  $p = .054^*$ ,  $d = .68$ ) sowie Transitionen zwischen Text und Bild ( $Mform = 0.18$ ,  $Mpers = 0.14$ ,  $F(1,35) = 3.646$ ,  $p = .064^*$ ,  $d = .63$ ) unter der persönlichen Ansprache sind geringer, als unter der formalen Bedingung.

Die Ergebnismuster legen nahe, dass personalisierte Ansprache nicht für alle Lerninhalte geeignet ist. Entgegengesetzt zu den Hypothesen finden sich Unterschiede hinsichtlich der Aufmerksamkeitsprozesse hauptsächlich im Bezug auf die Bildmaterialien, aber nicht auf den Texterläuterungen. Mögliche Ursachen werden diskutiert.

## **Instrumentalitätserwartungen in der Lernregulation. Selektive Fixationszeiten aus Blickbewegungsdaten als Indikatoren einer motivationalen Fehlsteuerung?**

**Klaus-Peter Wild, Sarbina Viehauser**

Universität Regensburg

Im Zentrum zahlreicher Untersuchungen zum selbstgesteuerten Lernen steht die Frage nach der motivationalen Steuerung des Lernverhaltens (Wild, 2010). In empirischen Studien stützt sich sehr häufig sowohl die Erfassung der motivationalen Bedingungen wie auch die Erfassung des kognitiven Lernverhaltens auf Selbstberichtsverfahren. Seit einigen Jahren werden zunehmend vor allem zur Operationalisierung von kognitiven Prozessen, zuweilen aber auch zur Erfassung von motivationalen Konstrukten, Messverfahren und Indikatoren verwendet, die mit Eye-Tracking-Systemen gewonnen werden (Holmqvist, 2011). Durch die Verwendung von Eye-Tracking Daten in einem experimentellen Lernumfeld ergibt sich die Möglichkeit, das individuelle Lernverhalten durch Indikatoren zur selektiven Allokation von Lernzeiten spezifischer erfassen zu können. Gleichzeitig bietet sich damit auch eine innovative Form zur Untersuchung motivationaler Steuerungsprozesse.

Auf dem Hintergrund von Erwartungs-Wert-Modellen und im Einklang mit zahlreichen Befunden wird in der vorliegenden Studie von einer motivational förderlichen Wirkung von Nützlichkeits- und Kostenkalkulationen in der Lernregulation ausgegangen (z.B. Wigfield & Eccles, 2000; Canning & Harackiewicz, 2015). Nützlichkeits- und Kostenkalkulationen ergeben sich u.a. durch eine explizite Benennung von Lern- und Prüfungsthemen in einem Textkorpus. Kosten ergeben sich aus dem Umfang der mit hoher Aufmerksamkeit gelesenen Passagen eines Textes. Aus der Sicht der Lernenden sollte sich dadurch im Kontrast zu den nicht als relevant benannten Themenbereichen eine höhere Instrumentalität und damit mittelbar eine erhöhte Lernmotivation ergeben. Im Fokus dieser Studie steht nun die Frage, ob sich durch eine verstärkte Fokussierung auf „relevante“ Teilgebiete gleichzeitig eine unerwünschte schwächere Fokussierung auf die „irrelevanten“ Gebiete ergibt. Hypothesenkonforme Befunde wären zudem ein wichtiger Hinweis auf die Validität von Blickbewegungsindikatoren zur Untersuchung motivationaler Prozesse in Lernsituationen.



In einer experimentellen Laborstudie bearbeiteten 93 Studierende ein textbasiertes Lernprogramm. Sie hatten in einer ersten Trainingsphase die Aufgabe, aus einem Lehrtext die wesentlichen Informationen für eine spätere Arbeitsphase zu entnehmen. Der vorgelegte Text war zunächst ohne eine themenspezifische Vorinstruktion insgesamt zu lesen. Eine der beiden experimentellen Variationen betraf die daran anschließende selektive Information darüber, welche Teilbereiche für das weitere Lernprogramm „wirklich“ von Bedeutung seien. Diese Information betraf aus Plausibilitätsgründen jeweils die Mehrzahl der vorgelegten Informationen, lies aber jeweils die „Lücke“ einer nicht explizit als relevant deklarierten Informationseinheit. Abschließend war der gleiche Text noch einmal zu lesen.

Die zweite experimentelle Variation betraf die Untersuchung eines möglichen Reihenfolgeeffektes, der sich durch die zweifache Durchführung dieser Instruktionsabfolge mit einem neuen Text ergeben könnte. Es wurde angenommen, dass sich die Instrumentalitätseinschätzungen und damit auch die Motivation der Studierenden bei einem zweiten Durchlauf abweichend vom ersten Durchlauf auswirken müssten, da die Probanden aufgrund des ersten Durchlaufs nun bereits korrekt vermuten konnten, dass sie nach dem ersten Lesen die besonders relevanten Abschnitte des Textes explizit benannt erhalten. Für die jeweils zweite Lese-phase wird daher ein abgeschwächter Effekt der Instrumentalitätswahrnehmung erwartet.

Die Auswertung der Blickdaten beruht auf den an der Zeichenzahl relativierten Fixationszeiten der als relevant vs. nicht-relevant deklarierten Textabschnitte in beiden Texten. Die zentralen Analysen basieren aufgrund der genesteten Datenstruktur entweder auf varianzanalytischen oder auf mehrebenenanalytischen Verfahren (HLM, R). Die Befunde zeigen zunächst erwartungsgemäß, dass die summativen Fixationszeiten zwischen dem jeweiligen ersten und zweiten Lesedurchgang insgesamt drastisch auf weniger als die Hälfte der Lesezeiten absinken. Dabei sind die Lesezeiten zwischen dem ersten und zweiten Lesedurchgang mittelhoch und signifikant korreliert. Dies verweist darauf, dass in der Regel kein kompensatorischer Leseaufwand vorliegt. Die signifikanten Wechselwirkungen zwischen Lesedurchgang und der Relevanzdeklaration verweisen zudem darauf, dass für die „nicht-deklarierten“ Lernbereiche dieser Abfall der Lesezeiten nochmals stärker ausgeprägt ist. Dies verweist hypothesenkonform darauf, dass auf die weniger instrumentellen Textpassagen eine besonders reduzierte Aufmerksamkeit entfällt.

Die Befunde und Grenzen der Studie werden in einem motivationspsychologischen Kontext und im Hinblick auf die Unterrichts- und Lehrpraxis kritisch diskutiert.

## G13

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H118

### Unterrichtsqualität: Die offene Frage nach den Perspektiven

Chair(s): **Benjamin Fauth** (Universität Tübingen, Deutschland), **Anna-Katharina Praetorius** (Universität Zürich)

DiskutantIn(nen): **Kurt Reusser** (Universität Zürich)

Die Qualität des Unterrichts ist eine der wichtigsten Determinanten des Lernerfolgs von Schülerinnen und Schülern. Entsprechend ist das professionelle Handeln von Lehrpersonen ein hochrelevantes Forschungsfeld. Dieses Forschungsfeld steht aktuell jedoch vor einer Reihe theoretischer und methodischer Herausforderungen. Dazu gehört auch, dass bislang nur unzureichend geklärt ist, wie die Qualität von Unterricht optimal erfasst werden kann. In unterschiedlichen empirischen Studien wurden immer wieder verhältnismäßig geringe Zusammenhänge zwischen Urteilen zur Unterrichtsqualität aus unterschiedlichen Perspektiven gefunden – meist Urteile von Schüler/innen, Lehrkräften und externen Beobachter/innen (Clausen, 2002; Fauth et al., 2014; Kunter & Baumert, 2006; Wagner et al., 2015; Wettstein et al., 2016). Diese Befunde werden oft mit einer perspektivenspezifischen Validität der einzelnen Maße erklärt.

Eine solche Erklärung greift jedoch erstens nur dann, wenn empirisch auch hinreichend nachgewiesen werden kann, dass es sich bei den Abweichungen nicht um Messproblematiken und Einschränkungen der Validität handelt. Zweitens bleibt auch bei einer solchen Einordnung unklar, welche Schlussfolgerungen nun konkret für die Erfassung von Unterrichtsqualität gezogen werden können und welche Perspektive entsprechend für welche Aspekte von Unterrichtsqualität optimalerweise herangezogen werden sollte. Mit dieser Herausforderung beschäftigt sich das Symposium aus theoretischer und empirischer Sicht.

In dem Symposium geht es schwerpunktmäßig um die Klassenführung als einer von drei Basisdimensionen von Unterrichtsqualität. Diese Dimension eignet sich besonders gut als Ausgangspunkt einer Diskussion zur Perspektivenspezifität, da hier die Annahme eines gemeinsamen true scores über verschiedene Perspektiven hinweg plausibel erscheint, so dass Abweichungen eindeutiger interpretiert werden können. Die Erklärung dieser Abweichungen steht im Fokus des Symposiums: Wie kommen Unterschiede zwischen den Perspektiven zustande?

Das Symposium sucht Antworten auf diese Frage in einem theoretischen Beitrag und zwei empirischen Beiträgen. Im Beitrag von Fauth et al. wird zunächst gezeigt, dass in gängigen Fragebogen- und Beobachtungsinstrumenten in unterschiedlichem Maß nach dem Verhalten von Schüler/innen, Lehrkräften oder Lehrer-Schüler-Interaktionen gefragt wird (unterschiedliche Referenten). Im Rückgriff auf Modelle aus der Persönlichkeitspsychologie werden Asymmetrien zwischen Selbst- und Fremdbeurteilung beschrieben, die in Kombination mit den unterschiedlichen Itemreferenten zu Unterschieden in der Bearbeitung von Items führen können.

Die beiden empirischen Beiträge untersuchen zentrale Annahmen, die im ersten Beitrag zur Erklärung von perspektivenspezifischen Urteilen gemacht wurden. So zeigen Göllner et al., dass es bei der Erfassung von Klassenführung durch Schülerurteile einen Unterschied in Bezug auf die prädiktive Validität macht, ob nach dem Verhalten der Lehrkraft (Monitoring) oder der Schüler/innen (Störungen) gefragt wird. Zudem zeigen Monitoring und Störungen unterschiedliche Zusammenhänge mit der Klassenkomposition. Praetorius et al. vergleichen Items, die eine für Lehrkräfte relevante Bewertungskomponente (ego involvement) beinhalten mit Items ohne eine solche Komponente. Es zeigt sich hypothesenkonform, dass die Zusammenhänge zwischen Lehrer- und Schülerurteilen deutlich geringer ausfallen, wenn Items eine solche Komponente beinhalten. Allerdings können interindividuelle Unterschiede in der Höhe der Übereinstimmung nicht durch Lehrermerkmale (z.B. soziale Erwünschtheit und Selbstwirksamkeit) erklärt werden.

#### Beiträge des Symposiums

### Warum (und inwiefern) ist Unterrichtsqualität eine Frage der Perspektive? Auf dem Weg zu einem besseren Verständnis perspektivenspezifischer Unterrichtsbeurteilungen

**Benjamin Fauth<sup>1</sup>, Richard Göllner<sup>1</sup>, Gerlinde Lenske<sup>2</sup>, Anna-Katharina Praetorius<sup>3</sup>, Wolfgang Wagner<sup>1</sup>**

<sup>1</sup>Universität Tübingen, <sup>2</sup>Universität Koblenz-Landau, <sup>3</sup>Universität Zürich

In dem Beitrag sollen Überlegungen zu einem theoretischen Modell angestellt werden, das perspektivenspezifische Mechanismen bei der Bearbeitung von unterschiedlichen Fragebogen- und Beobachtungsitens zur Erfassung von Unterrichtsqualität beschreibt. Es soll helfen die verhältnismäßig geringen Übereinstimmungen zwischen den unterschiedlichen Perspektiven (z.B. Clausen, 2002; Fauth et al., 2014; Kunter & Baumert, 2006) zu erklären: Wie kommen perspektivenspezifische Unterschiede bei der Beurteilung von Unterrichtsqualität zustande?

Ausgangspunkt und erster Baustein des Modells sind Überlegungen aus der Persönlichkeitspsychologie, wo grundlegende Asymmetrien zwischen Selbst- und Fremdbeurteilung beschrieben werden (z.B. Funder, 1995; Spain et al., 2000), die zu unterschiedlichen Einschätzungen aus unterschiedlichen Perspektiven führen. In Vazires (2010) SOKA-Modell (Self-Other Knowledge Asymmetry) werden zwei Asymmetrien unterschieden:

(1) „Informational difference“: Unterschiedlichen Urteilern stehen unterschiedliche Informationen zur Verfügung und Informationen sind unterschiedlich salient (vgl. Actor-Observer-Unterschiede).

(2) „Motivational significance“: Items sind in unterschiedlichem Maße für den Selbstwert der sich selbst einschätzenden Person bedeutsam (ego-involvement).

Diese Überlegungen lassen sich auf die Beurteilung von Unterricht anwenden: (1) Schüler/innen, Lehrkräften und Beobachter/innen stehen unterschiedliche Informationen zur Verfügung, z.B. werden Schüler/innen vermutlich Probleme haben, ein Item wie „In Mathematik weiß unsere Lehrerin immer genau, was in der Klasse vor sich geht“ einzuschätzen. (2) Für das professionelle Selbstverständnis einer Lehrkraft können unterschiedliche Unterrichtsaspekte unterschiedlich relevant sein. Beispielsweise wird es problematischer sein anzugeben, dass man keinen „guten Überblick“ über das habe, „was in der Klasse geschieht“ als anzugeben, dass es im Unterricht „viele Störungen“ durch die Schüler/innen gibt.

Der zweite Baustein des Modells ergibt sich aus den Ergebnissen eines Literaturüberblicks zu in der Unterrichtsforschung gängigen Fragebogen und Beobachtungsitens. Hier zeigt sich, dass die Items in unterschiedlichem Maß nach dem Verhalten von Lehrkräften, Schüler/innen oder aber dem Unterrichtsgeschehen als Schüler-Lehrer-Interaktion fragen. Das heißt, die Items haben ihren Fokus auf unterschiedlichen Referenten und erfassen somit auch unterschiedliche Aspekte des

Unterrichtsprozesses (Beispiele: „Dem Lehrer gelingt es, Störungen schnell zu beenden“, Thiel et al., 2013; „Student behavior in this class is a problem“, Wallace et al., 2016; „In Mathematik wird im Unterricht viel Zeit vertrödelte“, Baumert et al., 2008).

Welchen Referenten ein Item fokussiert, hat bedeutsame Konsequenzen: Bei der Bearbeitung durch Schüler oder Lehrkräfte entscheidet der Referent eines Items darüber, ob es sich bei der Einschätzung um eine Selbst- oder eine Fremdbeurteilung handelt. Welche Informationen einem Urteiler zur Verfügung stehen, aber auch wie bedeutsam ein Item für den Selbstwert eines Urteilenden ist, hängt vom jeweils referenzierten Aspekt des Unterrichtsgeschehens ab. Die oben beschriebenen Asymmetrien wirken also referenten- und perspektivenspezifisch.

Auf der Grundlage dieser Beobachtungen und Überlegungen wird in dem Vortrag ein Raster aus Beurteilerperspektive und Itemreferent entwickelt. Mit diesem Raster lassen sich bereits vorliegende Forschungsergebnisse zum Vergleich unterschiedlicher Perspektiven kritisch aufarbeiten. Dabei zeigt sich erstens, dass Fragebogenitems mit unterschiedlichen Referenten in empirischen Analysen unterschiedlich funktionieren (z.B. Wallace et al., 2016; Wagner et al., 2016; Wettstein et al., 2016). Zweitens zeigt sich, dass das Raster eine hohe Erklärungskraft für in früheren Studien gefundene bzw. nicht-gefundene Zusammenhänge zwischen unterschiedlichen Perspektiven aufweist.

Auf der Grundlage dieser Befunde lassen sich nicht nur die in der Literatur berichteten Zusammenhänge besser verstehen, sondern es lassen sich auch Vorhersagen über zu erwartende Zusammenhangsmuster bei der multiperspektivischen Erfassung von Unterrichtsqualität machen. Aus den im Vortrag präsentierten Befunden lassen sich somit auch Empfehlungen für den Einsatz von Erhebungsinstrumenten in künftigen Studien ableiten.

### **Die Erfassung von Klassenführungsmerkmalen aus Schülersicht: Ein Vergleich von Äpfeln und Birnen?**

**Richard Göllner<sup>1</sup>, Wolfgang Wagner<sup>1</sup>, Gerlinde Lenske<sup>2</sup>, Anna-Katharina Praetorius<sup>3</sup>, Benjamin Fauth<sup>1</sup>**  
<sup>1</sup>Universität Tübingen, <sup>2</sup>Universität Koblenz-Landau, <sup>3</sup>Universität Zürich

Die Erfassung der Klassenführung einer Lehrkraft erfolgt im Rahmen von Large-Scale-Assessments häufig anhand von Schülerbeurteilungen des Unterrichts (z.B. Fraser, 1991). In vielen Studien werden die Schülerinnen und Schüler einer Klasse dabei zu unterschiedlichen Aspekten von Klassenführung befragt, wie etwa dem Monitoring, der Disziplin, der Regelklarheit, oder der Störungsfreiheit (Kunter et al., 2013; Wallace, Kelcey, & Ruzek, 2016). Allerdings verweisen bereits die Konstruktbezeichnungen in gängigen Fragebogeninstrumenten darauf, dass sich die verwendeten Items nicht nur im Hinblick auf den jeweiligen Klassenführungsaspekt unterscheiden, sondern den Beurteilungsgegenstand in unterschiedlicher Weise fokussieren. So beziehen sich einige der bekannten Skalen unmittelbar auf die Lehrkraft (z.B. Monitoring: „In Mathematik achtet unser Lehrer/ unsere Lehrerin sehr darauf, dass wir aufpassen.“), während andere Skalen keinen expliziten Referenten nennen oder sich eher auf die Schülerinnen und Schüler einer Klasse beziehen (z.B. Störungen oder Disziplinprobleme: „In Mathematik wird der Unterricht oft sehr gestört.“; z.B. Prenzel et al., 2006). Inwiefern Skalen ohne eine explizite Referenz zur Lehrkraft nicht nur die durch die Lehrkraft realisierte Klassenführung abbilden, sondern auch lehrkraftunabhängige Schülermerkmale widerspiegeln (z.B. Unterschiede der „Störungspotenziale“ verschiedener Schulklassen), ist Gegenstand der vorliegenden Studie.

Hierbei wurden basierend auf Daten von PISA 2003 (nationale Ergänzungsstudie; N = 5.628 Schülerinnen und Schüler aus 259 Mathematikklassen der 10. Jahrgangsstufe) die Klassenführungsmerkmale Monitoring, Störungen, sowie die Zeitverschwendung im Mathematikunterricht untersucht. Die Ergebnisse eines faktorenanalytischen Zweiebenenmodells zeigten auf der Klassenebene zunächst, dass der Zusammenhang zwischen dem Faktor Störungen und Zeitverschwendung äußerst hoch ausfiel ( $r = .97$ ), so dass beide Faktoren zusammengefasst wurden. Weiterführende Analysen auf der Grundlage eines nested factor-Modells mit einem Generalfaktor über sämtliche Indikatoren sowie einem spezifischen Faktor für Störungen/Zeitverschwendung zeigten, dass auf der Klassenebene ausschließlich der Faktor Störungen/Zeitverschwendung mit der Mathematikleistung (adjustiert für die Vorjahresleistung) assoziiert war ( $\beta = -0.09$ ,  $p = .003$ ), während dies für das Monitoring einer Lehrkraft nicht zutraf ( $\beta = 0.01$ ,  $p = .766$ ). In einem erweiterten Modell wurde der spezifische Faktor Störungen/Zeitverschwendung – unter Kontrolle der Schulform und des sozioökonomischen Status – sowohl durch die Geschlechtskomposition der Klasse (d.h. Anteil Jungen;  $\beta = 0.23$ ,  $p = .031$ ) als auch durch die Leistungscomposition der Klasse im Vorjahr ( $\beta = -0.32$ ,  $p = .023$ ) vorhergesagt. Zudem medierte der Faktor Störungen/Zeitverschwendung die Effekte beider Kompositionsmerkmale auf die Mathematikleistung ( $ps' < .05$ , einseitig). Für das Monitoring fanden sich wiederum keine Zusammenhänge mit Kompositionsmerkmalen der Klasse.

Zusammenfassend legen die Ergebnisse nahe, dass Schülerbeurteilungen bestimmter Klassenführungsmerkmale je nach Referent der Fragebogenitems mit der Schülerkomposition einer Klasse in Zusammenhang stehen. Inwiefern dies zumindest in Teilen auf eine Interaktion von Lehrkraftverhalten im Unterricht und der Schülerschaft einer Klasse zurückzuführen ist oder weitgehend lehrkraftunabhängig ist, stellt eine überaus bedeutende Fragestellung im Bereich der empirischen Unterrichtsforschung dar.

### **Ego Involvement als Erklärung für unterschiedliche Einschätzungen von Klassenführung durch Lehrkräfte und Schüler/-innen?**

**Anna-Katharina Praetorius<sup>1</sup>, Benjamin Fauth<sup>2</sup>, Oliver Dickhäuser<sup>3</sup>, Stefan Janke<sup>3</sup>, Markus Dresel<sup>4</sup>**  
<sup>1</sup>Universität Zürich, <sup>2</sup>Universität Tübingen, <sup>3</sup>Universität Mannheim, <sup>4</sup>Universität Augsburg

Theoretischer Hintergrund und Fragestellung

Lehrkräfte und Schüler/-innen weisen oft nur geringe bis mittlere Übereinstimmungen bei der Einschätzung von Unterrichtsqualität auf (Clausen, 2002; Fauth et al., 2014; Kunter & Baumert, 2006). Erklärt wird dies mit den der jeweiligen Perspektive zur Verfügung stehenden Informationen und deren spezifischen Sichtweisen, auch als perspektivenspezifische Validität bezeichnet. Bislang fehlen jedoch Modelle, die diese spezifische Validität erklären und von Messfehlern abgrenzen (Wagner, 2008). Mittels solcher Modelle könnte entsprechend bestimmt werden, welche Unterrichtsaspekte durch welche Perspektive(n) optimal erfassbar sind und welche Konsequenzen dies für das Verständnis der erfassten Konstrukte hat. Fauth, Göllner, Lenske, Praetorius und Wagner (siehe Beitrag 1 des Symposiums) haben ein Raster entwickelt, das diese Unterschiede erklären soll. Bezug nehmend auf das Self-Other Knowledge Asymmetry Modell von Vazire (2010) lautet eine zentrale Annahme, dass Zusammenhänge zwischen Lehrer- und Schülerperspektive dann geringer ausfallen sollten, wenn Items bedeutsam für die eigene professionelle Kompetenz der Lehrkraft sind (ego involvement). Im vorliegenden Beitrag soll für das im Fokus des Symposiums stehende Unterrichtsmerkmal Klassenführung zunächst überprüft werden, ob Items, die für eine Lehrkraft aufgrund ihres professionellen Selbstverständnisses ego involvement mit sich bringen sollten, in geringerem Maß mit den Schülerschätzungen zusammenhängen als Items, für die das nicht der Fall ist. Anschließend soll getestet werden, ob die Höhe des Zusammenhangs zwischen Lehrer- und Schülerschätzungen für Items mit ego involvement durch solche

Lehrermerkmale moderiert wird, für die Unterschiede zwischen Lehrkräften in dem Ausmaß der Beeinflussung durch ego involvement angenommen werden können. Konkret wird vermutet, dass motivationale Merkmale (Selbstwirksamkeit und Vermeidungsleistungszielorientierung) sowie die Tendenz sozial erwünscht zu antworten hierfür bedeutsam sind.

#### Methode

Es wurden Fragebogendaten von 1544 Schüler/-innen (54% weiblich, M Alter = 11.6, SD = 0.41) aus 70 Gymnasialklassen der 6. Jahrgangsstufe und deren Lehrkräften (52% weiblich, M Berufserfahrung = 15 Jahre, SD = 13) einbezogen. Die Schüler/-innen und Lehrkräfte wurden mittels 9 Items zu verschiedenen Aspekten der Klassenführung in der jeweiligen Klasse befragt ( $\alpha_{\text{Schüler}} = .90$ ,  $\alpha_{\text{Lehrkräfte}} = .60$ ). Zusätzlich wurden die Lehrkräfte in Bezug auf ihre soziale Erwünschtheit (Stöber, 1999, 16 Items,  $\alpha = .76$ ), ihre Selbstwirksamkeit (Schwarzer & Schmitz, 1999, 10 Items,  $\alpha = .75$ ) sowie ihre Vermeidungsleistungszielorientierung (Nitsche et al., 2011, 12 Items,  $\alpha = .93$ ) befragt. Die Klassenführungsitems wurden von zwei Ratern mittels Konsens im Hinblick auf das Ausmaß an ego involvement für die Lehrkraft kodiert und basierend auf der Kodierung in zwei Gruppen eingeteilt (ego involvement vs. non involvement). Anschließend wurde mittels Mplus auf der Klassenebene für Schüler- sowie Lehrereinschätzungen getrennt jeweils ein ego-involvement- sowie ein non-involvement-Faktor gebildet. Die latenten Einschätzungen wurden anschließend zwischen den beiden Perspektiven korreliert. Abschließend wurden unter Hinzunahme der o.g. Lehrermerkmale Moderationsanalysen durchgeführt, um potenziell Unterschiede in der Höhe des Zusammenhangs zwischen Lehrer- und Schülereinschätzungen für den ego-involvement-Faktor zu erklären.

#### Ergebnisse und Diskussion

Der Modellfit war zufriedenstellend (chi-square = 180.62, df = 129; CFI = .95; RMSEA = .07). Die Korrelation des ego-involvement-Faktors mit dem non-involvement-Faktor lag für die Lehrkräfte bei  $-.24$ ,  $p = .10$ , bei den Schülern bei  $-.53$ ,  $p < .001$ . Die Korrelation zwischen Lehrer- und Schülereinschätzungen des ego-involvement-Faktors beläuft sich auf  $.14$ ,  $p = .40$ , die des non-involvement-Faktors auf  $.66$ ,  $p < .001$ . Von den einbezogenen Lehrermerkmalen zeigte sich in den Moderationsanalysen keines als signifikanter Moderator. Die Befunde unterstützen entsprechend die Vermutung, dass Items, die ego involvement bedingen, mit geringen Korrelationen der Lehrer- und Schülerperspektive bei der Beurteilung von Unterrichtsqualität einhergehen. Hingegen ermöglichen die einbezogenen Moderatorvariablen keine Erklärung interindividueller Unterschiede zwischen Lehrkräften in der Höhe des Zusammenhangs ihrer Einschätzungen mit denjenigen der Schüler/-innen in Bezug auf den ego-involvement-Faktor.

## G14

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H117

### Einstellungen und Überzeugungen gegenüber einer heterogenen Schülerschaft

Chair(s): **Ineke Pit-ten Cate** (University of Luxembourg, Luxembourg)

DiskutantIn(nen): **Sabine Glock** (Bergische Universität Wuppertal)

Lehrpersonen werden in allen Schularten mit einer zunehmend heterogenen Schülerschaft konfrontiert. Mehr und mehr Schülerinnen und Schüler mit unterschiedlichen kulturellen, ethnischen und sprachlichen Hintergründen aber auch sonderpädagogischem Förderbedarf werden gemeinsam in Regelklassen unterrichtet (breites Verständnis von Inklusion). In diesem Kontext fokussiert sich nicht nur die Aus- und Weiterbildung der Lehrpersonen auf die Grundbedingungen der Inklusion unterschiedlicher Schülerinnen und Schüler, sondern auch in der Forschung wird den Faktoren, die mit einer erfolgreichen inklusiven Praxis assoziiert werden, verstärkte Aufmerksamkeit gewidmet. Dabei wird es als wesentlich erachtet, dass diese Bemühungen sich nicht nur auf Wissen und Kompetenzen sondern auch auf Einstellungen und Werte konzentrieren (Borg, Hunter, Sigurjonsdottir, & D'Alessio, 2011). Diese Einstellungen können für den Erfolg der inklusiven Praxis entscheidend sein (Avramidis, Bayliss, & Burden, 2000; Gebhardt et al., 2011), da sie das Handeln beeinflussen können (z.B. Eagly & Chaiken, 1993). Einstellungen sind mit positiven oder negativen Erwartungen und Beurteilungen verbunden, welche unterschiedliche Verhaltensweisen hervorrufen können (z.B. Fazio, 1990), was wiederum die erfolgreiche Einbeziehung unterschiedlicher Schülerinnen und Schüler in Regelklassen ermöglichen oder verhindern kann (Bender, Vail, & Scott, 1995; Villa, Thousand, Meyers, & Nevin, 1996).

Mit diesem Symposium soll das Thema „Einstellungen zum Unterrichten einer heterogenen Schülerschaft“ aufgegriffen werden. Der Schwerpunkt liegt dabei auf den unterschiedlichen Prädiktoren der Einstellungen gegenüber verschiedenen Aspekten der Heterogenität und ihren Auswirkungen auf – intendiertes - Lehrverhalten. Die empirische Forschung, die im Symposium präsentiert wird, verbindet dabei Perspektiven der Erziehungswissenschaften und der pädagogischen sowie der Sozialpsychologie und bezieht sich auf Lehramtsstudierende und erfahrene Lehrkräfte aus unterschiedlichen (Bundes-)Ländern. Ebenso greifen die Beiträge des Symposiums sowie dieses selbst das Tagungsthema „Professionelles Handeln als Herausforderung für die Bildungsforschung“ auf, da sowohl Aussagen zur professionellen Entwicklung und zum Handeln abgeleitet werden als auch methodische Fragen, beispielsweise zur Erfassung und Trennung der als wichtig erachteten Einstellungen von Lehrkräften, diskutiert werden.

In dem Symposium werden vier unterschiedliche Studien präsentiert. Im ersten Beitrag werden die Ergebnisse einer Studie zum Einfluss von Personen- und Herkunftsmerkmalen auf Einstellungen zu unterschiedlichen Dimensionen der Heterogenität der Schülerschaft bei Lehramtsstudierenden vorgestellt. Der zweite Beitrag berichtet über die Einstellungen von erfahrenen Lehrkräften zur Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf im Zusammenhang mit ihrer Aus- und Weiterbildung sowie ihres Kompetenzgefühls. Im dritten Beitrag werden die Einstellungen zur kulturellen und leistungsbezogenen Heterogenität der Schülerschaft im Zusammenhang mit dem intendierten Lehrverhalten bei Lehramtsstudierenden untersucht. Im vierten Beitrag wird berichtet, inwiefern das Lesen unterschiedlicher Lehrbuchtexte zum Thema Heterogenität im Unterricht einen Einfluss auf die Überzeugungen von Lehramtsstudierenden hat. Weiterhin werden in diesem Beitrag Ergebnisse zum Zusammenhang von Überzeugungen, der Selbstwirksamkeit sowie der Motivation in heterogenen Klassen zu unterrichten und dem antizipierten Unterrichtshandeln dargestellt. Abschließend diskutiert unsere DiskutantIn die verschiedenen (inter)nationalen Beiträge vor dem Hintergrund ihrer Expertise im Kontext der Einstellungsforschung gegenüber verschiedenen Heterogenitätsformen der Schülerschaft.

*Beiträge des Symposiums*

### Einstellungen von Lehramtsstudierenden zum Umgang mit Heterogenität und Inklusion: Ausprägungen, Schulartenunterschiede und Prädiktoren

**Marcus Syring, Teresa Tillmann, Sabine Weiß, Ewald Kiel**

Ludwig-Maximilians-Universität München

#### Theoretischer Hintergrund

In kompetenztheoretischen Professionsmodellen (z.B. Baumert & Kunter, 2006) nehmen Einstellung von Lehrkräften eine prominente Stelle ein, da sie ein wichtiger Prädiktor für (erfolgreiches) unterrichtliches Handeln und Lernen sind (z.B. Jacobs, Yoshida, Stigler & Fernandez, 1997) und somit auch darüber entscheiden, welche Praktiken von Lehrkräften im Umgang mit Heterogenität angewendet werden. Einstellungen produzieren bei der Bewertung von Objekten Zu- und Abneigungen (Eagly & Chaiken, 1993) und sind die Summe aus kognitiven, motivational-affektiven und verhaltensbezogenen Aspekten. Die Einstellungen zum Umgang mit Heterogenität und Inklusion können sich dabei auf unterschiedliche Heterogenitätsdimensionen der Schülerschaft beziehen, die relevantesten sind dabei die ethnisch-kulturelle, die soziale und die Leistungsheterogenität sowie der sonderpädagogische Förderbedarf.

Studien zu den Einstellungen von Lehrkräften v.a. zur Inklusion zeigen häufig einen Widerspruch: Trotz einer grundsätzlich positiven Einstellung gegenüber inklusiver Beschulung begegnet man einer zurückhaltenden bis ablehnenden Haltung mit Blick auf die konkrete Umsetzung. Dieser Widerspruch lässt sich teilweise dadurch erklären, dass bestehende Messinstrumente häufig nur Teilaspekte von Einstellungen, vor allem die kognitive Komponente, abbilden.

#### Forschungsfragen

Über die Einstellungen explizit von Lehramtsstudierenden zum Umgang mit Heterogenität und Inklusion ist bisher wenig bekannt, weder über die genannten Aspekte noch über differenzierte Ausprägungen je nach Heterogenitätsdimension. Daher wurden in der zu präsentierenden Studie verschiedene Fragestellungen verfolgt:

1. Welche Einstellungen lassen sich bei Lehramtsstudierenden der verschiedenen Schularten hinsichtlich verschiedener Heterogenitätsdimensionen (ethnisch-kulturelle, soziale, leistungsbezogene Heterogenität, sonderpädagogischer Förderbedarf) finden?
2. Gibt es dabei signifikante Unterschiede hinsichtlich der Einstellungen, die sich auf die Lehramtsart zurückführen lassen?
3. Welche Effekte haben Personen- und Herkunftsmerkmale (Geschlecht, Migrationshintergrund, SES, kulturelles Kapital) sowie die Unterrichtserfahrung auf die Einstellungen?

#### Methode

Zur Beantwortung der Fragestellungen wurden Lehramtsstudierende (n=294) für Grundschule, Förderschule, Mittelschule, Realschule und Gymnasium befragt (72.8% weiblich; Alter M=21.84 (SD=3.70); höchstes Semester: M=3.15 (SD=1.85); Migrationshintergrund < 7%). Dabei kam ein von Gebauer, McElvany und Klukas (2013) entwickelter Fragebogen zu Einstellungen von Heterogenität zum Einsatz, der bereits anhand anderer Daten auf Konstruktvalidität überprüft wurde (Merk, Cramer, Dai, Bohl & Syring, eingereicht). Die Skalen wurden für das Feld der Inklusion adaptiert (Syring, Tillmann, Weiß & Kiel, eingereicht). Die Daten wurden mithilfe von MANOVAS mit Post-Hoc-t-Tests sowie multiplen Regressionen ausgewertet.

#### Ergebnisse

Es zeigt sich, dass sich das adaptierte Instrument für den Einsatz im Kontext Inklusion eignet und das Gesamtbild der Einstellungen mit Blick auf Aspekte wie Mehrwert, Motivation und wahrgenommene Kompetenz ausdifferenziert. Die Ergebnisse deuten ebenso darauf hin, dass sich innerhalb von Personen die Einstellungen je nach Heterogenitätsdimension z.T. signifikant unterscheiden. Zusätzlich zeigen sich zwischen den Lehramtsarten keine signifikanten Unterschiede bezüglich der Einstellungen zur Heterogenität und Inklusion. Bezüglich der Prädiktoren konnte nur ein geringer Effekt des Geschlechts auf die Einstellungen gefunden werden.

#### Implikationen und Ausblick

Anhand der Ergebnisse der Studie sollen Maßnahmen für eine wirksame Gestaltung der ersten Phase der Lehrerbildung mit Blick auf den Umgang mit Heterogenität und Inklusion abgeleitet werden. Insbesondere liegt der Fokus dabei auf Lehr-Lern-Angeboten zur Reflexion und Veränderung von Einstellungen.

### **Einstellungen von Lehrkräften zur Inklusion von SchülerInnen mit sonderpädagogischem Förderbedarf**

**Ineke Pit-ten Cate, Mireille Krischler**  
Universität Luxemburg

Die Ratifizierung der UN-Konvention über die Rechte von Menschen mit Behinderungen bedeutet für die Unterzeichnerstaaten, ein inklusives Bildungssystem zu etablieren, um die bestmögliche Bildung, Erziehung und Förderung aller SchülerInnen mit und ohne sonderpädagogischem Förderbedarf an wohnortnahen und gemeinsamen Lernorten zu garantieren. Das erfolgreiche Umsetzen der Inklusion hängt insbesondere von den Lehrkräften ab. Inklusiver Unterricht konfrontiert Lehrkräfte mit neuen und komplexen Herausforderungen, zu deren Bewältigung positive Einstellungen eine zentrale Ressource darstellen (Avramidis, Bayliss, & Burden, 2000; Lütje-Klose & Urban, 2014). Einstellungen haben eine handlungsleitende und handlungsregulierende Funktion (Eagly & Chaiken, 1993) und beeinflussen das Lehrverhalten (Reusser & Pauli, 2014). Die bisherigen Forschungsbefunde in diesem Bereich sind eher widersprüchlich. Während de Boer und Kollegen (2011) in ihrem Review berichten, dass Lehrkräfte eher eine neutrale bis leicht negative Einstellung zur schulischen Inklusion haben, weisen Beacham und Rouse (2012) beispielsweise darauf hin, dass die meisten Studien positive Einstellungen zur gemeinsamen Beschulung von Schülern mit und ohne sonderpädagogischen Förderbedarf zeigen.

Untersucht wurde demzufolge die Frage der Einstellungswalenz der Lehrkräfte zur Inklusion von SchülerInnen mit sonderpädagogischem Förderbedarf und welche Faktoren diese Einstellungen möglicherweise vorhersagen. Konkret wurde der Einfluss der Thematisierung von Inklusion in Aus- und Weiterbildung (Leipziger, Tretter, & Gebhardt, 2012), bisher gemachter Erfahrungen mit Menschen mit sonderpädagogischem Förderbedarf (Avramidis & Norwich, 2002), der Berufserfahrung (de Boer, Pijl, Post, & Minnaert, 2012) sowie des Kompetenzgefühls (Sermier Dessemontet, Benoit, Bless, 2011) auf die Einstellungen analysiert.

Auf Basis eines Online-Fragebogens wurden 126 Grundschullehrkräfte befragt inwieweit sie sich auf den Unterricht einer heterogenen Schülerpopulation vorbereitet fühlen und wie sie der Inklusion von SchülerInnen mit sonderpädagogischem Förderbedarf gegenüberstehen. Der Fragebogen beinhaltete die deutsche Übersetzung des Fragebogens Opinions Relative to Integration (ORI; Antonak & Larrivee, 1995; Benoit & Bless, 2014) sowie des Fragebogens „Teacher Knowledge and Skill Survey for positive Behavior Support (TKSS-PBS; Blum und Cheney, 2009).

Erste Ergebnisse zeigen, dass die Grundschullehrkräfte die potentiellen Vorteile der Inklusion sehen, jedoch Zweifel daran haben, ob die SchülerInnen mit sonderpädagogischem Förderbedarf in der Regelschule optimal gefördert werden können. Obwohl sich die Grundschullehrkräfte in der Ausübung ihres Berufes als kompetent ansehen, zweifeln sie an ihrer Unterrichtsfähigkeit in Bezug auf SchülerInnen mit sonderpädagogischem Förderbedarf. 66% der Grundschullehrkräfte glauben ihr Studium habe sie nicht ausreichend auf die inklusive Praxis vorbereitet und halten zusätzliche Weiterbildungsangebote bezüglich der inklusiven Praxis für erforderlich. Ausserdem zeigten sich signifikante Korrelationen zwischen den Einstellungen zur Inklusion, dem Kompetenzgefühl, der Rolle der inklusiven Pädagogik in der Ausbildung und der Anzahl der Schüler mit sonderpädagogischem Förderbedarf in der Klasse. Die Berufserfahrung wurde hingegen nicht mit den Einstellungen zur Inklusion assoziiert.

Anhand dieser Ergebnisse könnten Maßnahmen für eine wirksamere Gestaltung der Lehrerbildung mit Blick auf den Umgang mit Heterogenität abgeleitet werden. Insbesondere liegt der Fokus dabei auf Lehr-Lern-Angeboten zur Reflexion und Veränderung von Einstellungen. Ausgehend von der Verbindung zwischen der Kompetenz und den Einstellungen kann man erwarten, dass die Erhöhung der spezifischen Kompetenz (z.B. im Bereich individueller Förderung) durch zusätzliche Bildungsangebote eine positive Auswirkung auf die Einstellungen gegenüber der Inklusion haben wird.

### **Einstellungen von Lehrenden für intendiertes Unterrichtsverhalten**

**Miriam Gebauer, Nele McElvany**  
Technische Universität Dortmund

#### Theoretischer Hintergrund

Einstellungen, die angehende Lehrende bezogen auf das Unterrichten von heterogenen Schülergruppen haben, können, neben anderen zentralen Faktoren, bedeutsam dafür sein, wie sie dieser Herausforderung im schulischen Unterricht begegnen (Ajzen, 2001; Eley, 2006; Kunter & Trautwein, 2013). Dabei sind Einstellungen die psychische Tendenz, die sich im Ausmaß an Zuneigung oder Abneigung - bezogen auf ein bestimmtes Objekt oder eine Meinung - niederschlägt (Eagly & Chaiken, 1993, 2007). Differenzielle Analysen liefern empirische Evidenzen für Geschlechterunterschiede und zeigen unter anderem, dass weibliche Studierende eher positive Einstellungen in Bezug auf das Unterrichten haben (Yildirim, 2012). Wiederholt wurden auch Unterschiede zwischen Studierenden festgestellt, die unterschiedliche Bildungsgänge studieren. So konnten zwischen angehenden Lehrenden, die Grund-Haupt-Realschulbildungsgänge studieren, und Gymnasiallehramtsstudierenden substantielle Unterschiede in Bezug auf lehr-lerntheoretische Vorstellungen festgestellt werden (Müller, Felbrich & Blömeke, 2008 a,b). Einstellungen von angehenden Lehrenden bezogen auf heterogene Schülergruppen können sich auf

unterschiedliche Aspekte der Heterogenität ihrer Schülerschaft beziehen. Untersuchungen, die Einstellungen von Lehrenden bezogen auf z.B. kulturelle Heterogenität zum Gegenstand haben, betrachten weniger den Zusammenhang zwischen den Einstellungen von Lehrenden und dem schulischen Lernerfolg von Schülerinnen und Schülern, sondern vielmehr multikulturelle versus egalitäre Haltungen von Lehrenden in Bezug auf ethnisch-kulturell heterogene Schülergruppen (Bryan & Atwater, 2002; Zeichner & Hoeft, 1996).

Forschungsfragen

Forschungsfrage 1a: Sind die Einstellungen von angehenden Lehrenden, die diese in Bezug auf das Unterrichten von kulturell-heterogenen Schülergruppen haben, empirisch von ihren Einstellungen zu leistungsheterogenen Schülergruppen trennbar?

Forschungsfrage 1b: Inwieweit unterscheiden sich die Dimensionen Einstellungen im Bereich kulturelle Heterogenität von den Einstellungen zur leistungsbezogenen Heterogenität?

Forschungsfrage 2: Sind die Einstellungen von angehenden Lehrenden, die diese in Bezug auf das Unterrichten von kulturell- und leistungsheterogenen Schülergruppen haben, von Bedeutung für das intendierte unterrichtsbezogene Verhalten?

Forschungsfrage 3: Bleiben die gefundenen Zusammenhänge nach Kontrolle individueller Merkmale der Lehramtsstudierenden (Geschlecht, studierte Schulform und Studiendauer) bestehen?

Methode

Es wurden N = 345 angehende Lehrende befragt (72.2 % weiblichen) mit einem mittleren Alter von 24.5 Jahren (SD = 2.9). 51.6 % studierten Grundschullehramt, 16,4% Lehramt für Haupt- und Realschulbildungsgänge und 26.9 % studierten Gymnasiallehramt. Die Analysen zur Beantwortung der Forschungsfragen wurden anhand von Strukturgleichungsmodellen durchgeführt.

Ergebnisse

Die heterogenitätsspezifischen Modellprüfungen ließen auf differenzielle Einstellungen der angehenden Lehrenden schließen. Dabei konnte die theoretisch angenommene Trennbarkeit der Einstellungsskalen Nützlichkeit, Kosten und negative Emotionen bezogen auf die kulturelle und leistungsbezogene Heterogenität von Schülerinnen und Schülern empirisch bestätigt werden. Bei der Überprüfung der Dimensionsunterschiede zeigen sich über alle Skalen hinweg statistisch signifikante Differenzen zwischen den kulturell- und leistungsbezogenen Einstellungsskalen.

Zwischen den Einstellungsskalen und den Skalen zur Erfassung des intendierten Unterrichtsverhaltens konnten – theoriekonform - heterogenitätsspezifische und substantielle Zusammenhänge identifiziert werden. Die Ergebnisse stehen im Einklang mit anderen Forschungen, bei denen die Einschätzung der Nützlichkeit für leistungsbezogenes Verhalten relevant waren, hingegen die Kosten aber nicht bedeutsam waren (Steinmayr & Spinath, 2010; Wigfield & Eccles, 2002). Die Merkmale Geschlecht und Studiengang erklärten das intendierte Unterrichtsverhalten gemäß den theoretischen Erwartungen. Es war aufgrund des in der Literatur beschriebenen gender gap von geschlechterdifferenziellen Zusammenhängen und einer höheren Bereitschaft der weiblichen Studierenden auszugehen Heterogenität zu berücksichtigen, (Yildirim, 2012). Die Bedeutsamkeit des Studiengangs entsprach erwartungskonform der theoretischen Vermutung, dass Studierende der höheren Semester, die möglicherweise über Wissen verfügen wie Unterricht für heterogene Schülergruppen zu gestalten ist, über positivere Einstellungen verfügen. Dies würde die theoretischen Annahmen empirisch stützen, dass Einstellungen durch Wissenserwerb veränderbar sind.

## **Wie verändern sich die Überzeugungen von Lehramtsstudierenden zur Heterogenität durch das Lesen von Lehrbuchtexten**

**Charlotte Dignath, Jonathan Fink, Bettina Retzbach, Mareike Kunter**  
Goethe-Universität Frankfurt

Heterogenität kann verschiedene Dimensionen betreffen, beispielsweise kultureller Hintergrund, sonderpädagogischer Förderbedarf, sozioökonomischer Status, aber auch Geschlecht, Leistung oder Motivation. Insbesondere im Zuge der Umsetzung der UN-Menschenrechtskonvention zur Inklusion sowie im Zuge von Migrationsbewegungen sind Lehrkräfte zunehmend mit neuen Formen von Heterogenität in ihrem Unterricht konfrontiert. Ein reflektierter Umgang mit Heterogenität nimmt darum in der Lehrerbildung eine zunehmende Rolle ein. Dabei ist noch wenig bekannt über die Entwicklung und Veränderung von Überzeugungen zum Unterrichten in heterogenen Klassen bei Lehramtsstudierenden. Im Rahmen von Modellen zur professionellen Lehrkompetenz (Baumert & Kunter, 2006; Kunter et al., 2010) wird davon ausgegangen, dass Überzeugungen mit Motivation und Selbstwirksamkeit zusammenhängen und Prädiktoren sind für professionelles Lehrverhalten. Obwohl die professionelle Kompetenz von Lehramtsstudierenden und dabei insbesondere die Überzeugungen zum Unterrichten in heterogenen Klassen in der aktuellen Forschung zunehmend Berücksichtigung finden, wurde sich bisher nur wenig damit beschäftigt, die Entwicklung der Determinanten professioneller Kompetenz genauer zu untersuchen.

Zunächst wurde die Veränderung von Überzeugungen zu Inklusion sowie von Überzeugungen zu kultureller Heterogenität untersucht. In einem experimentellen Setting wurde getestet, ob das Lesen eines Lehrbuchtextes zum Thema Heterogenität im Unterricht Effekte auf die Überzeugung von Lehramtsstudierenden hat (Fragestellung 1). In einem nächsten Schritt wurde untersucht, ob eine Veränderung der Überzeugung Auswirkungen hat auf die Selbstwirksamkeit und die Motivation, in heterogenen Klassen zu unterrichten (Fragestellung 2). Außerdem wurde getestet, ob sich hieraus auch Effekte für das antizipierte Unterrichtshandeln der Lehramtsstudierenden ausmachen lassen (Fragestellung 3).

Das experimentelle Paradigma bestand aus einem Lehrbuchtext, der in der positiven Bedingung eher positiv über Heterogenität im Unterricht berichtete („Heterogenität als Chance“), während der Text in der negativen Bedingung eher negativ gefärbt über das Thema informierte (Heterogenität als Herausforderung/Problem“). Die Texte waren im Aufbau möglichst parallel gehalten, und die Studierenden wurden den Bedingungen randomisiert zugewiesen. An den beiden Studien nahmen 83 bzw. 57 Studierende mehrerer Lehrveranstaltungen in den Bildungswissenschaften teil.

(1) Regressionsanalysen zeigten, dass das experimentelle Paradigma zu einer Überzeugungsveränderung führte: Studierende aus der positiven Bedingung berichteten im Posttest signifikant positivere Überzeugungen zum Unterrichten in heterogenen Klassen als diejenigen, die den negativen Text gelesen hatten; auch unter Berücksichtigung der Prätestwerte. Dabei machte die Art der Heterogenitätsdimension (Überzeugung zu Inklusion vs. Überzeugung zu kultureller Heterogenität) keinen Unterschied.

(2) Die Überzeugung korrelierte zudem signifikant mit der Selbstwirksamkeitserwartung und der Motivation, in heterogenen Klassen zu unterrichten.

(3) Es zeigten sich allerdings nur vereinzelte Effekte auf das antizipierte Unterrichtshandeln der Studierenden.

Die Ergebnisse zeigen, dass das Lesen von Texten zu einer zumindest kurzfristigen Überzeugungsänderung führen kann, die allerdings keine unmittelbare Auswirkungen auf die Selbstwirksamkeit oder die Motivation hat, und auch das antizipierte Verhalten nur bedingt vorhersagt. Die Ergebnisse liefern neue Befunde zum Zusammenspiel der Determinanten professioneller Kompetenz im Bereich der Heterogenität und legen nahe, dass Wissensvermittlung zum Thema Heterogenität in der ersten Phase der Lehrerausbildung eine zentrale Rolle spielen sollte.



## G15

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H116

### **Professional competence in teaching and learning from a domain specific research perspective. Empirical assessments of prescriptive models and assumptions.**

Chair(s): **Irene Pieper** (Universität Hildesheim, Deutschland)

DiskutantIn(nen): **Bernard Schneuwly** (Université de Genève)

In this symposium we take up a dilemma that educational research in Reading and Literature has kept busy for some time: Models we design for classroom practice are based on theoretically sound and normative principles. However, actual classroom teaching is based on potentially diverse attitudes, beliefs and aims. They are often unknown to the discipline and whether or not they fit the theoretical paradigms can only be assumed.

The challenge that theoretically based models may not be accepted and implemented in every day's teaching has been subject to reflection in domain specific educational research (German Fachdidaktik) for quite some time. Tenorth (2012) has commented on this problem from the point of view of educational research and Wieser (2016) demanded with regard to professional research in language and literature teaching: Research on professional competences should deal not only with prescriptive concepts, but should also describe actual classroom teaching and learning and reconstruct the attitudes and beliefs of teachers.

Research in reading and literature education thus has to face the task of taking up developments and methods from research in professional action in order to tackle specific tasks around the teaching and learning in the literature classroom. An interdisciplinary approach which brings together psychology, educational research and domain specific research in reading and literature education is needed and can take up previous work such as outcomes of the research alliance of Groeben, Hurrelmann and others (e.g. Groeben/Hurrelmann 2006).

We take up the desideratum of domain specific research into teaching and learning in the literature classroom and take an inter- and transdisciplinary perspective. At the same time the potentials of different research approaches and instruments is assessed.

The symposium consists of the following three papers:

1

Iris Winkler and Frederike Schmidt take up a well established general criterion of teaching quality, the aspect of cognitive activation. In the project KOALA a specification of this criterion with regard to domain specific learning in literature education is developed on the basis of videotaped literature lessons in Lower Secondary/Grade 8.

2

Marloes Schrijvers, Tanja Janssen and colleagues present a study which has investigated effects of literature teaching in upper secondary on the perception of self and others with the students, a major aim of literature education. Their team used the instrument 'learner report' to assess the learning outcomes of 297 students in Dutch Upper Secondary with regard to this aim.

3

Irene Pieper and colleagues assess a concept of paradigms for literature education (Witte & Samihaian 2012) on an empirical basis. The four paradigms, Witte and Samihaian distinguish, were developed in an international comparative perspective. The assessment is carried out via the study TAMoLi (Texts, Activities and Motivations in Reading and Literature Education), which investigates attitudes, aims and procedures of teachers in lower secondary in German speaking Switzerland and Lower Saxony as well as the reading motivation of their students.

*Beiträge des Symposiums*

### **Cognitive activation in L1 literature classes. Domain specific conceptualization of a prominent construct in empirical classroom research**

**Iris Winkler, Frederike Schmidt**  
Friedrich Schiller Universität Jena

#### Theoretical Background

Empirical classroom research proved three general core dimensions of teaching quality: classroom management, supportive climate, cognitive activation. All of them support desirable effects of learning (Seidel & Shavelson, 2007; Baumert et al., 2013). From a view that stresses subject teaching, cognitive activation is the most interesting of these dimensions: It allows the description of domain specific characteristics of teaching that very probably support the pupils' learning outcome. However, most existing studies about cognitive activation focus on mathematics and sciences classes. For L1 classes there are suggestions for the conceptualization and operationalization of cognitive activation concerning reading instruction (e. g. Lotz, 2016). To apply the concept of cognitive activation to L1 literature classes it has to be specified for this domain that includes reading but still more than reading (Klieme, 2006; Klieme & Rakoczy, 2010). The conceptualization and operationalization has to aim at connecting the domain specific discussion about 'good' literature teaching with the generic view of empirical classroom research and the question of 'effective' teaching (Kunter & Ewald, 2016).

#### Research question

Against this background, the leading question of the paper is what "cognitive activation" can mean in literature classes. The context of this question is the research project KoALa (Cognitive Activation by Tasks in Literature Lessons). In the KoALa project we are going to investigate possible relations between teacher beliefs, the potential of cognitive activation of their literature lessons and the pupils' learning results (Winkler & Steinmetz, 2016). For this kind of mixed-methods research a careful specification of constructs is fundamental. Thus, in conceptualising and operationalising "cognitive activation" for literature lessons the paper presents the basis for further research in KoALa.

#### Method

The first step of the specification is a theoretical approach to the research question. This theoretical conceptualization is empirically underlaid with data of a pilot study for KoALa. For the pilot study, six literature lessons about the same short story

with different teachers and classes (107 pupils, grade 8, German "Gymnasium") were videotaped. In an iterative process data from the pilot study and theoretically based rating categories were brought into a balance.

## Results and Discussion

The operationalization regards tasks as indicators for the potential of cognitive activation. This is usual in studies focussing on cognitive activation, because tasks are essential tools for stimulating and structuring domain specific learning processes (Neubrand et al., 2013; Drollinger-Vetter, 2006, Drollinger-Vetter & Lipowsky, 2006). However, selected and realised tasks need to be distinguished in teaching processes. The operationalization is based on this crucial distinction. The coding is carried out based on events. An event is determined by a task (selected task plus task realization).

The main categories concerning the selected tasks are the focus and the cognitive demands of a task. Concerning the task realization the coding monitors if the characteristics of the selected task are modified during the teaching process. Additionally there is paid attention to the way of discussing and solving the task in the class plenum (depth of processing). The connection to normative discussions about desirable aims and ways of teaching literature is considered especially in two categories: the focus of the selected task (text and/or reader, e.g. Spinner, 2006; Vischer Bruns, 2011; Zabka, 2015a;) and the depth of processing (e.g. Zabka, 2015b).

The offered operationalization of cognitive activation for literature classes allows to examine characteristics of literature teaching very carefully. At the same time, it stimulates the discussion which characteristics of literature teaching should be focused on.

## **What students learn about themselves and others in the literature classroom: An analysis of written learner reports**

**Marloes Schrijvers, Tanja Janssen**  
Universität Amsterdam

### Theoretical Background

According to the Dutch Institute for Curriculum Development literature education 'has an important value for developing citizenship, [for instance by] broadening social and cultural horizons and developing empathic capabilities' (2015, p. 15). Others have contended that reading literary fiction can be self-modifying for readers and thus may have an impact on their self-development (Fialho, 2012; Miall & Kuiken, 2002). However, most studies are about avid adult readers. Little is known about what adolescent students gain from literature and literature teaching at school, and whether such encounters actually contribute to their personal and social development.

The aim of our study was to gain insight into students' personal and social learning experiences as a result of literature education in secondary schools. We based our approach on De Groot's (1980) views on educational objectives and outcomes, and on how to evaluate them. According to De Groot, the objective of education is teaching mental programs: the readiness to use knowledge and skills at one's own discretion, freely and consciously. This implies that ultimately, the learner himself is the expert regarding what he has learned.

Thus, while some learning outcomes can be measured through knowledge tests or demonstrations of skills, others call for written or verbal reports of students. For this, De Groot designed an instrument: the learner report. As learning outcomes related to personal and social development tend to resist objective measurement, we investigated whether students would be able to report them in a learner report.

### Research questions

Our research questions were:

- 1) What are the affordances and pitfalls of the learner report as appears from applications in previous research?
- 2) Are students able to report learning experiences concerning their self-perceptions and social perceptions attributed to literature education, and if so, which kinds of learning experiences are these?

### Method

First, we consulted educational studies which used the learner report as an evaluation instrument, in particular studies of literature education. Next, we collected 297 learner reports, written by students in Dutch upper secondary education (age 14-20, grades 10-12) (Schrijvers, Janssen, Fialho & Rijlaarsdam al., 2016). These learner reports focused on the impact of literature teaching on students' perceptions of themselves and others. To develop a coding scheme, we used a comparative qualitative method based on linguistics and phenomenology (Fialho, 2012). Ultimately, we distinguished nine categories of learning experiences that reflected personal and social insights.

### Results and Discussion

The learner report has been used in various Dutch educational studies, especially during the 1980's and 1990's. Applications take different forms, varying from written to verbal reports, and from open to more closed versions. The instrument appears to be adequate to trace learning experiences that cannot be detected in another way. Moreover, we found some indications that the learner report can be a valid and reliable instrument under certain conditions.

Findings of our own study showed that nearly all students (99%) reported to have learned something about themselves and others through literature education. Most frequently, students reported personal characterizations of themselves and others (80%), learning about themselves and others as literary readers (65%), descriptions and evaluations of people's behaviours in fiction and in real life (49%), and lessons for life (33%). Other experiences were reported less frequently, for instance, learning about one's future desired or undesired selves (13%).

A question that remains is: which type(s) of literature education may best promote this kind of personal and social learning experiences? We found some indications that a reader-oriented, text-experiencing approach to literature is more effective than an analytical-interpretative approach in this respect. However, this issue deserves more research.

## **How do teachers' orientations differ with regard to literature education in lower secondary? Paradigms of literature education re-assessed via the bi-national study TAMoLi (texts, activities and motivation in the literature classroom)**

**Irene Pieper<sup>1</sup>, Katrin Böhme<sup>2</sup>, Andrea Bertschi-Kaufmann<sup>3</sup>, Steffen Siebenhüner<sup>3</sup>, Nora Kernen<sup>3</sup>, Simone Depner<sup>1</sup>**  
<sup>1</sup>Universität Hildesheim, <sup>2</sup>Universität Potsdam, <sup>3</sup>PH FH Nordwestschweiz

## Theoretical Background

Since the late 20th century L1 education in the Western world has shown a strong turn towards learner orientation (Pieper [ed.] 2007). Current curricula for literature education formal and enacted reflect the shifts to a large extent. While student-orientation has become stronger, content-orientation remains important. International research has distinguished four major paradigms: The cultural paradigm aims at cultural literacy, the linguistic at aesthetic awareness, the social at social awareness and the personal growth paradigm at personal development (Witte & Sâmihãian 2013; Schrijvers et al. 2016). The model relates the four paradigms not only to aims, but also to the aspects criteria of text selection, approaches to texts, class management, teacher role and evaluation.

It is expected that the paradigms are not exclusively present in the literature classroom, but combined. Thus, from a developmental perspective it has been argued that the two student-oriented paradigms are central as a starting point. The linguistic paradigm should come in early in lower secondary and the cultural paradigm only late in lower secondary. However, little is known about how teachers' aims in lower secondary are related to these paradigms and how they deal with potential conflicts between content-oriented aims and student-orientation as professionals.

## Research Question

In our study TAMoLi we investigate how aims and procedures of teachers in Grade 8 and 9 in German speaking Switzerland and Lower Saxony relate to the four paradigms mentioned above. One aim of our study is to assess whether the model can be empirically validated.

We expect that the distinction between student-oriented paradigms on the one hand and content-oriented paradigms on the other hand has made room for a more permeable border between the two.

## Methods

The analyses are based on data from the TAMoLi-study. This study uses a mixed-methods sequential explanatory design, in which a quantitative phase is followed by a qualitative (Ivankova, 2006). In the first phase, we used teacher and student questionnaires to investigate teacher's beliefs and aims in lower secondary in German speaking Switzerland and Lower Saxony as well as the reading motivation of their students. The teachers' questionnaire included a scale covering the four paradigms for literature education with 21 items in total and at least 4 items per paradigm. The teachers answered the question "How important are the following aims of the German literature and reading classroom to you?" on a 4-point Likert scale. In Switzerland 64 classes of 8th and 9th grade from seven different Swiss cantons took part in our study, covering all types of secondary education. In Germany 62 classes (8th grade) from Lower Saxony participated in the study, again covering all types of secondary education in this federal state. As the German data is currently cleaned, the analyses are based on the Swiss subsample only. The mean age of the Swiss teachers was 43 years, 61% of them were female.

The model fit was evaluated investigating the psychometric properties of the scales and items (internal consistency reliability, item difficulty, item discrimination) as well as by evaluating the factor structure of the scale (EFA, CFA).

## Results and Discussion

The mean of the subscales for the four paradigms varies between  $M = 2.3$  ( $SD = .61$ ) for the cultural paradigm and  $M = 3.4$  ( $SD = .42$ ) for personal growth. Cronbach's Alpha coefficient as measure of consistency ranges between  $\alpha = .35$  and  $\alpha = .76$  for the four paradigms. The item discrimination was within satisfactory ranges. CFA neither supports a four nor a two-factor solution, showing unacceptable fit. EFA suggests a one-factor solution. Whether the results could suggest an integrated and balanced paradigm in Lower Secondary is discussed.

## G16

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H115

### **Bildungsrelevante Rationalitäten – Perspektiven einer professionstheoretisch ausgerichteten Bildungsforschung**

Chair(s): **Corrie Thiel** (Universität Münster, Deutschland), **Fabian Dietrich** (Ruhr-Universität Bochum, Deutschland)

DiskutantIn(nen): **Hedda Bennewitz** (Universität Kassel)

Über die letzten Jahrzehnte hinweg wurden in einer Vielzahl von Ländern Instrumente einer „Neuen“ bzw. „datenbasierten“ Steuerung etabliert. Instrumente wie Bildungsstandards oder schulische Qualitätsrahmen definieren Soll-Vorgaben für die schulische und unterrichtliche Arbeit. Instrumente wie zentrale Vergleichsarbeiten und Schulinspektionsverfahren überprüfen das Erreichen dieser Soll-Vorgaben mit Hilfe wissenschaftlicher Verfahren. Daneben wird die schulische Autonomie erweitert, um einzelschulisches Engagement im Bereich der Qualitätsentwicklung zu ermöglichen. Zentrale Aufgabe im Bereich der Qualitätsentwicklung ist dabei die Minimierung von Ist-Soll-Differenzen (Bellmann 2016). Diese Entwicklung lässt sich auch als Etablierung einer neuen bildungspolitischen Steuerungsrationaltät fassen. Sie trifft in schulischen Handlungsfeldern auf feldspezifische Rationalitäten, die bedingen, dass die Steuerungsinstrumente in schulischen Handlungsfeldern keineswegs bruchlos aufgegriffen, sondern auf sehr unterschiedliche Weise „rekontextualisiert“ (Fend 2008) werden.

Das Symposium versammelt Beiträge, in denen die Ergebnisse professionstheoretischer Analysen der skizzierten bildungspolitischen Entwicklungen sowie deren Rekontextualisierungen in der pädagogischen Praxis vorgestellt werden. Neben dem inhaltlichen Interesse an den Ergebnissen sollen anhand der präsentierten Untersuchungen die Erträge, Grenzen und methodischen Herausforderungen einer professionstheoretischen Forschung für die Rekonstruktion bildungsrelevanter Rationalitäten diskutiert werden.

Im Mittelpunkt des Beitrages von NN steht die Frage nach De-Professionalisierungspotenzialen einer gegenwärtig vorangetriebenen evidenzbasierten Steuerung im Schulsystem, die am Beispiel der Schulinspektion empirisch-fallrekonstruktiv bearbeitet und theoretisch diskutiert wird. In der dem Beitrag zugrunde liegenden rekonstruktiv reformulierten governance-analytischen Perspektive wird Professionalität nicht allein als besondere Form einer beruflichen Qualifikation gefasst, sondern als spezifischer Modus der Handlungskoordination. Dieser wird, so ein zentrales Ergebnis der auf einer Interviewstudie basierenden objektiv-hermeneutischen Rekonstruktionen, durch in die Schulinspektion eingeschriebene Konzeptionen von Professionalität und innerschulischer Handlungskoordination unter Druck gesetzt.

NN präsentiert die Ergebnisse einer diskursanalytischen Untersuchung, die rekonstruiert, wie Lehrerkooperation im Kontext von Schulentwicklungsdiskursen zu einem festen Bestandteil pädagogischer Professionalität geworden ist. Das Konzept der Professionalität kommt hier als bildungspolitisch wie bildungspraktisch bedeutsames diskursives Konzept in den Blick. Auf Basis von Interviews mit österreichischen Lehrkräften zeichnet sie nach, wie sich das Sprechen über Lehrerkooperation vollzieht und welche Sprechpositionen Lehrkräfte hierbei einnehmen. Die Untersuchung kommt zu dem Ergebnis, dass Lehrkräfte durchaus unterschiedliche Positionen einnehmen. Während auf der einen Position Lehrerkooperation als Dimension pädagogischer Professionalität angenommen ist, ist dies auf anderen Positionen nicht der Fall.

Auch im Beitrag von NN wird Professionalität als diskursives Konzept in den Blick genommen. Sie präsentiert drei Beispiele empirischer Untersuchungen, in denen das Konzept der Professionalität eingesetzt wird, um die Deutungen bildungspolitischer Initiativen von Seiten der Lehrkräfte zu analysieren. Anhand dieser Studien zeigt sie auf, in welcher Weise der Terminus Professionalität von verschiedenen Akteuren des norwegischen Schulsystems umgedeutet und genutzt wird. Anhand der Beispiele thematisiert sie auch Erträge und Herausforderungen des Gebrauchs des Konzeptes „Professionalität“ in den vorgestellten Studien.

Im Beitrag von NN werden die Ergebnisse einer Studie präsentiert, in der das Konzept der Professionalität nicht in seiner diskursiven Verwendung im Vordergrund stand, sondern in der auf Basis von Interviews und Gruppendiskussionen differente Formen des Umgangs mit Instrumenten Neuer Steuerung von Lehrkräften in Berlin und Thüringen aus der Perspektive strukturtheoretischer Professionstheorie rekonstruiert worden sind. Auch in diesem Beitrag werden die Erträge sowie die methodischen Herausforderungen des Rückgriffs auf das Konzept „Professionalität“ diskutiert.

*Beiträge des Symposiums*

### **Evidenzbasierte Steuerung als (De)Professionalisierung? Empirische Rekonstruktionen und theoretische Überlegungen am Beispiel der Schulinspektion**

**Fabian Dietrich**

Ruhr-Universität Bochum, Deutschland

Im Zuge der vielfältigen und z.T. weitreichenden Reformen der vergangenen zwanzig Jahre rückten zunehmend auch die Fragen der Steuerung und Steuerbarkeit des Schulsystems in den Fokus (Altrichter/Maag Merki 2016). Im Zentrum einer sich in diesem Zuge etablierenden „Neuen Steuerung“ steht die Programmatik einer „Evidenzbasierung“: Mithilfe von an Methoden der empirischen Sozialforschung orientierter Verfahren soll Steuerungswissen generiert werden, welches sowohl bildungspolitische Entscheidungen als auch Schulentwicklungsprozesse initiiert und anleitet sowie letztlich die unterrichtliche Praxis orientiert (Dedering 2012).

In einer dem Beitrag zu Grunde liegenden rekonstruktiv reformulierten governance-analytischen Perspektive (Dietrich 2014, 2017) stellt sich dabei jenseits der meist im Fokus der Diskussion stehenden Problematik, inwieweit sich die an die Steuerungsinstrumente geknüpften Hoffnungen und Erwartungen erfüllen oder aber enttäuscht werden, die Frage nach der (Neu)Relationierung von Bildungspolitik, Bildungsverwaltung und schulischer Praxis im Zuge der Etablierung von Elementen einer dieser Programmatik folgenden „Neuen Steuerung“ im als Mehrebenensystem gefassten Schulsystem. Die professionalisierungstheoretische Relevanz dieser Frage ergibt sich, wenn in Anlehnung an Schimank Professionalität nicht (allein) als besondere Form der Qualifikation des schulischen Personals gefasst wird, sondern gleichzeitig als spezifischer Modus der Handlungskoordination (Schimank 2014) – in Abgrenzung etwa zu marktförmigen oder bürokratischen Modi der Koordination. Davon ausgehend wird anhand des Beispiels der Schulinspektion gefragt, inwieweit diese in Richtung einer (De-)Professionalisierung des schulischen Personals arbeitet.

Ausgehend von Ergebnissen einer empirisch-rekonstruktiv angelegten Teiluntersuchung des vom BMBF geförderten Verbundprojektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene“ (Böhm-Kasper et al. 2016) richtet sich der Vortrag auf in das Steuerungsinstrument eingeschriebene

„weiche“ Formen des Steuerns (Börzel 2007). Auf der empirischen Grundlage von themenzentrierten Interviews mit auf unterschiedlichen Ebenen verorteten Akteuren (Lehrkräfte, Schulleitungen, SchulinspektorInnen), die mit Hilfe der Objektiven Hermeneutik (Oevermann 2000, Wernet 2006) rekonstruiert wurden, werden Adressierungen der schulischen Akteure herausgearbeitet, die jenseits der Inspektionsergebnisse tradierte berufliche Selbstverständnisse und Formen der innerschulischen Handlungskoordination unter Druck setzen. Die daraus ableitbare Deprofessionalisierungsthese (vgl. Tacke 2005) wird abschließend dahingehend problematisiert und professionalisierungstheoretisch eingeordnet, dass die Rekonstruktionen der Bezugnahmen auf der Ebene der Lehrkräfte gerade nicht auf ein stabiles und konsistentes professionelles Selbstverständnis verweisen.

### **Lehrerkooperation als Merkmal pädagogischer Professionalität**

**Cornelia Dinsleder**

Fachhochschule Nordwestschweiz

Lehrerkooperation wird als eine wichtige Dimension der Professionalisierung von Lehrer\*innen im Kontext von Schulentwicklung diskutiert (Esslinger 2002, Steinert et al. 2006, Terhart/Klieme 2006, MaagMerki/Steinert 2009, Boller 2009, Baum/Idel/Ullrich 2012, Huber/Ahlgrimm 2012, Bondorf 2013, Fussangel/Gräsel 2014, Massenkeil/Rothland 2016). Die Konzeptualisierung von pädagogischer Professionalität von Lehrer\*innen veränderte sich in den vergangenen Jahrzehnten durch die gesteigerte Bedeutung der Entwicklung der Einzelschule als pädagogische Handlungseinheit (Fend 1986).

Erst durch die Steuerungsveränderung des Schulsystems – das sich von einer zentralen Steuerung nach einer top-down-Logik distanzierte und zunehmend dezentral gesteuert wird – wurden individuell zu gestaltenden Schulprofilierungen wichtig: Die Gestaltungsaufforderung der einzelnen Schulen richtete sich an Lehrer\*innen, die zu einer kooperativen Schulentwicklung aufgerufen wurden (vgl. Rolf 1998). Steinwand (2012) analysierte Beiträge zum Diskurs der Lehrerkooperation und stellte im Anschluss an die Untersuchung die Frage: Inwiefern nur kooperierende Lehrer\*innen als gute Lehrer\*innen gelten können? (ebd., S. 32) Diese kritische Nachfrage wurde zum Anlass für die Forschungsarbeit, die in diesem Beitrag vorgestellt wird. Die zentrale These lautet: Die Neufassung der Professionalität von Lehrer\*innen bezieht Kooperation als wichtiges Merkmal von Professionalität ein (vgl. hierzu auch Terhart 2011) und distanziert sich von einer autonomen Berufsausübung im Sinne eines negativ konnotierten Einzelkämpfertums. In der Forschungsarbeit wird die Transformation der Fassung von pädagogischer Professionalität in Kontext der Bedeutungszunahme der Entwicklung von Einzelschulen durch die Analyse diskursiver Praktiken (Foucault 1981, Wrana 2012, 2015) zur Lehrerkooperation aus den 1960er bis 2016er Jahren untersucht.

Das Datenmaterial für die Analyse diskursiver Praktiken zur Lehrerkooperation besteht aus wissenschaftlichen Publikationen zu Diskursen der Schulentwicklung und Professionalisierung sowie Äusserungen von Lehrer\*innen über ihren Berufsalltag. Am Schnittpunkt der genannten Diskurse des wissenschaftlichen Feldes wird eine Fokussierung auf diskursive Ereignisse vorgenommen, die die Verschiebung der Konzeption von Professionalität von Lehrer\*innen genauer herausarbeitet. Für die Analyse diskursiver Praktiken (vgl. auch Langer 2008, Wrana 2015) ist die Untersuchung von unterschiedlichem Datenmaterial bedeutsam: In dieser diskursanalytischen Studie werden auch erste Ergebnisse aus der Analyse von Interviewmaterial mit Lehrer\*innen präsentiert, welche an vier österreichischen Schulen erhoben wurden.

Ziel des Beitrags ist, durch die Analyse diskursiver Praktiken in der Untersuchung von schulpädagogischen Diskursen sowie den Äusserungen im Interviewmaterial die Verschiebungen des Professionsverständnisses zu diskutieren. Die Äusserungen der Lehrer\*innen werden zu Veränderungen von aktuellen Steuerungsformen des Bildungssystems durch eine Bedeutungszunahme von Schulentwicklung in Beziehung gesetzt. Zur Diskussion gestellt werden die unterschiedlichen Positionierungen von Lehrer\*innen in ihrem Sprechen über Lehrerkooperation: Hierbei rückt das Spannungsverhältnis einer sich verändernden Institution Schule zwischen den «Vorzügen» einer bürokratischen Institution, die ein relativ ungestörtes Nebeneinanderarbeiten ermöglichte und der Entwicklung einer lernenden Organisation, die das Lehrerkollegium zunehmend in die Verantwortung nimmt, in den Fokus.

### **The concept of professionalism in studies on education policy and teacher work**

**Sølvi Mausethagen**

Oslo and Akershus University College of Applied Sciences

The concept of professionalism is often employed in studies on education policy and teacher work. However, professionalism can be construed with quite different meanings (Ozga & Lawn, 1981; Evetts, 2003; Mausethagen & Granlund, 2012). Professionalism can be used by both conservative and progressive positions to address teachers' work, to request changes or to defend the status quo (Hall, 2004). Another important aspect of this dimension of professionalism is how various actors construe the concept of professionalism differently, and how this is done differently over time. Moreover, professionalism can be studied both as a normative value system, that is, a "true" concept deriving from definitional aspects of professions as building upon a scientific knowledge base (Freidson, 2001; Abbott, 1988), and as an ideology (Larson, 1977). Evetts (2003) argues that these two perspectives are both necessary to understand occupational and organisational change within professions, and that they should be combined to enhance understanding about how the balance of normative and ideological elements varies between various actors. Distinguishing between professionalism as constructed from above and professionalism from within can assist in doing this work (Evetts, 2003). Professionalisation can take place from within if the profession can use normative aspects to construct occupational identity in ways that they can secure and maintain autonomy and discretionary power, while professionalisation from above can be employed to convince professionals to perform in ways seen as appropriate and effective. Already in 1981, Ozga and Lawn outlined how policy makers and professionals assert different agendas by using the concept of professionalism in different ways. Although these social groups use the same terminology and the conflict may seem disguised, the content is often disharmonious. This disharmony can trigger political tensions, in which the issue of professional autonomy is especially contested. Arguably then, the complex concept of professionalism needs to be located within a specific historical and national context to investigate how professionalism can be used both as control strategies (the state) and as an occupational strategy (the teachers). Accordingly, ideas about professionalism and the teacher profession are interrelated (Freidson, 2001; Gewirtz et al., 2008), and if the "quality" of the work of the teacher profession are questioned, this can result in increased professionalism from above taking place in the form of competence development and developments in teacher education, as well as accountability measures and increased external control of teachers work. Thus, tensions around state initiatives often derive from central aspects of teachers' didactic and relational work in classrooms arguably in need for a certain extent of autonomy, and national and local quests to intervene in this autonomy.

In this presentation, I will present examples on how the concept of professionalism can be used in studies on teachers' interpretations of different state initiatives to strengthen the teacher profession, taking into account both the normative and

ideological aspect of the concept. I will discuss the usefulness of the concept yet also the challenges attached to using a concept that are often used differently both among researchers, policymakers and teachers. I will also discuss issues that challenge dichotomies that are often used, such as professionalism from above and within, and organizational and occupational professionalism, for example related to how teachers often would also seek to construct legitimacy within an organisation and not only the profession, and that government policy and values in the public sector often resonate with professional ideals and thus also are valued by teachers. Use of dichotomies may then also contribute to underscore tensions where tensions are not necessarily that prominent.

## **Zu den transformatorischen Effekten pädagogischer Professionalität – Befunde einer vergleichenden Untersuchung des Umgangs von Lehrkräften mit den Instrumenten Neuer Steuerung in zwei Bundesländern**

**Corrie Thiel**

Universität Münster

Die Einführung Neuer Steuerungsinstrumente wird im wissenschaftlichen Diskurs unter Bezugnahme auf professionstheoretisches Vokabular analysiert. Eher selten werden den Instrumenten dabei Professionalisierungspotentiale zugesprochen (z.B. Bastian, Combe & Reh 2002). Weit häufiger ist demgegenüber von der Deprofessionalisierung pädagogischer Praxis durch Neue Steuerung die Rede (z.B. Beck 2008; Höhne 2011; Tacke 2005). Sowohl die Professionalisierungs- als auch die Deprofessionalisierungsdiagnosen sind dabei fast ausschließlich Ergebnis modelltheoretischer Auseinandersetzungen mit Neuer Steuerung. Der Umgang von Lehrkräften mit den Steuerungsinstrumenten wird dabei kaum in den Blick genommen. Linda Evans (2008) weist nun darauf hin, dass die Frage, inwiefern bildungspolitische Steuerung professionalisiert oder auch deprofessionalisiert, nicht allein mit Blick auf die bildungspolitischen Entwicklungen entschieden werden kann. Entscheidend ist der Umgang pädagogischer Akteure mit diesen.

Im englischsprachigen Diskurs findet sich bereits eine ganze Anzahl an Studien, die den Umgang von Lehrkräften mit Neuer Steuerung untersucht haben (z.B. Achinstein & Ogawa 2006; Ballet, Kelchtermans & Loughran 2006; Mintrop 2004, 2012; Moore et al. 2002; Olsen & Sexton 2009). Diese kommen zumeist zu dem Schluss, dass Lehrkräfte die Instrumente Neuer Steuerung keineswegs affirmativ adaptieren, sondern eigensinnig und widerständig mit diesen umgehen. In diesen Studien wird Eigensinn und Widerständigkeit dabei mehr oder weniger explizit an individuellen Merkmalen der untersuchten Lehrkräfte festgemacht. Hierüber hinausgehende Erklärungsansätze finden sich nicht.

In dem geplanten Beitrag werden die Ergebnisse einer qualitativen Studie präsentiert, die aus der Perspektive strukturtheoretischer Professionstheorie der Frage nachgegangen ist, inwiefern die Etablierung Neuer Steuerung mit der Deprofessionalisierung pädagogischer Praxis einhergeht. Hierfür wurde eine vergleichende Untersuchung in Berlin und Thüringen durchgeführt. Diese Bundesländer stellen kontrastive Fälle hinsichtlich der Art und Weise der Einführung Neuer Steuerungsinstrumente und deren Einbindung in die jeweiligen Steuerungsregime dar (Thiel, Bellmann & Schweizer 2017). Während Berlin zur Qualitätssicherung und -entwicklung auf die staatlich-administrative Kontrolle und Wettbewerb unter Schulen setzt, sind die Instrumente in Thüringen eingebunden in eine umfassendere Qualitätsentwicklungsstrategie, die primär auf die Qualifizierung und Vernetzung von Lehrkräften setzt.

In beiden Bundesländern wurden Interviews und Gruppendiskussionen zum Thema Neue Steuerung mit Lehrkräften erhoben. Insgesamt wurden 27 Interviews und 5 Gruppendiskussionen in den beiden Bundesländern durchgeführt. Diese wurden mit Hilfe der Dokumentarischen Methode (Bohnsack 2010; Nohl 2009) ausgewertet. Aus den Interviews wurden drei verschiedene Modi des Umgangs mit Neuer Steuerung rekonstruiert, aus den Gruppendiskussionen zwei. Lediglich in einem Modus, der aus den Interviews rekonstruiert wurde, werden die Defizitrückmeldungen in modelltheoretisch konformer Weise als Handlungsauftrag von Lehrkräften thematisiert. In den anderen Modi wird den Instrumenten die Bedeutung abgesprochen, schulisches Handeln auf substantielle Weise orientieren zu können. Dies geht mit der Aktualisierung professioneller Selbststeuerungs- und Kompetenzansprüche einher.

Von einer umfassenden Deprofessionalisierung pädagogischer Praxis in Kontexten Neuer Steuerung kann vor dem Hintergrund dieser Befundlage kaum die Rede sein. Man könnte auch davon sprechen, dass sich im schulischen Handlungsfeld pädagogisch-professionelle Handlungsrationitäten finden lassen, die die Deprofessionalisierung pädagogischer Praxis durch Neue Steuerung be- bzw. verhindern.

Im Rahmen der eben skizzierten Studie wurde das Konzept pädagogischer Professionalität eingesetzt, um den Umgang von Lehrkräften mit bildungspolitischen Anforderungen zu theoretisieren. Dies hat sich insofern als ertragreich erwiesen, als dass das Konzept eine Interpretationsfolie liefert, die es ermöglicht, spezifische Umgangsweisen von Lehrkräften mit den Instrumenten Neuer Steuerung nicht auf individuelle Eigenschaften einzelner Personen zurückzuführen, sondern an die berufliche Handlungspraxis zurückzubinden. Dieses Vorgehen ist nun aber auch mit Schwierigkeiten verbunden. So muss etwa festgelegt werden, wann von pädagogisch-professionellen Umgangsweisen mit Neuer Steuerung gesprochen werden soll. In Anschluss an die Darstellung der Studie werden im Vortrag sowohl die Erträge als auch die Schwierigkeiten, die mit dieser Einbindung des theoretischen Konzeptes verbunden sind, diskutiert.

## G17

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H114

### Visualisierungen in Mathematik

Chair(s): **Claudia Leopold** (Universität Fribourg, Schweiz), **Stanislaw Schukajlow** (Universität Münster)

DiskutantIn(nen): **Rolf Plötzner** (Pädagogische Hochschule Freiburg)

Visualisierungen spielen sowohl beim Verstehen von Texten (Mayer, 2009; Van Meter et al. 2006) als auch in der Mathematik eine bedeutende Rolle (Arcavi, 2003; Uesaka et al., 2007). Die Vermittlung eines angemessenen Umgangs mit Visualisierungen stellt eine wichtige Grundlage für professionelles Handeln in Lehrberufen dar. In der Mathematik werden Visualisierungen nicht nur zur Kommunikation und Präsentation quantitativer Informationen eingesetzt, sondern auch zur Förderung von Problemlöse- und Modellierungsprozessen. In diesem Symposium soll die Rolle von Visualisierungen aus unterschiedlichen Perspektiven betrachtet werden. Es werden sowohl kognitionspsychologische als auch mathematikdidaktische Ansätze integriert.

Im Beitrag von Ludewig, Lambert, Dackermann, Scheiter und Moeller "Influences of basic numerical competences on graphical literacy" wurde untersucht, welche Kompetenzen und Prozesse Einfluss auf das Verständnis graphisch präsentierter Informationen (Diagramme, Graphen etc.) nehmen. Bedeutsam ist die Differenzierung zwischen unterschiedlichen Facetten fachspezifischer Einflussfaktoren. Untersucht wurde, welchen Einfluss basale numerische Kompetenzen im Vergleich zu generellen kognitiven Fähigkeiten auf das Verständnis graphisch präsentierter Informationen haben.

Im Beitrag von Schukajlow, Rellensmann und Leopold "Erfassung des Strategiewissens über Visualisierungen beim Bearbeiten realitätsbezogener Mathematikaufgaben" steht die Erfassung des spezifischen Strategiewissens über Visualisierungen bei realitätsbezogenen Modellierungsaufgaben im Vordergrund. Im Vergleich zum ersten Beitrag wurde hier - umgekehrt - der Einfluss des Strategiewissens über Visualisierungen als ein Indikator für "graphical literacy" auf mathematische (Modellierungs)kompetenzen untersucht und mit dem Einfluss auf innermathematische Leistungen verglichen.

Während in den ersten beiden Beiträgen der Fokus auf grundlegenden Prozessen bzw. Kompetenzen liegt, werden in den folgenden beiden Beiträgen Effekte konkreter Interventionen bezüglich der von Schülern verwendeten Strategien und Aufgaben untersucht.

Im Beitrag von Thiede, Leuders und Holzäpfel "Der Prozentstreifen als Hilfsmittel beim Lösen von Textaufgaben" wird eine visuell-rechnerische Lösungsstrategie mit einer rein rechnerischen Strategie sowie einer Kontrollbedingung, die kein Strategietraining erhielt, verglichen. Erwartet wurde, dass die visuell-rechnerische Strategie insbesondere geeignet ist, die Konstruktion eines Situationsmodells der Aufgabe zu fördern.

Im Beitrag von Leopold, Hugo, Schmid und Schukajlow "Seductive Details in realitätsbezogenen Mathematikaufgaben" wurden Modellierungsaufgaben (Pythagoras) so gestaltet, dass sie einen emotionalen Fokus auf relevante oder irrelevante Informationen richteten oder keinen emotionalen Fokus enthielten. Die Untersuchung von seductive Details in Modellierungsaufgaben ist bedeutsam, weil diese die Bildung eines angemessenen Situationsmodells der Aufgabe behindern können.

Ein spezifischer Erkenntnisgewinn des Symposiums ist bezüglich der Wechselwirkungen zwischen graphical literacy und Modellierungskompetenzen sowie bezüglich der Förderung des Situations- und mathematischen Modells mittels Lösungsstrategien und Aufgabengestaltung zu erwarten.

Abschließend werden die Beiträge von Rolf Plötzner diskutiert.

#### Beiträge des Symposiums

### Influences of basic numerical competences on graphical literacy

**Ulrich Ludewig, Katharina Lambert, Tanja Dackermann, Katharina Scheiter, Korbinian Moeller**  
Universität Tübingen

Theory. "Graphs, charts, cartograms, thematic maps, etc., are common tools for handling and communicating quantitative information in contemporary society." (Åberg-Bengtsson, 2006, p. 112). During the last ten years, graphically presented data have played an increasingly important role in our life. In the present study we investigated the influence of basic numerical competences on students' ability to understand graphically presented information – that is graphical literacy.

Method. We analyzed data of 750 German students (grades 9 – 11) and evaluated the determinants of graphical literacy with multiple regression analysis using predictors of basic numerical competences (such as number line estimation, basic arithmetic operations, etc.) while controlling for influences of general cognitive ability, age, gender and various interaction terms.

Results. The multiple regression analysis was able to explain about 34 % ( $R^2 = .34$   $F(30,720) = 12.61$ ,  $p < .001$ ) of the variance in the graphical literacy test score. We observed that number line estimation ( $p < .001$ ), subtraction ( $p < .001$ ) and conceptual number knowledge ( $p = .026$ ) were significant predictors of graphical literacy beyond the influences of general cognitive ability ( $p = .001$ ) using a FDR p-value adjustment for multiple testing.

Discussion. These predictors of graphical literacy closely map with the distinction between arithmetical and non-arithmetical processes described by the Mixed Arithmetic-Perceptual Model of graph reading (Gillan, 1994). Furthermore, considering how number line estimation is achieved (by means of proportion-based strategies; Dackermann et al., 2015) and considering the observed influence of students' conceptual number knowledge, we conclude that the influence of basic numerical competences goes beyond the level of providing mere numerical and arithmetical prerequisites for more applied everyday graphical literacy tasks. Instead, influences of number line estimation and conceptual number knowledge indicate that mastery of basic numerical concepts allows for the application of more efficient problem solving strategies beyond mere arithmetical procedures. In future studies, it may be interesting to evaluate how basic numerical competences influence students' tendency to apply efficient short-cut strategies (e.g. by building a visual mean between two bars; Gillan, 1995) instead of mental arithmetic (by calculating the arithmetic mean).

In sum, it is remarkable that basic numerical competences acquired in primary school significantly contribute to later performance in complex real life graphical literacy problems. However, in secondary education, teachers usually expect students to have acquired such basic numerical competences. Therefore, it may be worth reviewing elementary school knowledge to improve students' performance in real-world mathematical problems.

## **Erfassung des Strategiewissens über Visualisierungen beim Bearbeiten realitätsbezogener Mathematikaufgaben**

**Stanislaw Schukajlow<sup>1</sup>, Johanna Rellensmann<sup>1</sup>, Claudia Leopold<sup>2</sup>**

<sup>1</sup>Universität Münster, <sup>2</sup>Universität Fribourg

Theorie. Das Strategiewissen spielt eine wichtige Rolle bei der Bearbeitung von realitätsbezogenen Mathematikaufgaben, die wiederum hohe Bedeutung für die Lebens- und Berufswelt der Lernenden haben (Blum et al., 2007). So zeigen verschiedene Befunde, dass das Strategiewissen Mathematikleistungen über den Einfluss von kognitiven Grundfähigkeiten und Interesse hinaus vorhersagen kann (Lingel et al. 2014). In theoretischen Ansätzen wird oft zwischen relationalem und spezifischem Strategiewissen unterschieden (Borkowski et al. 2000). Bisher wurde vor allem die Bedeutung des relationalen Strategiewissens untersucht, welches das Wissen darüber umfasst, wann eine Strategie im Vergleich zu anderen Strategien angewandt werden sollte. Hingegen wurde das spezifische Strategiewissen seltener untersucht. Das spezifische Strategiewissen beinhaltet das Wissen darüber, wie die entsprechende Strategie angewandt werden soll. Im Vortrag werden zwei Studien präsentiert, in denen die Validität einer Skala zur Erfassung des spezifischen Strategiewissens über Visualisierungen beim Bearbeiten von realitätsbezogenen Mathematikaufgaben (Modellierungsaufgaben) überprüft wurde. Visualisierungsstrategien standen im Fokus, weil sie besonders beim mathematischen Modellieren bedeutsam sind (Rellensmann et al. 2017).

Fragestellung. Wie kann spezifisches Strategiewissen über Visualisierungen valide erfasst werden?

Methode. Im selbsterstellten Visualisierungstest wurde sowohl das Wissen über Visualisierungen der in der Aufgabe beschriebenen Situation als auch des mathematischen Modells erfasst. Die Skala umfasste 16 Items (siehe Abb. 1). Das Strategiewissen bezog sich auf die Akkuratheit der Visualisierung bzgl. der mathematischen Lösungsstruktur und auf die Vollständigkeit der Zuordnung der Zahlenangaben zu den jeweiligen Objekten. In einer qualitativen Studie wurde der Inhaltsvalidität und in einer qualitativen Studie wurden die strukturelle, prädiktive und diskriminante Validität der Skala geprüft. In der qualitativen Studie wurde die Skala acht Neuntklässlern präsentiert und sie wurden im Rahmen eines Interviews zu den Gründen für die Auswahl eines Items befragt. Die Schülerantworten wurden mittels Inhaltsanalyse in Anlehnung an die von diSessa (2002) entwickelten Kategorien Akkuratheit, numerische Vollständigkeit und Ästhetik der Repräsentation analysiert (Interrater-Übereinstimmung Cohens  $\kappa = .88$ ).

An der quantitativen Studie nahmen 146 Neuntklässler (91 weiblich,  $M = 14,8$  Jahre) teil. Die Lernenden bearbeiteten einen Test zu kognitiven Grundfähigkeiten (KFT4-12), den Strategiewissenstest, gaben ihr Interesse in Mathematik an und lösten innermathematische Aufgaben. Im Anschluss erhielten sie acht Modellierungsaufgaben mit der Aufforderung, jeweils eine Skizze vor der Aufgabenbearbeitung zu erstellen, so dass sowohl ihre Modellierungsfähigkeit als auch die Qualität der gezeichneten Skizzen analysiert werden konnten. Alle Modellierungsaufgaben konnten mit dem Satz des Pythagoras gelöst werden. Die Reliabilität aller Skalen war gut. IRT (ConQuest) wurde verwendet, um Itemstatistiken und Dimensionalität der Strategiewissensskala zu prüfen.

Ergebnisse und Diskussion. Die Analyse der Begründungen von Lernenden zeigte, dass sich Lernende beim Bearbeiten des Tests auf die intendierten Kriterien Akkuratheit und numerische Vollständigkeit der Repräsentation bezogen. Die quantitative Analyse zeigte darüber hinaus, dass der Test gute WLE-Reliabilitäten von .83 hatte und alle Items – bis auf ein Item – hinreichend gute Kennwerte aufwiesen ( $0.8 < MNSQ < 1.2$ ;  $T < 1.96$ ;  $r > .30$ ). Eine konfirmatorische Faktorenanalyse bestätigte, dass das Strategiewissen als eindimensionales Konstrukt modelliert werden kann und die eindimensionale Modellierung der zweidimensionalen Modellierung – mit separaten Dimensionen für situationales und mathematisches Strategiewissen – überlegen war. Unter Kontrolle kognitiver Grundfähigkeiten und des Interesses war das Strategiewissen über Visualisierungen entsprechend unserer Annahmen prädiktiv für die Qualität der gezeichneten Skizzen ( $r = .29$ ,  $p < .01$ ) und für die Modellierungsleistungen ( $r = .19$ ,  $p < .01$ ). Die Zusammenhänge zwischen dem Strategiewissen und den innermathematischen Leistungen waren hingegen nicht signifikant ( $r = .01$ ), sodass auch die Annahme zur diskriminanten Validität der entwickelten Skala bestätigt werden konnte.

Zusammenfassend lässt sich festhalten, dass eine valide Erfassung von spezifischem Strategiewissen über Visualisierungen möglich ist. Mit Hilfe des entwickelten Messinstruments kann untersucht werden, in welchem Zusammenhang spezifisches Strategiewissen mit anderen Konstrukten steht und welchen Einfluss das spezifische Strategiewissen auf die Strategienutzung und die Mathematikleistungen hat.

## **Der Prozentstreifen als Hilfsmittel beim Lösen von Textaufgaben**

**Benjamin Thiede, Timo Leuders, Lars Holzäpfel**

Pädagogische Hochschule Freiburg

Theorie. Schüler der Sekundarstufe I erlernen unterschiedliche Strategien zum Lösen von Prozenttextaufgaben. Die häufigsten sind der Dreisatz- oder Formelanwendungen. Dies bestätigen unter anderem Kleine & Jordan (2007), die 795 Schüler der Klassen 8 bis 10 hinsichtlich auftretender Lösungsstrategien untersuchten. Sie ergänzen darüber hinaus individuelle (arithmetische), oft fehlerhafte Lösungsstrategien als weitere Kategorie. In diesem Kontext werden jedoch immer wieder unzureichende Lösungsraten verzeichnet (Berger, 1991; Hafner, 2012), auch abhängig vom Aufgabentyp selbst (Grundwert, Prozentwert oder Prozentsatz gesucht). Typische Fehler gründen dabei häufig auf Verständnisproblemen des Prozentkonzeptes selbst und Rechenproblemen, wie die Wahl der richtigen Rechenoperation (Dole et al., 1997; Rianaseri, 2012). Die Ergebnisse hinsichtlich der Schwierigkeiten in der Prozentrechnung können unter anderem durch die Notwendigkeit der Bildung eines geeigneten Situationsmodells, quasi als erstem Lösungsschritt erklärt werden. An dieser Stelle bedarf es der Identifizierung einer passenden Rechnung, die die grundlegende Struktur der Aufgabe widerspiegelt (z.B. Erhöhung, Verminderung etc.). Als vielversprechende, das Bilden des Situationsmodells unterstützende Strategie gilt dabei das Zeichnen einer Skizze, ferner eines Diagrammes, welches die Einsicht in die Aufgabenstruktur erleichtert (Arcavi, 2003; Van Essen & Hamaker, 1990). Eine geeignete Visualisierung, die neben den verschiedenen Größen und Prozentangaben auch deren Verhältnisse darstellt, ist der Prozentstreifen. Auf diesem werden die gegebenen und gesuchten Größen markiert, so dass die relationalen Zusammenhänge deutlich(er) werden. Verschiedene Fallstudien weisen darauf hin, dass der Prozentstreifen (auch: bar diagram oder percentage bar) das Verstehen der Situation und die Lösung der Textaufgabe unterstützen können (Hoven & Garelick, 2007; Van den Heuvel-Panhuizen, 2003).

Auf Grundlage dieser Ergebnisse wurde eine Unterrichtseinheit konzipiert, bei der Schüler angehalten wurden, den Prozentstreifen als visuelle Lösungsstrategie (zusätzlich zu den bereits erlernten Strategien) anzuwenden. Daraus resultieren folgende Forschungsfragen: (1) Inwiefern nutzen Schüler den Prozentstreifen als Lösungsstrategie bei Textaufgaben? (2) Hat die Anwendung des Prozentstreifens einen positiven Einfluss auf die Lösungsraten?



Methode. In einem quasi-experimentellen Design wurden Schüler der Experimentalgruppe einem 45-minütigen Prozentstreifen-Training unterzogen. Ebenso wurde der notwendige zweite Lösungsschritt, hier das Rechnen des Dreisatzes, integriert. Eine zweite Gruppe erhielt ein Dreisatz-Training (Wiederholung), ohne visuelle Unterstützung und eine dritte Gruppe erhielt kein Training, was zu folgenden Bedingungen führte: (a) Prozentstreifengruppe (PSG, N=86), (b) Dreisatzgruppe (DSG, N=98) und (c) Warte-Kontrollgruppe (KG, N=82). Die Schüler entstammten 15 unterschiedlichen Klassen der Stufen 7 und 8, unter der Prämisse, dass das Prozentrechnen bereits eingeführt wurde. Sie stammen aus den Schultypen Werkrealschule und Realschule (N=266,  $\bar{x}$  13,95 Jahre; 47 % weiblich / 53 % männlich). Die Fähigkeit entsprechende Aufgaben zu lösen wurde in einem Test mit fünf struktur- und kontextheterogenen Textaufgaben getestet (Cohen's  $\kappa = .689$  (pre),  $.655$  (post)).

Ergebnisse. (1) Zum einen ist festzuhalten, dass Dreisatz und Formel die prävalenten Lösungsstrategien sind. Zum anderen und im Hinblick auf die Schüler der Prozentstreifengruppe, ist zu konstatieren, dass diese die visuelle Lösungsstrategie im direkten Vergleich (Pre-Post) in fast 2/3 der Aufgaben (60,4%) nutzen. Im Follow-up-Test sinkt diese Nutzung zwar auf 33,5%, stellt aber immer noch die häufigste, angewandte Lösungsstrategie dar.

(2) Die Lösungsraten beider Experimentalbedingungen steigen hinsichtlich des direkten Lernzuwachses signifikant. Entgegen unserer Erwartungen zeigen sich hier keine Unterschiede zwischen diesen beiden Gruppen. Ein anderes Bild zeigt sich im Pre-Follow-up-Test-Vergleich. Dort profitieren insbesondere leistungsschwächere Schüler vom Prozentstreifen.

Diskussion. Zu diskutieren ist unter anderem die Frage, warum der Prozentstreifen im Vergleich zur rein rechnerischen Strategie im direkten Lernzuwachs keinen Vorteil hat, dafür aber einen nachhaltigen Effekt erzielt. Das könnte mit der Kürze der Intervention zusammenhängen, weist jedoch darauf hin, dass der Prozentstreifen stark rekonstruktiv agiert, während eine reine Wiederholung des Dreisatzes kaum Erfolg birgt. In diesem Kontext ist die mögliche Rolle von Moderatoren und ATI-Effekten zu klären. Als Limitation ist die Randomisierung auf Klassenebene zu nennen, da die Ergebnisse möglicherweise mit Klassenmerkmalen konfundieren.

### **Seductive Details in realitätsbezogenen Mathematikaufgaben**

**Claudia Leopold<sup>1</sup>, Anja Hugo<sup>1</sup>, Eva Schmid<sup>1</sup>, Stanislaw Schukajlow<sup>2</sup>**

<sup>1</sup>Universität Fribourg, <sup>2</sup>Universität Münster

Theorie. In diesem Projekt wurde untersucht, welche Wirkungen seductive details auf das Bearbeiten realitätsbezogener Mathematikaufgaben haben. Bei seductive details handelt es sich um interessante aber irrelevante Informationen, die Lernmaterialien hinzugefügt werden (Garner et al., 1989), um sie ansprechender zu gestalten. Etliche Befunde zeigen lernhinderliche Wirkungen für Lernmaterialien mit seductive details (Mayer, 2009; Rey, 2012). Unklar ist, wie seductive details wirken. Einerseits kann die Aufmerksamkeit der Lernenden durch affektive Elemente von relevanten Inhalten abgelenkt werden. Andererseits können irrelevante Informationen zu einer Überlastung des Arbeitsgedächtnisses führen. In beiden Fällen sollte die Bildung eines Situationsmodells der Aufgabe beeinträchtigt werden. Um potentielle Einflussfaktoren zu untersuchen, wurde der emotionale Fokus in realitätsbezogenen Modellierungsaufgaben variiert. Die Forschungsfrage lautet: Wie wirkt ein emotionaler Fokus auf relevante oder irrelevante Informationen im Vergleich zu einer Bedingung ohne emotionalen Fokus auf die Modellierungsleistung? Welche Rolle spielt das Zeichnen der in der Aufgabe beschriebenen Inhalte, welches sich - wie andere Befunde zeigen (Rellensmann et al., 2017) - förderlich auf die Modellierungsleistung auswirken kann?

Methode. Versuchspersonen und Design: 46 Schüler/innen der 9. Jahrgangsstufe (M = 13.8 Jahre; 54.3% weiblich), nahmen an Studie 1 und 47 an Studie 2 (M = 13.8 Jahre; 59.6% weiblich) teil. Beiden Studien lag ein within-subject-Design zugrunde, in welchem die Art der Aufgabe in drei Stufen variiert wurde (ohne emotionalen Fokus, Fokus auf relevante Info, Fokus auf irrelevante Info). Während in Studie 1 keine spezifische Instruktion zur Aufgabenbearbeitung gegeben wurde, wurden die Schüler/innen in Studie 2 zum Zeichnen einer Skizze pro Aufgabe aufgefordert.

Material und Vorgehensweise: Sechs Modellierungsaufgaben wurden so gestaltet, dass jede/r Schüler/in jeweils zwei Aufgaben ohne emotionalen Fokus, mit Fokus auf relevante Informationen und mit Fokus auf irrelevante Informationen bearbeitete. Vorlagen für die Aufgaben wurden von Rellensmann et al. (2017) übernommen und entsprechend den experimentellen Bedingungen adaptiert. Als Kontrollvariablen wurden themenspezifisches Vorwissen, innermathematische Fähigkeiten und verbale Fähigkeiten (KFT, Wortschatz) erfasst. Nach Bearbeitung der Vorwissenstests und des verbalen Fähigkeitstests, erhielten die Lernenden eine Instruktion zur Bearbeitung von Modellierungsaufgaben (in Studie 2 zusätzlich eine Anleitung zum Skizzieren). Im Anschluss bearbeiteten die Schüler/innen die Modellierungsaufgaben sowie innermathematischen Aufgaben.

Ergebnisse und Diskussion. In Studie 1 weisen die Ergebnisse einer Varianzanalyse keinen Effekt für die Art der Aufgabe aus,  $F(2, 90) < 1$ . Es konnte kein Einfluss des emotionalen Fokus auf die Modellierungsleistung festgestellt werden. Wenn Kontrollvariablen in die Analyse einbezogen werden, ändert sich die Ergebnisse nicht. Es liegen keine Interaktionen zwischen Kontrollvariablen und der Art der Aufgabe vor,  $F_s(2, 84) < 1$ .

In Studie 2 (Zeicheninstruktion) weisen die Ergebnisse einen Effekt der Art der Aufgabe aus,  $F(2, 92) = 4.78$ ,  $p = .011$ . Der emotionale Fokus der Aufgabe beeinflusste die Modellierungsleistung. Follow-up-Analysen weisen eine bessere Leistung für die Bedingung ohne emotionalen Fokus im Vergleich zur Bedingung mit emotionalem Fokus auf relevante Informationen aus,  $t(46) = 3.57$ ,  $p = .001$ . Kein Unterschied zeigte sich zwischen der Bedingung emotionaler Fokus auf irrelevante Aufgaben und den anderen Bedingungen,  $t(46) = 1.83$ ,  $p = .074$  bzw.  $t(46) = 1.11$ ,  $p = .274$ . Wenn Kontrollvariablen in die Analyse einbezogen werden, ändern sich diese Ergebnisse nicht. Es liegen keine signifikanten Interaktionen mit Kontrollvariablen vor.

Insgesamt zeigen die Ergebnisse einen seductive-details-Effekt, allerdings nur, wenn eine Zeicheninstruktion gegeben wurde. Die Aufgabe des Skizzierens stellt offensichtlich Bedingungen her, die die Wirkung von seductive details zur Geltung bringen. In diesem Kontext wirkt ein emotionaler Fokus auf relevante, nicht jedoch auf irrelevante Informationen, lernhinderlich. Unsere Ergebnisse zeigen, dass seductive details nicht als solche die Lernergebnisse beeinträchtigen, sondern vor allem dann, wenn sie von relevanten Information ablenken. Insofern stützen sie die Annahme, dass die Aufmerksamkeit der Lernenden durch seductive details von relevanten Informationen abgelenkt werden kann.

## G18

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H102

### **Verständnis über arithmetisches Lernen vom Kindergarten bis zur Sekundarstufe – Relevanz der Einflussfaktoren Arbeitsgedächtnis und Sprache**

Chair(s): **Anja Schröder** (TU Dortmund, Deutschland), **Alexander Röhm** (TU Dortmund, Deutschland)

DiskutantIn(nen): **Evelyn Hang** (TU Dortmund, Deutschland)

Die Beiträge des Symposiums befassen sich mit der zentralen Fragestellung, wie die Relevanz der Einflussfaktoren Arbeitsgedächtnis und Sprache auf arithmetisches Lernen über die Entwicklungsspanne von Vorschulalter bis zur Sekundarstufe beschrieben werden kann.

Grundlegend für diese Fragestellung sind folgende Annahmen: Arithmetisches Lernen nimmt innerhalb des mathematischen Lernens in der Schule eine zentrale Schlüsselposition ein, weil es entscheidend für die Bewertung des Bildungserfolges ist. Dabei beginnt die Entwicklung arithmetischen Lernens nicht erst in der Schule, sondern weit vorher im Säuglings- (z. B. Subitizing) und frühen Kleinkindalter (z. B. Erwerb erster Zahlwörter) (Clements, 1999; Grube, 2006). Der Bedarf der Förderung arithmetischer Kompetenzen bei deutschen Schülern und Schülerinnen wird u.a. in der jüngsten TIMS-Studie (2015) deutlich. Im Trendvergleich von 2007, 2011 und 2015 wird zudem deutlich, dass Arithmetik der mathematische Inhaltsbereich ist, der Grundschulkindern in Deutschland vor die größten Herausforderungen stellt: Weder im Bereich Geometrie/Messen noch im Bereich Umgang mit Daten zeigten sie in allen drei Erhebungszyklen im Durchschnitt schlechtere Leistungen (Selter, Walter, Walther & Wendt, 2016). Aufgrund dessen ist es erforderlich Ursachen für das schlechte Abschneiden in den arithmetischen Kompetenzen zu erforschen.

Zentrale Einflussfaktoren auf arithmetisches Lernen sind sprachliche Kompetenzen (z. B. Fazio, 1994) und Arbeitsgedächtnisleistungen (z. B. Gathercole et al., 2004). Allerdings zeigt die bisherige Forschung zu den beiden Einflussfaktoren sehr heterogene Befunde: Diese sind erstens zurück zu führen auf die forschungsmethodische Konzeption der Studien wie z. B. die Auswahl der Probanden (Alter, Entwicklungsvoraussetzungen der Lernenden), die abgefragten arithmetischen Kompetenzen (Faktenwissen, prozedurales Wissen, konzeptuelles Wissen) und die Spezifizierung der untersuchten Parameter für das Arbeitsgedächtnis sowie die Sprache. Zweitens zeigen bisherige Befunde nur Ausschnitte der Wirkrichtung von Sprache und Arbeitsgedächtnis auf arithmetische Kompetenzen, so dass die Systematisierung vorliegender Ergebnisse noch aussteht.

Zu diesem Zweck wurde ein interdisziplinäres Netzwerk gegründet, das aus Teilnehmenden mit wissenschaftlicher Expertise in Mathematikdidaktik, Rehabilitationswissenschaften, Psychologie, Neurowissenschaften und Erziehungswissenschaften besteht.

Aus der Arbeit dieses Netzwerks werden exemplarisch insgesamt fünf Studien in diesem Symposium vorgestellt: Im ersten Vortrag werden Hirnentwicklungsprozesse vorgestellt, um daran die neurophysiologische Zusammenhänge zwischen Sprach- und Mathematikleistungen besser zu verdeutlichen.

Der zweite Beitrag befasst sich mit einer Studie zu dem Einfluss von Arbeitsgedächtnisleistungen auf mathematisches Lernen, wobei der schnelle Faktenabruf als Mediator berücksichtigt wird.

Der dritte Beitrag umfasst drei Einzelstudien unter der zentralen Fragestellung, welche Zusammenhänge sich zwischen Sprachkompetenz, mathematischen Fähigkeiten sowie der Leistungsfähigkeit des Arbeitsgedächtnisses bei Kindern im Vorschul- und Grundschulalter mit und ohne Spracherwerbsstörungen zeigen.

*Beiträge des Symposiums*

### **Das Arbeitsgedächtnis als wichtiger Motor für die Entwicklung von Sprach- und Mathematikleistungen – Evidenz aus den Neurowissenschaften**

**Anja Fengler**

Universität Leipzig, Deutschland

Dank bildgebender Verfahren wie der Magnetresonanztomographie sind wir in der Lage neurophysiologische Entwicklungsprozesse im Gehirn zu identifizieren, die dem Erwerb kognitiver Entwicklungsmeilensteine zugrunde liegen. Gemeinsam zeigen neurowissenschaftliche und entwicklungspsychologische Erkenntnisse Parallelen in der Hirnentwicklung und der Erwerbsreihenfolge verschiedener kognitiver Kompetenzen (z.B. Diamond, 2001; Shaw et al., 2006). Im Kleinkind- und Schulalter existieren dabei wichtige Zeitpunkte, die unterschiedliche Lernprozesse möglich machen (Galvan, 2010). Beispielsweise sind perzeptuelle Fähigkeiten wie das Sehen auf stimulierenden Input innerhalb spezifischer früher Zeitfenster angewiesen, um die relevanten Hirnregionen für ihre Funktionen zu spezialisieren (Greenough, Black, & Wallace, 1987). Komplexere Verarbeitungsprozesse wie die Satzverarbeitung und das arithmetische Lernen hingegen erfordern die Kommunikation verschiedener Hirnareale, deren Verbindungen z.T. erst im Schulalter ausreichend reifen (Lebel et al., 2008). Um neurophysiologische Zusammenhänge zwischen Sprach- und Mathematikleistungen besser verstehen zu können, werde ich in meinem Vortrag diese wichtigen Hirnentwicklungsprozesse kurz darstellen. Insbesondere werde ich dabei hirstrukturelle und hirnfunktionelle Veränderungen beschreiben, die mit dem Erwerb kognitiver Fähigkeiten einhergehen, die sowohl für sprachliche als auch mathematische Kompetenzen bedeutsam sind. Denn dabei zeigt sich, dass vor allem die Reifung und Nutzung von Hirnstrukturen, die mit verschiedenen Arbeitsgedächtnisfähigkeiten assoziiert sind, für bessere Sprach- (Fengler, Meyer & Friederici, 2015, 2016) und Mathematikleistungen notwendig sind (Menon, 2010).

### **Wird der Zusammenhang zwischen dem Arbeitsgedächtnis und komplexen Mathematikleistungen durch die Verfügbarkeit arithmetischen Faktenwissens mediert?**

**Janin Brandenburg<sup>1</sup>, Christina Jörns<sup>2</sup>, Kirsten Schuchardt<sup>2</sup>, Marcus Hasselhorn<sup>1</sup>**

<sup>1</sup>Deutsches Institut für Internationale Pädagogische Forschung, Deutschland, <sup>2</sup>Universität Hildesheim, Deutschland

Theoretischer Hintergrund. Nach Baddeley (1986) besitzt das Arbeitsgedächtnis drei strukturelle Komponenten mit unterschiedlichen Funktionen: Während der visuell-räumliche Notizblock und die phonologische Schleife als modalitätsspezifische Speichersysteme für die kurzzeitige Aufrechterhaltung von Informationen zuständig sind, ist die zentrale Exekutive als übergeordnete und modalitätsunabhängige Steuerungszentrale konzipiert. Das Arbeitsgedächtnis leistet einen wesentlichen Beitrag zur Entwicklung mathematischer Leistungen bei Grundschulkindern. Dennoch ist die unterschiedliche

Relevanz der einzelnen Arbeitsgedächtniskomponenten für mathematisches Lernen noch unzureichend geklärt: Während sich in den meisten Studien die zentrale Exekutive übereinstimmend als relevanter Prädiktor herauskristallisiert (z. B. Bull & Scerif, 2001; Gathercole et al., 2004), wird der differentielle Einfluss der beiden Speichersysteme noch kontrovers diskutiert (z. B. Alloway & Passolunghi, 2011). Ebenfalls ungeklärt ist, ob das Arbeitsgedächtnis als ein eher unspezifischer Prädiktor einen direkten Einfluss auf komplexe Mathematikleistungen ausübt oder ob der Zusammenhang durch bereichsspezifisches Grundlagenwissen wie etwa dem basalen arithmetischen Faktenwissen mediiert wird. Unter dem basalen arithmetischen Faktenwissen versteht man die Fähigkeit die Ergebnisse einfacher Additions- und Subtraktionsaufgaben im Zahlenraum bis 20 nicht mehr aktiv berechnen zu müssen, sondern die Lösung aus dem Langzeitgedächtnis direkt abzurufen. Die Geschwindigkeit, mit der wir auf dieses Wissen zugreifen können, gilt mittlerweile als zentrale Grundlagenfertigkeit für das Lösen komplexer mathematischer Leistungen (Geary, 2004).

Fragestellung. Die Studie geht in einem längsschnittlichen Design der Frage nach, inwieweit der Zusammenhang zwischen dem Arbeitsgedächtnis und komplexen Mathematikleistungen (d.h. Sachrechnen, Geometrie und Arithmetik im Zahlenraum bis zu einer Million) durch die Geschwindigkeit des basalen arithmetischen Faktenabrufs mediiert wird.

Methode. An der Studie nahmen 100 lernunauffällige Kinder (50 Mädchen, 50 Jungen) teil. Zum ersten Messzeitpunkt (Mitte der 3. Klasse) bearbeiteten die Kinder Aufgaben zur phonologischen Schleife, zum visuell-räumlichen Notizblock und zur zentralen Exekutive des Arbeitsgedächtnisses. Beim zweiten Messzeitpunkt (Anfang der 4. Klasse) wurde die Geschwindigkeit des basalen arithmetischen Faktenabrufs im Zahlenraum bis 20 mit Hilfe eines Speedtests (DIRG; Grube, Weberschock, Blum & Hasselhorn, 2010) untersucht. Am Ende der 4. Klasse wurden die komplexen Mathematikleistungen der Kinder curricular valide mit dem Deutschen Mathematiktest für 4. Klassen (DEMAT 4+; Roick, Göllitz, & Hasselhorn, 2006) erhoben. Durch das längsschnittlich angelegte Design lässt sich die Mediationshypothese gerichtet interpretieren, weil zwischen der Messung des Arbeitsgedächtnisses und des basalen arithmetischen Faktenabrufs ein halbes Schuljahr verstrichen ist.

Ergebnisse. Zunächst wurde mit Hilfe eines Strukturgleichungsmodelles untersucht, welche Teilkomponenten des Arbeitsgedächtnisses einen prädiktiven Beitrag zur Erklärung der komplexen Mathematikleistungen am Ende der 4. Klasse leisten. Bedeutsam für die Vorhersage erwiesen sich dabei die zentrale Exekutive und der visuell-räumliche Notizblock, jedoch nicht die phonologische Schleife des Arbeitsgedächtnisses. Im zweiten Schritt wurde die Leistung im basalen arithmetischen Faktenabruf als ein möglicher Mediator der Zusammenhänge ins Modell aufgenommen: Während der direkte Einfluss der zentralen Exekutive auf die Leistungen im DEMAT 4+ erhalten blieb, zeigte der visuell-räumliche Notizblock lediglich einen indirekten Beitrag auf das mathematische Lernen, indem sein Einfluss vollständig durch die Verfügbarkeit des basalen arithmetischen Faktenwissens mediiert wurde. Die Ergebnisse werden vor dem Hintergrund des Arbeitsgedächtnismodells von Baddeley (1986) diskutiert.

## **Einflüsse sprachlicher Kompetenzen und Arbeitsgedächtnisleistungen auf die Entwicklung mathematischen Lernens im Vor- und Grundschulalter**

**Nurit Viesel-Nordmeyer, Alexander Röhm, Anja Starke, Anja Schröder, Ute Ritterfeld**  
TU Dortmund, Deutschland

Theoretischer Hintergrund. Sprachliche Kompetenzen stellen sowohl in der kindlichen Entwicklung wie auch im schulischen Kontext eine wesentliche Schlüsselqualifikation dar (Ritterfeld et al. 2013; Willems, Bos, Wilmans & Platz, 2013). Im Bereich des mathematischen Lernens nimmt Sprache unterschiedliche Funktionen ein, welche sich in Erwerb, Vermittlung und Speicherung abbilden (Schröder & Ritterfeld, 2014). Da nicht erst in schulischen Bildungsprozessen, sondern bereits beim vorschulischen Erwerb grundlegender mathematischer Lernprozesse Sprache eine große Bedeutung zukommt (z. B. Ritterfeld et al. 2013), zeigt sich der enge Zusammenhang zwischen sprachlichen Kompetenzen und mathematischem Lernen insbesondere bei Kindern, deren Sprache nicht altersgemäß entwickelt scheint. So zeigen beispielsweise Kinder mit sogenannten umschriebenen Sprachentwicklungsstörungen (USES) bereits deutliche Schwächen im Bereich mathematischer Basiskompetenzen (z. B. Fazio, 1994). Zahlreiche Studien belegen bei unauffälligen sowie auffälligen Kindern sowohl für sprachliche Fähigkeiten (z. B. Gathercole, Willis, Emslie & Baddeley, 1992; Graf Estes, Evans, & Else-Quest, 2007) als auch mathematische Kompetenzen (z. B. Friso van den Bos, van der Ven, Kroesbergen & van Luit, 2013) einen deutlichen Einfluss des Arbeitsgedächtnisses. Der Zusammenhang aller drei Komponenten – Sprache, Mathematik und Arbeitsgedächtnis – wurde bislang nur vereinzelt gemeinsam untersucht und kann daher nicht ausreichend erklärt werden. Zwar erwies sich das Arbeitsgedächtnis im Vergleich zur Intelligenz als stärkerer Prädiktor späterer Schulleistungen (z. B. Alloway & Alloway, 2010; Krajewski, Schneider & Nieding, 2008), dennoch gilt es zur Klärung bestehender Interdependenzen innerhalb der Entwicklung dieser drei Kompetenzen auch die Rolle der Intelligenz weiter zu untersuchen.

Fragestellungen. In diesem Beitrag werden daher anhand der Befunde von drei Studien folgende Fragestellungen untersucht:

1. Welche Zusammenhänge zeigen sich zwischen Sprachkompetenz, mathematischen Fähigkeiten sowie der Leistungsfähigkeit des Arbeitsgedächtnisses bei Kindern im Vorschul- und Grundschulalter?
2. Welche sprachlichen, mathematischen und kognitiven Fähigkeiten (inklusive einzelnen Arbeitsgedächtnisprozessen) gehen sprachlichen und mathematischen Fertigkeiten der Schulanfangsphase voraus?
3. Zeigen sich Unterschiede in den Zusammenhängen bei Kindern mit und ohne sprachliche Auffälligkeiten?

Methode. In Studie 1 (vgl. Röhm, Starke & Ritterfeld, 2017) wurden unter Anwendung eines Within-subject-Designs bei 30 Kindergartenkindern ohne sprachliche Auffälligkeiten Sprachentwicklung, mathematische Basiskompetenzen, nonverbale Intelligenz und Arbeitsgedächtnis mit standardisierten Testverfahren erhoben.

Studie 2 untersuchte ebenfalls die Sprachentwicklung, mathematische Basiskompetenzen, nonverbale Intelligenz und Arbeitsgedächtnis sowie die phonologische Bewusstheit von 48 Kindern mit USES am Ende des ersten Grundschuljahres mit standardisierten Tests.

Anhand der monolingual aufwachsenden Kinder der Startkohorte Kindergarten (SC2) des Nationalen Bildungspanels (N = 268) der Altersspanne vom Kindergarten (4/5-Jahre) bis zum Primarbereich (1. Klasse) werden in Studie 3 die Zusammenhänge zwischen Sprache, Mathematik und kognitiven Fähigkeiten längsschnittlich betrachtet.

Ergebnisse. Die Ergebnisse verdeutlichen zahlreiche Auswirkungen der Sprachentwicklung wie auch unterschiedlicher sprachlicher Vorläuferfähigkeiten auf mathematisches Lernen. In allen drei Studien zeigen sich indirekte Einflüsse der phonologischen Schleife als Mediation über die Sprachkompetenz der Kinder auf mathematische Kompetenzen. Die Zentrale Exekutive wiederum beeinflusst mathematische Kompetenzen im Vor- und Grundschulalter direkt. Längsschnittlich stellen sich in Studie 3 komplexe Wechselbeziehungen zwischen Arbeitsgedächtnis und Sprachentwicklung dar. Zudem kann mathematisches Vorwissen spätere Leistungen der Kinder erklären.

Diskussion. Die Ergebnisse der vorliegenden Untersuchung leisten einen Beitrag zum Verständnis der komplexen kognitiven Strukturen und deren Einfluss auf die Entwicklung sprachlicher und mathematischer Kompetenzen. Dies ermöglicht eine Gegenüberstellung der mathematischen Kompetenzen und Arbeitsgedächtnisleistungen von Kindern mit und ohne Spracherwerbsstörungen. Besonders im mathematischen Bereich zeigen die Arbeitsgedächtnisleistungen einen multifaktoriellen Einfluss in dessen Rahmen sich sprachliche Kompetenzen auf unterschiedlichen Ebenen als Schlüsselqualifikation darstellen. Die Bedeutung derartiger Kenntnisse für die Entwicklung von Interventionsmöglichkeiten wird diskutiert.

## G19

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H001

### Förderung von Diagnosekompetenzen mit Simulationen an der Hochschule

Chair(s): **Nicole Heitzmann** (LMU München, Deutschland), **Frank Fischer** (LMU München, Deutschland)

DiskutantIn(nen): **Anna Südkamp** (TU Dortmund)

In vielen beruflichen Handlungsfeldern gilt Diagnostizieren als Voraussetzung für das erfolgreiche Bewältigen von professionellen Aufgaben, z.B. das Identifizieren einer passenden Therapie für einen Patienten. Auch in der Schule spielen Diagnosen von Lernvoraussetzungen, Schülerleistungen oder auch von Unterrichtsmerkmalen im Klassenzimmer eine wichtige Rolle für das professionelle Handeln. An der Hochschule werden viele angehende Fachkräfte unzureichend auf diese spätere Anforderungssituation vorbereitet, da es wenige Gelegenheiten gibt, Wissen in diagnostischen Situationen anzuwenden. Studien aus der Medizindidaktik zeigen, dass Simulationen zur Lösung dieses Problems beitragen könnten.

In dieser interdisziplinären Arbeitsgruppe mit Forscher/innen aus den Didaktiken der Medizin, der Mathematik, der Physik und der Biologie sowie aus der pädagogisch-psychologischen Lehr-/Lernforschung und der Informatik soll ein Brückenschlag zwischen der schulbezogenen Forschung und der Forschung im medizinischen Kontext unternommen werden. Strategisch bringt die Arbeitsgruppe eine DFG-Forschergruppe, ein DFG-Netzwerk sowie einen vom BMBF-geförderten Verbund zur Thematik der Diagnosekompetenzen und ihrer Förderung zusammen. Die Diagnosesituationen stammen aus der Medizin, der Psychologie sowie verschiedenen Lehramtsfächern.

Diagnostizieren wird in der Arbeitsgruppe verstanden als das zielgerichtete Sammeln und Integrieren von Informationen zur Unsicherheitsreduktion zum jeweils vorliegenden Fall beim Treffen medizinischer bzw. pädagogischer Entscheidungen. Bislang ist unklar, wie Simulationen gestaltet werden können, um diagnostische Kompetenzen effektiv zu fördern und inwiefern individuelle Lernvoraussetzungen und Lernprozesse sowie die Domäne des Diagnostizierens diese Effekte moderieren. Diese Aspekte sollen in der vorgeschlagenen Arbeitsgruppe thematisiert werden.

Im ersten Beitrag werden zwei Metaanalysen vorgestellt. In der ersten steht die Förderung von Diagnosekompetenzen mit Intervention im Mittelpunkt. In einer zweiten wird untersucht wie effektive Simulationen zur Förderung von komplexen Fähigkeiten gestaltet sind.

Der zweite und dritte Beitrag der Arbeitsgruppe fokussieren die in der DFG-Forschergruppe Cosima entwickelten Simulationen aus sieben Teilprojekten. In beiden Beiträgen werden Ergebnisse aus Validierungsstudien vorgestellt.

Im zweiten Beitrag stehen die Diagnosequalität und der Diagnoseprozess im Vordergrund. Die Diagnosequalität wird über Akkuratheit, Effizienz und Effektivität der Diagnose operationalisiert. Der Diagnoseprozess wird mittels acht epistemisch-diagnostischen Aktivitäten (Fischer et al., 2014) beschrieben. Es wird analysiert inwiefern die verwendeten Instrumente und Simulationen die Diagnosequalität und den Diagnoseprozess valide erfassen konnten und inwiefern der Diagnoseprozess und die Diagnosequalität mit dem Professionswissen zusammenhängen.

Der dritte Beitrag fokussiert die erlebte Authentizität in Simulationen in unterschiedlichen Domänen. Dabei wird analysiert inwiefern sich Lernende in die simulierten diagnostischen Situationen einfühlen können und wie dies mit dem Diagnosekompetenzerwerb zusammenhängt.

Im vierten Beitrag werden Simulationen zum Rückenschmerz (Medizin) und zu Lernstörungen (Psychologie) aus dem BMBF-Verbundprojekt FAMULUS vorgestellt. Der Fokus des Beitrags ist die automatisierte Analyse der Interaktion mit der Simulation und darauf aufbauendes adaptives Feedback zur Förderung der Diagnosekompetenzen.

*Beiträge des Symposiums*

### Förderung von Diagnosekompetenzen Simulationen: Zwei Metaanalysen

**Nicole Heitzmann, Venance Timothy, Frank Fischer**

LMU München, Deutschland

Diagnosekompetenzen sind wichtige Kompetenzen für das spätere berufliche Handeln. Bei der instruktionalen Förderung von Diagnosekompetenzen wurden jedoch deutliche Hinweise auf bestehende Defizite identifiziert. Als wichtige Ursache für den bestehenden Förderbedarf wird gesehen, dass das Hochschulstudium zwar die Aneignung der deklarativen Wissensbasis unterstützt, bislang aber oft zu wenig Möglichkeiten für eigenes Handeln von Studierenden in praktischen Diagnosesituationen bereithält. Das eigene Handeln wird neben dem Erwerb deklarativen Wissens als notwendige Voraussetzung für den Diagnosekompetenzerwerb betrachtet (vgl. Anderson & Lebiere, 1998; z.B. Okuda et al., 2009; Schank, Berman, & Macpherson, 1999; VanLehn, 1996). Damit das Handeln in Diagnosesituationen geübt werden kann, ist es aus ökonomischen, pragmatischen sowie teilweise auch aus ethischen Gründen erforderlich, berufliche Situationen mit Blick auf den Zweck der Diagnose zu simulieren (Ziv, Wolpe, Small, & Glick, 2003). Jedoch ist unklar wie diese Simulationen gestaltet sein sollten um Diagnosekompetenzen bestmöglich zu fördern und inwiefern der Kontext diese Effekte moderiert. Metaanalytisch lassen sich diese Fragen aktuell nicht in einer Metaanalyse beantworten, da die Datengrundlage zur Förderung von Diagnosekompetenzen mit Simulationen nicht ausreichend ist. Darum werden im folgenden zwei Metaanalysen vorgestellt, die jeweils einen Teilbereich der Fragen abdecken.

In der ersten Metaanalyse liegt der Schwerpunkt auf der Förderung von Diagnosekompetenzen mit Interventionen. Obwohl es bereits einige Interventionsstudien zur Förderung von Diagnosekompetenzen gibt, zeigten sich bislang unklare Effekte. Eine offene Frage ist darum, inwiefern Diagnosekompetenzen durch zusätzliche instruktionale Maßnahmen gefördert werden können und welche Faktoren (Instruktionsansatz, Vorwissen, Dauer der Intervention, kooperatives Lernen vs individuelles Lernen, Domäne) darauf Einfluss nehmen.

Dazu wurde eine Metaanalyse mit 22 Primärstudien (45 Effekte) durchgeführt. Die Daten wurden mit R ‚metafor‘ & ‚robumeta‘ analysiert. Es wurde ein Random-effects model und eine Meta-Regression verwendet.

Die Ergebnisse zeigten einen kleinen Gesamteffekt ( $g = 0.38$ ,  $p = 0.002$ , 95% CI von 0.16 bis 0.61), eine Gesamtheterogenität von  $\tau^2 = 0.216$  und eine Variabilität von 76,2 %. Der instruktionale Ansatz Problemlösen zeigte sich in einer Moderationsanalyse beispielbasiertem Lernen als überlegen ( $g = 0.09$ ,  $p = 0.0097$ , CI: [0.144, 0.893]). Die anderen untersuchten Moderatoren (Dauer der Intervention, Vorwissen, kooperatives Lernen vs individuelles Lernen, Domäne) zeigten keine signifikanten Effekte. Es konnte auch kein Interaktionseffekt zwischen dem Vorwissen und dem verwendeten Instruktionsansatz gezeigt werden.

Diagnosekompetenzen scheinen demnach durch Interventionen gefördert werden können, jedoch sind die Effekte klein. Problembasiertes Lernen scheint dabei dem beispielbasierten Lernen überlegen, auch bei Lernern mit geringem Vorwissen. Dies erscheint zunächst im Widerspruch zu aktuellen Theorien des beispielbasierten Lernens (z.B. Renkl, 2014). Möglicherweise ist der Effekt von beispielbasiertem Lernen in schlecht strukturierten Domänen eingeschränkt. Ein besseres Verständnis über die Wirkmechanismen von Interventionen zur Förderung von Diagnosekompetenzen erscheint notwendig um die bisher eher kleinen Effekte zu vergrößern. Dazu wird aktuell eine erweiterte Analyse von weiteren Moderatoren wie z.B. weiteren Instruktionsansätzen, der Messmethode von Diagnosekompetenzen und kontextuellen Variablen durchgeführt. Diese erweiterte Metaanalyse wird bis Februar 2018 abgeschlossen sein.

In einer zweiten Metaanalyse wird aktuell untersucht inwiefern komplexe Fähigkeiten mit Simulationen gefördert werden können und welche Gestaltungsmerkmale dabei eine Rolle spielen. In aktuellen Metaanalysen zeigten sich deutliche Hinweise über die Wirksamkeit von Simulationen verglichen mit einer Kontrollgruppe die ohne Simulationen lernte (z.B. Cook, 2014). Unklar ist jedoch welche Gestaltungsmerkmale zur Effektivität von Simulationen beitragen. Dazu sollen die Lernvoraussetzungen, der verwendete Instruktionsansatz sowie die Kontextvariablen genauer analysiert werden. Diese zweite Metaanalyse wird aktuell durchgeführt. Erste Ergebnisse sind bis zum Symposium im Februar 2018 zu erwarten.

### **Erfassung des Diagnoseprozesses und der Diagnosequalität im Rahmen von simulationsbasierter Lernumgebungen**

**Angelika Wildgans<sup>1</sup>, Maximilian Christian Fink<sup>2</sup>, Amadeus Jonathan Pickal<sup>3</sup>, Constanze Weber<sup>1</sup>, DFG-Forscherguppe Cosima FOR 2385<sup>4</sup>**

<sup>1</sup>Technische Universität München, <sup>2</sup>Klinikum der Universität München, <sup>3</sup>LMU München, <sup>4</sup>LMU München, Klinikum LMU, TU München

Für eine adaptive Unterrichtsgestaltung und individuelle Kompetenzförderung von Lernenden sowie für eine adäquate Therapie von PatientInnen sind Diagnosekompetenzen im Kontext von Schule und Medizin unerlässlich (Graber, Wachter & Cassel, 2012; Südkamp & Praetorius, 2017). Unter „Diagnose“ wird in diesem Zusammenhang „das zielgerichtete Sammeln und Integrieren von Informationen zum Treffen medizinischer bzw. pädagogischer Entscheidungen“ (Fischer et al., 2016, S. 4) verstanden. Aufgrund ihrer großen Bedeutung sollten Diagnosekompetenzen, welche als kognitive, kontextspezifische, und veränderbare Leistungsdispositionen verstanden werden können, einen zentralen Baustein der universitären Ausbildung darstellen (Helmke & Weinert, 2015; KMK, 2004; Koeppen, Hartig, Klieme & Leutner, 2008). Trotzdem zeigt sich, dass eine Förderung von Diagnosekompetenzen bislang – insbesondere in der ersten Phase der Ausbildung von Lehramtsstudierenden – vernachlässigt wurde (Oser & Oelkers, 2001). Auch im Medizinstudium gibt es an dieser Stelle Aufholbedarf; Diagnosekompetenzen werden vor allem in Praxisphasen implizit von ÄrztInnen an Studierende vermittelt (Graber et al., 2012).

Besonders geeignet für die Förderung von Diagnosekompetenzen sind Simulationen (Cook et al., 2012). Diagnosesituationen können in Simulationen realistisch gestaltet, beliebig oft wiederholt und zu Beginn vereinfacht werden. Außerdem kann die Vermittlung von Diagnoseprozessen nach einem standardisierten Muster erfolgen. Für die Förderung der Diagnosekompetenz ist es sinnvoll sowohl den Diagnoseprozess als auch die Diagnosequalität zu betrachten. Die Qualität einer Diagnose zeigt sich nach dem Verständnis dieser Forschergruppe in ihrer Akkuratheit, Effizienz und Effektivität.

Da gerade die Beobachtung des Diagnoseprozesses in der bisherigen Forschung in beiden Domänen vernachlässigt wurde (Fischer et al., 2016), wird ein Modell zu wissenschaftlichem Schlussfolgern und Argumentieren von Fischer et al. (2014) auf den Diagnoseprozess im Bereich der Medizin und des Lehramts übertragen. Der Diagnoseprozess setzt sich demnach aus acht epistemisch-diagnostischen Aktivitäten (EDAs) zusammen: 1) Problemidentifikation 2) Fragen stellen 3) Hypothesen generieren 4) Artefakte konstruieren 5) Evidenzen generieren 6) Evidenzen evaluieren 7) Schlussfolgerungen ziehen und 8) Prozess- bzw. Ergebnisse kommunizieren. Häufigkeit und Reihenfolge der Aktivitäten können in Abhängigkeit von der Diagnosesituation variieren.

Die Forschergruppe analysiert und fördert Diagnoseprozesse und Diagnosequalität von Studierenden mithilfe simulationsbasierter Lernumgebungen. In den ersten Teilstudien werden dabei folgende Fragestellungen bearbeitet:

- 1) Welche Möglichkeiten eignen sich zur Beobachtung und Kategorisierung der EDAs?
- 2) Inwiefern sind die Messinstrumente zur Erfassung des Diagnoseprozesses und der Diagnosequalität reliabel und valide?
- 3) Inwiefern hängen Professionswissen, Diagnoseprozess und Diagnosequalität zusammen?

In den sieben Teilprojekten werden die entwickelten Simulationen in Pilotstudien mit insgesamt 180 Versuchspersonen eingesetzt. Die Messung des Diagnoseprozesses erfolgt anhand von Beobachtungen, Protokollen lauten Denkens und Logdatenerhebungen. In den Teilprojekten werden aus digitalen Notizen, kollaborativen Handlungen und gestellten Anamnesefragen die einzelnen EDAs herausgearbeitet.

Die Diagnosequalität wird anhand der drei bereits genannten Aspekte erfasst. Nach Fischer et al. (2016) wird Akkuratheit gemessen als eine Übereinstimmung der Diagnose mit dem unabhängig davon eingeschätzten realen Profil der zu diagnostizierenden Person. Die Effizienz und die Effektivität werden anhand der Erstellung der Diagnose und deren Aufwand beurteilt. Darüber hinaus werden mithilfe von Professionswissenstests Zusammenhänge des Wissensstandes mit Diagnoseprozess und Diagnosequalität überprüft.

Ausgewählte Ergebnisse aus den Pilotstudien mehrerer Teilprojekte werden vorgestellt. Beispielsweise wird anhand eines medizinischen Teilprojekts erläutert, wie in einer Simulation zum Anamnesegespräch aus gestellten Fragen auf die EDAs Hypothesengenerierung, Evidenzgenerierung und Evidenzevaluation geschlossen wird. In einem anderen Teilprojekt der Mathematikdidaktik verfassen Studierende während der Auseinandersetzung mit Schülerlösungen digitale Notizen. Diese Notizen dienen als Grundlage, um die im Diagnoseprozess vorkommende EDAs zu ermitteln. Darüber hinaus werden in Bezug auf die Diagnosequalität erste Ergebnisse zur Akkuratheit und Effektivität vorgestellt.

### **Validierung von simulationsbasierten Lernumgebungen zur Förderung von Diagnosekompetenzen. Die Wahrnehmung von Presence in simulierten Diagnosesituationen in verschiedenen Domänen**

**Anika Radkowsch<sup>1</sup>, Elias Codreanu<sup>2</sup>, Frank Fischer<sup>1</sup>, Christian Förtsch<sup>1</sup>, Birgit Neuhaus<sup>1</sup>, Julia Schauburger<sup>1</sup>, Tina Seidel<sup>2</sup>, Kathleen Stürmer<sup>3</sup>, DFG-Forscherguppe Cosima FOR 2385<sup>4</sup>**

<sup>1</sup>LMU München, Deutschland, <sup>2</sup>TU München, <sup>3</sup>Universität Tübingen, <sup>4</sup>LMU München, Klinikum LMU, TU München

Theoretischer Hintergrund

Die Forschergruppe COSIMA hat sich zum Ziel gesetzt, die Diagnosekompetenzen von angehenden Lehrkräften und Mediziner\*innen mithilfe von simulationsbasierten Lernumgebungen zu analysieren und zu fördern. Die Lernumgebungen enthalten Videos und Dokumente, die verschiedene Herausforderungen in Diagnosesituationen darstellen. Die Lernenden können in die jeweilige Situation eingreifen und damit das weitere Geschehen beeinflussen (Fischer et al., 2014). In insgesamt 25 experimentellen Studien wird die instruktionale Gestaltung der Lernumgebungen systematisch variiert. Zur Vorbereitung dieser Studien werden zunächst in allen sieben Teilprojekten simulationsbasierte Lernumgebungen entwickelt und validiert. In diesem Beitrag werden Fragestellungen, Methodik und Hauptergebnisse dieser Phase berichtet. Um eine hohe Qualität der Lernumgebungen zu gewährleisten, werden bereits bei der Entwicklung Erkenntnisse aus der Forschung und Expertenurteile mit einbezogen. Bei simulationsbasierten Lernumgebungen wird außerdem das wahrgenommene Eintauchen (Presence) als wichtiges Kriterium für ihre Qualität hervorgehoben (z.B. Scoresby & Shelton, 2011). Nach Schubert, Friedmann und Regenbrecht (2001) handelt es sich bei Presence um eine subjektive Empfindung, die durch das Gefühl in der virtuellen Umgebung zu sein (räumliche Präsenz), durch eine kognitive Fokussierung auf die Lernumgebung (Involviertheit) sowie die wahrgenommene Authentizität der Lernumgebung definiert wird. Obwohl die Wichtigkeit von Presence häufig betont wurde, sind bisherige Forschungsarbeiten über den Einfluss von Presence auf die Lernleistung uneindeutig (z.B. Winn, Windschitl, Fruland & Lee, 2002; Witmer & Singer, 1998). Jedoch scheint die wahrgenommene Presence einer Lernumgebung ausschlaggebend zu sein, damit in der Simulation gezeigtes Verhalten in reale Diagnosesituationen transferiert werden kann (Frank, 2015). Das gilt insbesondere für das Subkonstrukt Authentizität, da eine hohe Authentizität eine valide Messung und Förderung derjenigen Kompetenzen gewährleistet, die in späteren beruflichen Anforderungssituationen benötigt werden (Stürmer & Seidel, 2015). Daher ist ein erstes Ziel zu evaluieren, ob die Versuchsteilnehmer im Sinne von Presence das Gefühl haben, in die entwickelten Simulationsumgebungen einzutauchen. Ein weiteres Ziel ist es, den Zusammenhang zwischen Presence und der Diagnoseakkuratheit als ein Maß für erfolgreiche professionelle Performanz zu erfassen.

#### Fragestellungen

Zur Validierung der simulationsbasierten Lernumgebung werden folgende Fragestellungen untersucht:

- (1) Können die entwickelten simulationsbasierten Lernumgebungen den Lernenden ein Gefühl von Presence vermitteln?
- (2) Inwiefern besteht ein Zusammenhang zwischen der wahrgenommenen Presence und der Diagnoseakkuratheit?

#### Methode

Derzeit werden in der DFG-Forschergruppe sieben Fallsimulationen zur Analyse und Förderung von Diagnosekompetenzen entwickelt. Fünf Lernumgebungen sind im MINT-Bereich des schulischen Kontextes angesiedelt und fördern Kompetenzen der Individualdiagnostik und der Unterrichtsdiagnose. Zwei weitere Projekte fokussieren individuelle und kooperative Diagnosekompetenzen in der inneren und radiologischen Medizin. Bis Ende des Jahres werden in den Projekten Pilotierungsdaten zur Validierung der einzelnen Lernumgebungen erhoben. Insgesamt umfasst die angestrebte Stichprobe N=180 Experten und Studierende des Lehramts und der Medizin. Das Presence-Gefühl wird jeweils im direkten Anschluss an die Fallbearbeitung in der Lernumgebung anhand einer fünfstufigen Likertskala (1 trifft nicht zu, 5 trifft zu) gemessen. Die Probanden schätzen ihre wahrgenommene Presence anhand der drei Subskalen räumliche Präsenz (5 Items), Involviertheit (4 Items) und Authentizität (3 Items) ein (Schubert et al., 2001). Für die Beantwortung der Fragestellung (1) werden deskriptive Skalenwerte der einzelnen Subskalen berechnet.

Für jede simulationsbasierte Lernumgebung wird der Zusammenhang zwischen Presence und Diagnoseakkuratheit regressionsanalytisch bestimmt. Die Diagnoseakkuratheit wird unter anderem durch Vergleiche mit Expertenlösungen projektspezifisch erfasst. Dieses Verfahren wird für die Subskalen räumliche Presence, Involviertheit und Authentizität durchgeführt. Explorativ werden außerdem Unterschiede in der wahrgenommenen Presence zwischen den simulationsbasierten Lernumgebungen anhand verschiedener Kriterien (z.B. first- versus third-person Lerner; Domänen) analysiert.

#### Erwartete Ergebnisse

Trotz der Unterschiede in der Gestaltung der einzelnen Simulationsumgebungen wird erwartet, dass die Probanden ein Gefühl von Presence in allen Simulationsumgebungen empfinden. Dabei erwarten wir hohe Werte (Mittelwert > 3) aller drei Subskalen. Des Weiteren erwarten wir positive Effekte der wahrgenommenen Presence auf die Diagnoseakkuratheit. Es werden Ergebnisse unterschiedlicher Teilprojekte vorgestellt und exemplarisch näher betrachtet.

### **Automatische Analyse von Diagnosekompetenzen in Fallsimulationen**

**Claudia Schulz<sup>1</sup>, Michael Sailer<sup>2</sup>, Jan Kiesewetter<sup>3</sup>, Christian M. Meyer<sup>1</sup>, Elisabeth Bauer<sup>2</sup>, Iryna Gurevych<sup>1</sup>, Martin Fischer<sup>3</sup>, Frank Fischer<sup>2</sup>**

<sup>1</sup>TU Darmstadt, <sup>2</sup>LMU München, Deutschland, <sup>3</sup>Klinikum der Universität München

Diagnosekompetenzen spielen in vielen beruflichen Situationen eine wichtige Rolle. Diagnostizieren ist das zielgerichtete Sammeln und Integrieren von Informationen zum Zweck der bestmöglichen Entscheidungsfindung. Um Diagnosekompetenzen von Studierenden zu fördern und diese dadurch hinreichend auf ihre beruflichen Aufgaben vorzubereiten, bieten sich Fallsimulationen an. Dabei werden Studierenden berufsnahe Situationen geschildert, in denen diagnostisches Denken und Schlussfolgern erforderlich ist. Ohne hinreichendes inhaltliches Feedback ist es Studierenden jedoch nur schwer möglich ihre Diagnosekompetenzen zu verbessern. Klassische Methoden des Feedbacks, insbesondere für große Gruppen an Studierenden, sind entweder wenig Lehrkraft-intensiv jedoch allgemein (z.B. Muster- bzw. Expertenlösungen) oder adaptiv jedoch Lehrkraft-intensiv (z.B. individuelles Feedback von Experten).

Automatisches adaptives Feedback bezüglich Diagnosekompetenzen kann helfen die Nachteile klassischer Feedback Methoden zu überwinden. Ziel ist es, Methoden zu entwickeln, die Diagnosekompetenzen automatisch analysieren und evaluieren, und daher zur automatischen Erstellung von adaptivem Feedback genutzt werden können. Methoden aus der Künstlichen Intelligenz (KI) wurden bereits erfolgreich in Lehr-Lernumgebungen eingesetzt und haben hohes Zukunftspotential (Luckin, Holmes, Griffiths & Forcier, 2016), da sie sich zur automatischen Analyse von Daten eignen. Die hier untersuchte Forschungsfrage ist, inwieweit Methoden aus der KI auf Studierendenlösungen in Fallsimulationen anwendbar sind, um diese automatisch bezüglich Diagnosekompetenzen zu analysieren.

Um Diagnosekompetenzen zu untersuchen, bieten sich als Falllösungen diagnostische Essays an. Studierende schreiben dabei ihr Diagnosevorgehen während der Fallsimulation detailliert auf. Zu erwartende epistemisch-diagnostische Elemente sind u.a. (Fischer et al., 2014):

- a) Hypothesen generieren,

- b) Evidenz generieren,
- c) Evidenz evaluieren, und
- d) Schlussfolgerungen ziehen.

Um epistemisch-diagnostischen Elemente automatisch in Falllösungen zu erkennen, werden Methoden aus der KI zur automatischen Sprachanalyse verwendet. Diese Methoden lernen anhand von Beispielen, wie eine bestimmte Aufgabe zu bearbeiten ist. Die zu lernende Aufgabe ist hier das Erkennen von epistemisch-diagnostischen Elementen in einer Falllösung. Dazu werden als Beispiele diagnostische Essays verwendet, die von Studierenden während einer Fallsimulationsstudie geschrieben wurden. Damit die KI Methode lernen kann, was epistemisch-diagnostische Elemente sind, müssen diese Elemente von fachspezifischen Lehrkräften in den Beispiel-Essays annotiert werden.

Die hier gewählte Methode aus der KI ist ein Deep Learning neuronales Netzwerk namens LSTM (Long Short-Term Memory) (Hochreiter & Schmidhuber, 1997). Ein LSTM liest einen Text, ähnlich wie Menschen, Wort für Wort, hat jedoch ein „Gedächtnis“, in dem der bereits gelesene Kontext (die vorherigen Worte) gespeichert ist. Dies ermöglicht es dem LSTM, die Bedeutung eines jeden Wortes im Zusammenhang des ganzen Textes zu verstehen. Während der Lernphase stehen dem LSTM die annotierten epistemisch-diagnostischen Elemente in den Beispiel-Essays zur Verfügung. Es lernt dadurch den Zusammenhang zwischen Wörtern (im Kontext anderer „gemerkter“ Wörter) und epistemisch-diagnostischen Elemente. Nachdem es mehrere hundert beispielhafte Falllösungen gelesen hat, kann das LSTM so epistemisch-diagnostische Elemente in neuen Falllösungen erkennen.

Je mehr Beispiel-Essays ein LSTM während der Lernphase liest, desto besser kann es danach epistemisch-diagnostische Elemente in neuen Falllösungen erkennen. Jedoch bedarf die Erstellung von Beispiel-Essays eines hohen Aufwands, da diese von Fachexperten annotiert werden müssen. In dieser Arbeit werden LSTMs daher mit einem weiteren Ansatz aus der KI kombiniert, dem sogenannten Transfer Learning. Dieser Ansatz basiert auf der Idee, dass ähnliche Aufgaben ähnlich gelöst werden. Zum Beispiel ist das Erkennen von epistemisch-diagnostischen Elementen ähnlichen dem Erkennen von Argumentationsstrukturen. Es gibt bereits hunderte Texte, in denen die Argumentationsstruktur annotiert ist. Ein LSTM kann anhand dieser Beispiele also lernen, Argumentationsstrukturen in neuen Texten zu erkennen. Wie zuvor beschrieben, erlernt das LSTM den Zusammenhang zwischen Wörtern und Argumentationsstruktur. Unter der Annahme, dass dieser Zusammenhang ähnlich dem für diagnostisch-epistemische Elemente ist, kann das LSTM das erlernte Zusammenhangswissen zu Argumentationsstrukturen auf diagnostisch-epistemische Elemente „transferieren“. Dadurch benötigt das LSTM weniger beispielhafte Falllösungen, da es bereits „vorgelernt“ hat.

Zum Zeitpunkt der Tagung liegt eine erste Auswertung des beschriebenen LSTMs bezüglich des Erkennens von epistemisch-diagnostischen Elementen in Falllösungen vor.



## G20

Zeit: Samstag, 17.02.2018: 14:45 - 16:30 · Ort: Kollegienhaus H120

### Prozessdatenanalyse als Zugang zur Untersuchung von Geschlechterdisparitäten

Chair(s): **Carolin Hahnel** (Deutsches Institut für internationale Pädagogische Forschung, Zentrum für internationale Vergleichsstudien)

DiskutantIn(nen): **Ramona Lorenz** (Technische Universität Dortmund, Institut für Schulentwicklungsforschung)

Untersuchungen zu Geschlechterunterschieden gehen häufig mit einer heterogenen Ergebnislage einher und lösen zum Teil heftige Kontroversen aus. Obwohl Unterschiede zwischen Jungen und Mädchen beispielsweise für verbale Fähigkeiten und mathematische Leistungen verhältnismäßig gering ausfallen (Hyde, 2014), werden sie – und häufig auch Geschlechtsstereotype – herangezogen um Unterschiede im beruflichen Erfolg von Männern und Frauen zu erklären. Jüngere Befunde bestätigen, dass Jungen und Mädchen sich eher geringfügig in zentralen Kompetenzbereichen wie Lesen, Mathematik und Naturwissenschaften unterscheiden, zeigen aber auch, dass Jungen überdurchschnittlich häufig Positionen am oberen und unteren Leistungsspektrum besetzen (Bergold et al., 2017). Als Folge der indifferenten Befundlage werden Differenzen auch vermehrt auf die methodische Ebene zurückgeführt (vgl. Hatlevik et al., 2017). Festzuhalten ist, dass die Erforschung von geschlechtsspezifischen Unterschieden wichtig ist – einerseits um bestehende Differenzen und ihre Veränderung zu dokumentieren, andererseits um Konsequenzen für die Gestaltung von Testverfahren zu ziehen, damit Personen einer bestimmten Gruppe gegenüber einer anderen nicht bevorzugt bzw. benachteiligt werden (Hartig et al., 2012).

Durch Erhebungsmethoden des computerbasierten Testens werden Prozessdaten zugänglich, deren Auswertung als zusätzlicher objektiver Zugang zur Kompetenzerfassung diskutiert wird (Keller et al., 2016). Allgemein beschreiben Prozessdaten Verhaltensdaten, die während des Testverlaufs gesammelt werden. Dazu zählen unter anderem Logdaten, die die Interaktionen von Probanden mit einer Testumgebung erfassen, und Zeitdaten, die Auskunft darüber geben, wie lange eine Person eine bestimmte Aufgabe bearbeitet hat. In diesem Symposium gehen wir der gemeinsamen Fragestellung nach, inwiefern Geschlechterunterschiede anhand von Indikatoren der Test- und Aufgabebearbeitung in Kompetenztests identifiziert und erklärt werden können. Dabei sollen für die Bildungswissenschaften relevanten Aspekte der aktuellen Genderdebatte mit dem Erkenntnisgewinn vereint werden, der durch die systematische Auswertung von Prozessdaten erlaubt wird.

Im ersten Beitrag betrachten Riedel und Hahnel Unterschiede im grundlegenden Umgang von Schülerinnen und Schülern mit Computeroberflächen – einer Teilfähigkeit der kompetenten Nutzung von Informations- und Kommunikationstechnologien. Sie untersuchen hierbei die Frage, ob spezifische Bearbeitungssequenzen in den Aufgaben identifiziert werden können, die vorliegende Geschlechterdifferenzen erklären können. Im zweiten Beitrag des Symposiums präsentieren Eichmann und ihre Kollegen Befunde zu Geschlechterdisparitäten im Bereich des komplexen Problemlösens. Sie konzentrieren sich dabei auf die Frage, ob sich geschlechtsspezifische Unterschiede im komplexen Problemlösen auf das Bearbeitungsverhalten zurückführen lassen, das Schülerinnen und Schüler beim Explorieren eines Problems zeigen. In letzten Beitrag widmen sich Thums und Wolter den Lesezeiten, die Erwachsene für das Lesen von literarischen Texten und Sachtexten aufwenden. Die Autorinnen gehen hierbei der Frage nach, ob sich bekannte Unterschiede in der Lesegeschwindigkeit zwischen Männern und Frauen auch für Lesezeiten in Kompetenztests replizieren lassen.

#### Beiträge des Symposiums

### Erklärung geschlechtsspezifischer Unterschiede in Aufgaben zur Erfassung basaler Computerfähigkeiten durch die Analyse von Bearbeitungssequenzen

**Carolin Riedel<sup>1</sup>, Carolin Hahnel<sup>2</sup>**

<sup>1</sup>Universität Kassel, <sup>2</sup>Deutsches Institut für internationale Pädagogische Forschung, Zentrum für internationale Vergleichsstudien

Geschlechtsspezifische Unterschiede werden in vielen bildungswissenschaftlichen Themenfeldern beforscht und diskutiert. Unter anderen werden Genderdiskurse häufig in Bezug auf das Konstrukt der ICT (Information and Communication Technology) Literacy geführt, das Fähigkeiten zur kompetenten Anwendung und Nutzung von Informations- und Kommunikationstechnologien wie Computern, Tablets und Smartphones abbildet (Goldhammer, Naumann & Keßel, 2012). Studien zu geschlechtsspezifischen Unterschieden in ICT Literacy erklärten Differenzen häufig anhand von (fehlenden) Zugangsmöglichkeiten zum Medium und konzentrierten sich auf Aspekte von Einstellungen und Interessen (u.a. Durndell & Haag, 2002; Morris & Trushell, 2014; Vollmeyer & Imhof, 2007). Laut heutigem Kenntnisstand lassen sich zumindest zu erstgenanntem Punkt kaum noch Differenzen verzeichnen. So verfügte laut JIM-Studie (Jugend, Information, (Multi-) Media) 2016 nahezu jeder deutsche Haushalt (98%) über eine entsprechende Ausstattung (Feierabend, Plankenhorn & Rathgeb, 2016). Die Aussagekraft von Unterschieden in ICT Einstellungen und Selbstwirksamkeitserwartungen zur Erklärung von Geschlechterdisparitäten weist aber ebenfalls Schwächen auf (Hatlevik, Scherer & Christophersen, 2017). Zentrale Ergebnisse der ICILS-Studie (International Computer and Information Literacy Study) 2013 – einer großangelegten, internationalen Vergleichsstudie zu Computer- und Informationskompetenzen von Schülerinnen und Schüler der 8. Jahrgangsstufe – zeigten beispielsweise einen signifikanten Leistungsvorsprung für Mädchen gegenüber Jungen, obwohl Mädchen auch eine bedeutsam geringere Selbstwirksamkeitserwartung gegenüber ICTs attestiert wurde (Bos et al., 2014). Insgesamt betrachtet ist die Befundlage sehr heterogen wie auch Ergebnisse zum Beispiel von Goldhammer und Kollegen (2012) zeigen, die für grundlegende Fertigkeiten mit Computern eine bessere und schnellere Leistung für Jungen feststellten.

Der vorliegende Beitrag konzentriert sich auf die Erklärung von Geschlechterunterschieden in Aufgaben zur Erfassung basaler Computerfähigkeiten. Hierfür soll zunächst geprüft werden, ob Unterschiede in den Testwerten zwischen Jungen und Mädchen gefunden werden können. Anschließend sollen konkrete Aufgaben identifiziert werden, die unterschiedliche Lösungswahrscheinlichkeiten für Jungen und Mädchen aufweisen. Durch eine inhaltliche Analyse der Aufgaben und anhand von Logdaten, die den Bearbeitungsverlauf von Probanden in computerbasierten Tests abbilden, sollen vorliegende Geschlechterunterschiede erklärt werden.

Zur Untersuchung wurden die Daten von 618 15-jährigen Schülerinnen und Schülern betrachtet (48.9% weiblich), die 2012 an einer deutschen Zusatzstudie zur Einführung computerbasierter Assessments in PISA (Programme for International Student Assessment) teilnahmen. Die Probanden bearbeiteten 20 simulationsbasierte Items der Basic Computer Skill Scale (BCS; Goldhammer et al., 2012), die den grundlegenden Umgang mit Computeroberflächen beim Finden, Verwalten und Kommunizieren von Informationen erfasst (WLE Reliabilität = .82). Zur Analyse werden die dichotomen Ergebnisdaten (richtig

vs. falsch gelöst; Anteil korrekter Antworten pro Item zwischen 15.4% – 83.5%) sowie aus der Erhebung gewonnene Klickdaten auf Links, Buttons und Menüeinträge herangezogen.

Im Gegensatz zu den Befunden von Goldhammer und Kollegen (2012) ergab eine Regression der manifesten Geschlechtsvariable auf die latent rasch-modellierte BCS-Variable signifikant niedrigere Fähigkeitswerte für Jungen ( $b = -0.26$ ,  $p < .001$ ). Zur Identifikation von Items, die geschlechterspezifische Unterschiede aufweisen, wurden Differential Item Functioning Analysen (DIF) durchgeführt. Dabei wiesen insgesamt fünf Items DIF auf (vgl. Magis, Beland, Tuerlinckx & DeBoeck, 2010). Im Weiteren werden die Klickdaten der Schülerinnen und Schüler mit Hilfe von N-Grammen (d.h. Bearbeitungssequenzen; vgl. He & von Davier, 2015) analysiert. Die Ergebnisse dieser Analysen werden Gegenstand des Beitrags sein und hinsichtlich der aktuellen Genderdebatte kritisch diskutiert.

### **Die Erklärung von Geschlechterunterschieden beim komplexen Problemlösen durch Explorationsverhalten**

**Beate Eichmann<sup>1</sup>, Frank Goldhammer<sup>2</sup>, Samuel Greiff<sup>3</sup>, Liene Pucite<sup>4</sup>, Johannes Naumann<sup>4</sup>**

<sup>1</sup>Deutsches Institut für internationale Pädagogische Forschung, <sup>2</sup>Deutsches Institut für internationale Pädagogische Forschung, Zentrum für internationale Vergleichsstudien, <sup>3</sup>Universität Luxemburg, <sup>4</sup>Goethe Universität Frankfurt

Die Fähigkeit komplexe Probleme zu lösen ist eine wichtige Kompetenz in der heutigen Gesellschaft. Das Fortschreiten von Globalisierung und Technisierung konfrontiert den Einzelnen regelmäßig mit unbekanntem Situationen, die es zu meistern gilt (Fischer, Greiff & Funke, 2012). Wegen seiner praktischen Relevanz wurde das komplexe Problemlösen (KPL) 2012 in die fünfte Runde des Programme for International Student Assessment (PISA) aufgenommen. Die Ergebnisse von PISA 2012 zeigen Leistungsunterschiede beim KPL zwischen den Geschlechtern (OECD, 2014): Mädchen zeigten signifikant schlechtere Leistungen als Jungen. Im Durchschnitt betrug die Differenz sieben Punkte auf der PISA-Skala (MW=500; SD=100). Wir wollen die Frage beantworten, ob die Leistungsunterschiede beim KPL auf unterschiedliches Bearbeitungsverhalten der Schülerinnen und Schüler zurückgeführt werden kann.

Problemlösen kann als die Transformation eines gegebenen Zustandes in einen Zielzustand betrachtet werden, wenn zunächst keine offensichtliche Lösung verfügbar ist (Mayer & Wittrock, 2006). Es kann zwischen KPL und nicht-komplexem Problemlösen unterschieden werden (Greiff, Kretschmar & Leutner, 2014). Ein Problem gilt als komplex, wenn nicht alle nötigen Informationen gegeben, sondern durch Exploration des Problems gewonnen werden müssen. Daher ist Exploration grundlegend für eine erfolgreiche Lösung solcher Probleme. Auch empirische Studien deuten auf die Wichtigkeit von Exploration beim KPL hin. Je mehr Exploration die Versuchspersonen in einer Studie von Dormann und Frese (1994) zeigten, desto besser war ihre Leistung in einem anschließenden Software-Test. Bell und Kozlowski (2008) konnten zeigen, dass Personen, die während eines Trainings im Umgang mit komplexen Systemen zur Exploration angeregt wurden, eine bessere Leistung in einem anschließenden Test zeigten als Personen, die ein Training ohne diese Anregung erhielten. Naumann, Goldhammer, Rölke und Stelter (2014) fanden einen quadratischen Zusammenhang zwischen der Anzahl der durchgeführten Interaktionen (ADI) und der Leistung beim technologiebasierten Problemlösen, wobei die optimale ADI eine Standardabweichung über dem Durchschnitt lag. Naumann et al. (2014) interpretieren den positiven Zusammenhang von ADI und Leistung (links vom Optimum) als adäquates Problemlöseverhalten, während der negative Zusammenhang von ADI und Leistung (rechts vom Optimum) Desorientierung widerspiegelt.

Im vorliegenden Beitrag haben wir den Zusammenhang von Exploration und Leistung beim KPL genauer betrachtet, um die Frage zu beantworten, ob der Leistungsunterschied zwischen Mädchen und Jungen durch unterschiedliches Explorationsverhalten erklärt werden kann. Zu diesem Zweck haben wir zwei Indikatoren für Explorationsverhalten gebildet: Zum einen haben wir den von Naumann et al. (2014) vorgestellten Indikator ADI angewendet. Zum anderen wollten wir zwischen zielgerichtetem und explorativem Verhalten unterscheiden, indem wir die Anzahl der nicht-zielgerichteten Interaktionen (ANI) als zusätzlichen Indikator verwendet haben. Mit ANI werden im Gegensatz zu ADI nur Interaktionen gezählt, die für eine erfolgreiche Aufgabenbearbeitung nicht unmittelbar notwendig sind.

Für unsere Analysen benutzten wir Daten des computerbasierten Problemlöse-Assessments der deutschen Substichprobe aus PISA 2012. Wir nutzten die Daten aller Items, die der Definition des komplexen Problemlösens nach Greiff et al. (2014) entsprechen und die die nötigen Berechnungen zuließen. Dadurch erhielten wir eine Stichprobe von  $n=1325$  Schülerinnen und Schülern und 12 Items. Aus den Logdaten dieser Schülerinnen und Schüler wurden die beiden Verhaltensindikatoren ADI und ANI gebildet und Scores für richtige und falsche Antworten vergeben. Anschließend wurden mit generalisierten linearen Mischmodellen (GLMMs) die Zusammenhänge von Prozessindikatoren, Leistung und Geschlechterunterschieden überprüft.

In den Ergebnissen zeigen sich erwartungsgemäß ein signifikanter Geschlechterunterschied zugunsten der Jungen und ein positiver Effekt beider Prozessindikatoren auf die Leistung. In Kombination mit der ADI verkleinert sich der Geschlechtereffekt, bleibt jedoch signifikant. In Kombination mit der ANI verkleinert sich der Geschlechtereffekt ebenfalls und ist nicht länger signifikant. Die Ergebnisse zeigen, dass Exploration positiv prädiktiv für die Leistung beim KPL ist und den Leistungsunterschied zwischen Mädchen und Jungen erklären kann. Somit kann der Geschlechterunterschied tatsächlich auf direkt abbildbare Verhaltensunterschiede zurückgeführt werden: Jungen schneiden offenbar besser beim KPL ab, weil sie mehr explorieren als Mädchen.

### **Lesen Frauen schneller? Die Analyse von Geschlechtsunterschieden in Erstlesezeiten von Texten eines Lesekompetenztests anhand von Prozessdaten**

**Kathrin Thums, Wolter Ilka**

Leibniz-Institut für Bildungsverläufe

Die Relevanz der Lesegeschwindigkeit in der Erklärung von Unterschieden in der Lesekompetenz ist in vielen Studien empirisch aufgezeigt worden (z. B. Artelt, Schiefele & Schneider, 2001). Nach Rosebrock und Nix (2012) ist eine angemessen schnelle Lesegeschwindigkeit neben der exakten Dekodierfähigkeit von Wörtern und der Automatisierung der Dekodierprozesse ein Indikator der Leseflüssigkeit. Es konnte gezeigt werden, dass die Leseflüssigkeit wiederum für das Gesamtverständnis eines Textes von zentraler Bedeutung ist (z. B. Rosebrock & Nix, 2012; Garbe, Holle & Jesch, 2010). Empirische Studien zeigen bereits im Schulalter Geschlechtsunterschiede zugunsten von Mädchen in der Lesegeschwindigkeit (z. B. Wassenburg, de Koning, de Vries, Boonstra & van der Schoot, 2017), welche auch im Erwachsenenalter weiterhin bestehen (z. B. Hayes & Waller, 1994; Roivainen, 2011). In dieser Studie gehen wir der Frage nach, ob die bisherigen Befunde zu Geschlechtsunterschieden in der Lesegeschwindigkeit repliziert werden können, wenn wir die Information zur Lesegeschwindigkeit nicht aus einem expliziten Speed-Test ableiten, sondern die individuellen Erstlesezeiten von Lesetexten aus den Prozessdaten eines Lesekompetenztests (Power-Test) heranziehen. Zudem untersuchen wir vor dem Hintergrund von

empirischen Studien zu Textsorten-Präferenzen von Frauen und Männern (Charlton, Burbaum & Sutter, 2004; Philipp, 2011), ob sich Unterschiede in literarischen Texten und Sachtexten aufzeigen lassen.

Zunächst zeigte sich im Lesegeschwindigkeitstest (Zimmermann, Artelt & Weinert, 2014) in den Daten der Studierenden- und Erwachsenenkohorte des Nationalen Bildungspanels (NEPS, SC5: N=5842; SC6: N=10451; Blossfeld, Roßbach & von Maurice, 2011) erwartungskonforme Mittelwertsunterschiede in der Lesegeschwindigkeit zwischen Frauen und Männern. Anschließend wurden aus einer Entwicklungsstudie des NEPS für eine computerbasierte Lesekompetenzmessung mit N=830 Erwachsenen zwischen 19-73 Jahren (M=33.04; SD=12.53) die Erstlesezeiten in literarischen und Sach-Texten aus den Prozessdaten ausgelesen. Insgesamt wurden den Studienteilnehmenden in dieser Studie sechs Sachtexte und drei literarische Texte vorgelegt (vgl. NEPS Rahmenkonzeption zur Lesekompetenz; Gehrler, Zimmermann, Artelt & Weinert, 2013). Aus der Erstlesezeit wurde zunächst die Komplexität der Texte, in Form des Lesbarkeitsindex (LIX, Lenhard & Lenhard, 2017), sowie die Anzahl der Wörter des Textes, herauspartialisiert. Für die Analysen wurde zusätzlich die Lesekompetenz über alle Texte über eine 2-PL-Skalierung ermittelt (Birnbaum, 1968).

In Regressionsmodellen wurden Geschlechtsunterschiede in der Lesezeit der Texte, unter Kontrolle des Alters und der Lesekompetenz der Teilnehmenden, betrachtet. Entgegen unserer Erwartung zeigten sich keine Geschlechtsunterschiede in der Erstlesezeit in den literarischen Texten und in den Sachtexten. Erwartungskonform zeigte sich ein positiver Zusammenhang der Lesekompetenz mit der Erstlesezeit der Texte in dem Kompetenztest sowohl für die literarischen Texte ( $\beta = -.091$ ,  $p < .013$ ,  $d = .14$ ) als auch für die Sachtexte ( $\beta = -.074$ ,  $p < .040$ ,  $d = .17$ ). Es konnte somit gezeigt werden, dass anders als vermutet, keine Geschlechtsunterschiede in den Erstlesezeiten von Texten zu finden sind, jedoch die Lesekompetenz einen Einfluss auf die Erstlesezeiten hat. In der Diskussion wird ein Ausblick auf zukünftige Studien gegeben, in welchen auch Bearbeitungsmuster der im Anschluss an die Texte präsentierten Aufgaben aufgenommen werden. Die vorliegenden Ergebnisse sollen weiterführend der Testkonzeption dienen und hinsichtlich diagnostischer Implikationen diskutiert werden.